

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

REFICS: Assimilating Data-Driven Paradigms into Reverse Engineering and Hardware Assurance on Integrated Circuits

RONALD WILSON, HANGWEI LU, MENGDI ZHU, DOMENIC FORTE (SENIOR MEMBER, IEEE) and DAMON L. WOODARD (SENIOR MEMBER, IEEE)

Florida Institute for Cybersecurity Research (FICS), University of Florida, Gainesville, FL 32601, USA

Corresponding author: Ronald Wilson (E-mail: ronaldwilson@ufl.edu).

ABSTRACT Comprehensive hardware assurance approaches guaranteeing trust on Integrated Circuits (ICs) typically require the verification of the IC design layout and functionality through destructive Reverse Engineering (RE). It is a resource intensive process that will benefit greatly from the extensive integration of data-driven paradigms, especially in the imaging and image analysis phase. Although obvious, this uptake of data-driven approaches into RE-assisted hardware assurance is lagging due to the lack of massive amounts of high-quality labelled data. In this paper, a large-scale synthetic Scanning Electron Microscopy (SEM) dataset, REFICS, is introduced to address this issue. The dataset, the first open-source dataset in the RE community, consists of 800,000 SEM images over two node technologies, 32nm and 90nm, and four cardinal layers of the IC, namely, doping, polysilicon, contact and metal layers. Furthermore, a framework, based on uncertainty and risk, is introduced to compare the efficacy and benefits of existing RE workflows utilizing ad-hoc steps in its execution. These developments are critical in developing RE-assisted hardware assurance into a scalable, automated and fault-tolerant approach. Finally, the work is concluded with the performance analysis of existing machine learning and deep learning approaches for image analysis in RE and hardware assurance.

INDEX TERMS Computer Vision, Dataset, Deep Learning, Hardware Assurance, Image Processing, Integrated Circuits, Machine Learning, Reverse Engineering, Scanning Electron Microscopy

I. INTRODUCTION

In the age of the Internet-of-Things (IoT), finding a product that doesn't incorporate an Integrated Circuit (IC) into its design and functionality is an extremely challenging task. ICs are semiconductor devices that convert a bunch of input signals into useful output signals. Being mass-produced, these devices are affordable and well-utilized in products ranging from low-cost IoT devices to high-performance computing clusters. Due to their ubiquity, they are exposed to almost all the data that flows through the internet. With the nature of data ranging from trivial pleasantries to personal and highly sensitive information, its accidental exposure due to faulty/compromised ICs can have severe consequences in the real world. Apart from faulty hardware design that leads to compromised data, there are also flaws that are introduced in the design, by adversaries, to compromise the design and, consequently, the data or the functionality of the IC at will. These malicious modifications made to the source design are called hardware Trojans. Hardware assurance approaches

ensure trust in these devices by ensuring that there are no malicious modifications installed on the IC.

As shown in Figure 1, the likely culprit behind compromised ICs are the use of third party services in the manufacturing workflow. Although these issues can be resolved by moving the manufacturing process to an in-house facility, the cost associated with the process is usually debilitating for IC designers, especially for small-scale designers. Similarly, usage of third party intellectual properties (IP) in the IC design also introduces a potential source of vulnerability. Hence, hardware assurance measures are critical in ensuring trust in the devices. Existing techniques for trust and assurance in these scenarios are limited and ineffective [1]–[4]. For example, run-time monitoring on the ICs increases the resource requirements such as power consumption, memory utilization, and area overhead on ICs due to on-chip sensors used to detect anomalous activities. Furthermore, the hardware used by run-time methods may also contain hardware Trojans. In test time methods, generating test vectors that

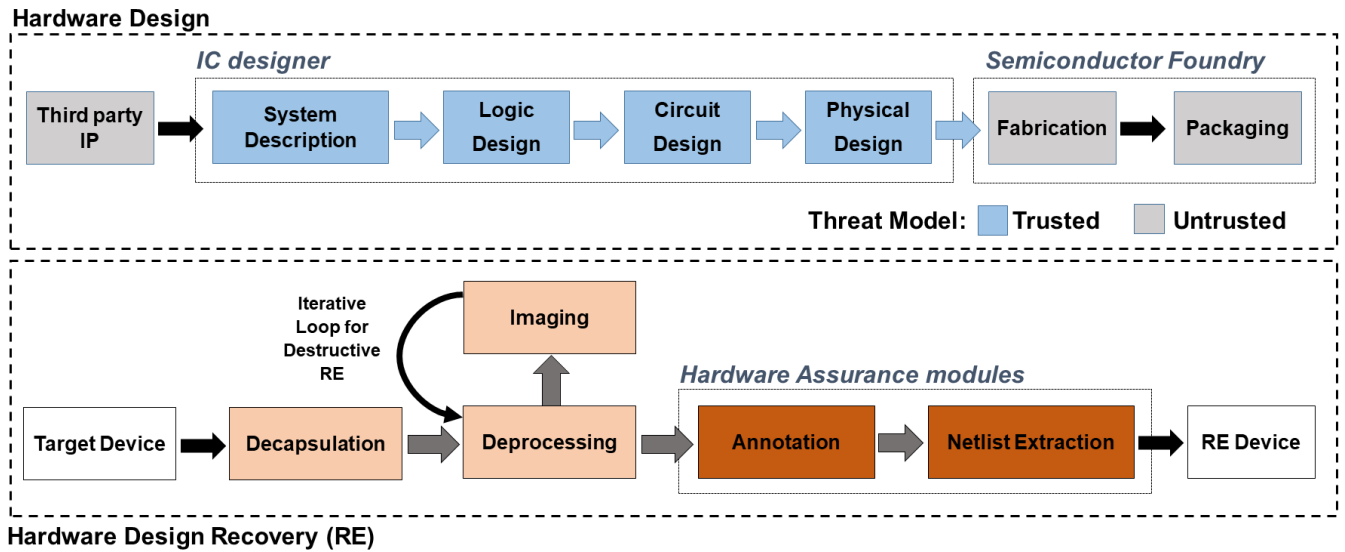


FIGURE 1: The hardware design and recovery workflow assumptions in this proposal. The commercial entities involved in the design process and the threat models are also provided.

can trigger stealthy, well-placed hardware Trojans in billion-transistor chips is usually a near-impossible task. In side-channel signal analysis approaches, process variations and measurement noise undermine the probability of detecting small-scale Trojans. As a result, the confidence level in detecting Trojans using existing techniques are quite low. At present, Reverse Engineering (RE), the process of acquiring the source hardware design by destructive physical analysis of the end product, is the only approach that can assist in these scenarios and guarantee trust. As a result, RE has gained attention in recent years and experienced community-wide acceptance as an effective approach for hardware assurance [5]–[7].

Although RE is a versatile tool for the hardware assurance community, the pace of its adoption into mainstream use is lagging. The primary obstacle is the negative connotation behind RE as an attack mechanism used by an adversary to illegally acquire design schematics and counterfeit IP rather than a hardware assurance tool. The concerns from utilizing RE as an attack mechanism can be prevented by the use of IC obfuscation and camouflaging [8]–[10]. There are several obfuscation methods available for securing layout-level and netlist-level information rendering them hard to decipher [11], [12]. Some approaches generate functionally identical logic gates with very different physical/layout realizations [13], [14]. Consequently, they hinder the RE workflow. However, they do not hinder the process for the IC designers with access to the source design files. i.e. the golden data. In contrast, it provides a significant advantage and incentive for the IC designers to adopt RE as a tool for hardware assurance. Another concern raised in the adoption of RE is the considerable investment required in terms of infrastructural, computational and human resources [6], [15]–[18]. Although, some infrastructural investment is inevitable, the

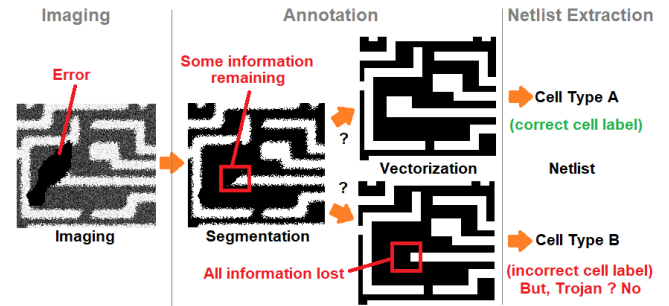


FIGURE 2: Exemplary RE workflow depicting accumulation of errors in each successive step.

computational and human resources constraint can be addressed through process automation and efficient algorithms. With the negative connotation behind RE addressed, the key requirement for developing RE-based hardware assurance into a popular mainstream approach is the incorporation of data-driven paradigms, such as machine learning (ML) and deep learning (DL), into the RE workflow. Consequently, this raises the question: **Can data-driven paradigms be directly integrated into RE and hardware assurance?**

It is a common belief that the widespread adoption of data-driven approaches in recent years is primarily fueled by the availability of massive amounts of data. This is not entirely true. In addition to data, a deeper understanding of the target domain is necessary. For instance, consider the exemplary RE workflow shown in Figure 2. It can be seen that a few missing pixels in the annotation stage causes a standard cell misclassification down the line in the netlist extraction stage. In terms of hardware assurance, it results in an ambiguous situation where the error cannot be resolved into a true Trojan detection or an error associated with

the image acquisition process. In contrast, errors in a few pixels, typically, do not result in severe consequences in the image analysis domain. Similarly, there are several image segmentation evaluation metrics, such as the Intersection-over-Union (IoU) and Structural Similarity Index Measure (SSIM), that are capable of evaluating segmentation based on shapes but are inherently incapable of incorporating electrical connectivity information into account. They lead to short-circuits or modification of the intended functionality resulting in significant inflation of RE process time frame to account for manual error resolution [19]. As demonstrated in the examples, a deeper understanding of the domain and its inherent challenges needs to be acquired before data-driven approaches can be fully integrated into the hardware assurance problem. The prominent challenges are:

- **Fault-intolerance of individual RE modules:** In the exemplary RE workflow in Figure 1, every module in the workflow expects the output from the preceding module to be error-free. This is seldom the case. At present, there are no approaches suggested in literature that can handle faulty data. With the RE workflow being sequential, errors accumulate at every successive step making the conclusion obtained from the process uncertain.
- **Ad-hoc nature of the RE workflow:** RE is not a fully defined formal process. Although the modules in the workflow, shown in Figure 1, has its steps well defined as part of the process, the approach taken to achieve them may be different. For instance, the challenges introduced by layering the IC through Computer Numerical Control (CNC) milling is different from executing the step using a Focused Ion Beam (FIB). Similarly, the existence of several varieties of gas and acid chemistry in deprocessing may have different effects on different ICs. They are not guaranteed to be repeatable. Repeatability, reproducibility and comparability are key requirements for any well-defined process.
- **Sensitive nature associated with the design data:** There is a significant shortage of data dedicated for RE and hardware assurance applications. This can be attributed to two reasons. First, the time and resource cost associated with performing RE on ICs and labelling the data manually. With approaches such as DL requiring several hundred thousand labelled images, this is a significant undertaking. Although this can still be achieved, the legal ramifications associated with disclosing sensitive design data, the IP of the IC designers, is much more severe. Without proper obfuscation of the design data, such as through privacy-preserving transforms [20]–[22] or zero-shot learning [23]–[25], disclosing them may cause more harm than good.

To effectively assimilate data-driven paradigms into RE and hardware assurance, these challenges have to be resolved. The purpose of this paper is to introduce a clear path to resolve these issues and facilitate the formal transition into

data-driven approaches. The rest of the paper is structured as follows: *Section II* lists, in detail, the noise sources commonly encountered in the RE workflow. Emphasis is done on classifying these sources as predictable (modelable) and random (non-modelable). *Section III* leverages the curated noise taxonomy to generate synthetic SEM images for use in RE and hardware assurance applications. *Sections II* and *III* collectively assist in inserting domain knowledge and generating synthetic data to address the **fault-intolerance of the individual RE modules** and the **sensitive nature associated with the design data**. *Section III* further expands on a risk analysis approach to handle the uncertainty introduced by the **ad-hoc nature of the RE workflow**. *Section IV* performs a thorough performance analysis and discusses on the generalizability of existing ML and DL approaches in RE and hardware assurance using the generated synthetic dataset. These would serve as a baseline for building better data-driven approaches for RE and hardware assurance. Finally, the work is concluded in *Section V*.

A. CONTRIBUTIONS

Our goal in this paper is to introduce a clear path to resolve the issues associated with RE domain and successfully integrate data-driven approaches into the RE workflow to provide for a scalable, automated and fault-tolerant workflow for RE-based hardware assurance. Our key contributions are as follows:

- A detailed taxonomy of errors that can affect the efficacy of RE-based hardware assurance. These error sources are collated from previously published literature on hardware assurance and several RE case studies.
- A synthetic open-source¹ Scanning Electron Microscopy (SEM) image dataset, called **REFICS**, generated using various imaging parameters and two node technologies (32nm and 90nm). The dataset also includes augmentation using error sources from the taxonomy for developing robust data-driven approaches. The dataset has samples for the doping, polysilicon and the metal layer culminating in 800,000 images. REFICS is the first and the largest SEM image dataset ever introduced in the RE and hardware assurance community.
- Benchmarks and performance analysis for existing ML algorithms and DL models used in RE. Additionally, for the first time in the RE community, the generalization capability of data-driven approaches across node technologies and IC layers is investigated.
- Introduction of a risk analysis approach for facilitating comparison between RE executed using ad-hoc steps and assessing its influence on hardware assurance.

II. UNDERSTANDING NOISE INTERACTIONS IN RE

As illustrated in Figure 1, a typical execution of the RE workflow begins with decapsulation of the IC package –

¹Published under Creative Commons Attribution (CC BY 4.0). Link: <https://creativecommons.org/licenses/by/4.0/>

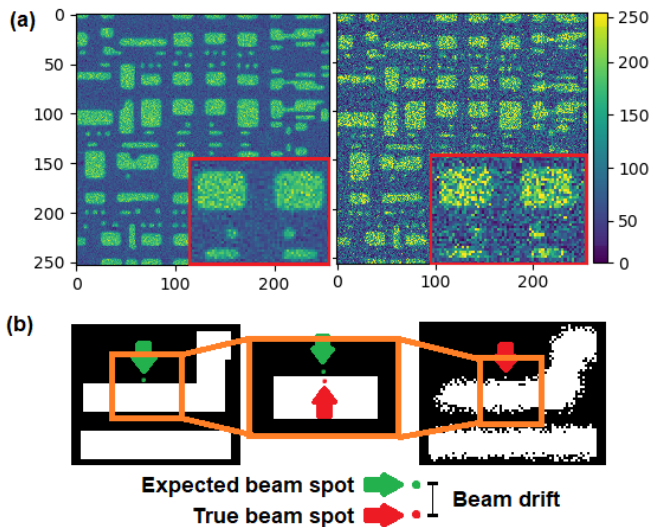


FIGURE 4: Exemplary cases of imaging noise. (a) Pixel intensity variation associated with beam interactions. The SEM image on the left has a higher dwelling time per pixel than the image on the right. (b) Demonstration of beam drift in SEM imaging.

the scanning beam is commonly known as *shot noise* and can be represented using a Poisson modulated process. A similar issue happens with the secondary emissions as well. This leads to noise in the beam interaction which can be modelled using either a compound Poisson process or a Gaussian-Poisson process (by approximation) [26]–[30]. In addition, the electronic components of the SEM (e.g., amplifiers and scan generators) induce an Additive White Gaussian Noise (AWGN) into the response but its influence was found to be negligible as compared to that of beam interaction noise [31], [32]. The consequence of these interactions can be visualized as pixel intensity variations in pixels belonging to the same material under test. This affect can be overcome by sampling the same area for a longer amount of time and averaging out the response. In terms of SEM imaging, a longer dwelling time per pixel will accomplish this task. An exemplary case is shown in Figure 4(a). It was suggested that the noise introduced by imaging modalities can be used to hide hardware Trojans [18].

Beam drift: The scanning beam excites/irradiates a small spot on the material. However, in some cases, random fluctuations in the scanning beam and mechanical creep induced by the staging platform of the SEM causes the spot to move from its intended position on the material. If this situation happens on the transition boundary between materials, then the transition edge gets corrupted with interchanged response between the two materials. An example of the phenomenon is shown in Figure 4(b). Drift is more pronounced with a larger Field-of-View [33]. Although the beam drift is random, the process can be modelled using a Gaussian distribution for the true position of the beam with respect to the intended position

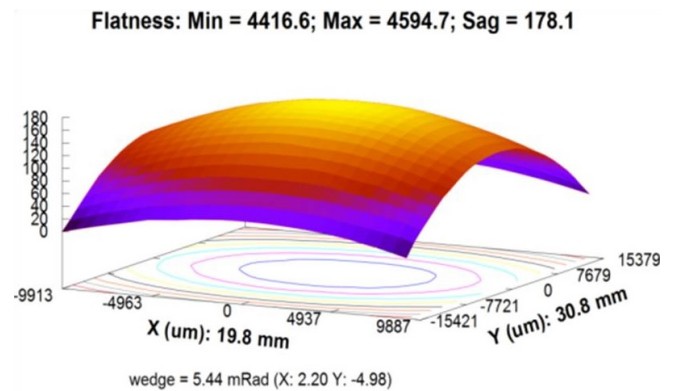


FIGURE 5: Warpage induced in the IC die due to accumulated mechanical stresses [35].

[34]. It is present in every electron microscopy modality and cannot be fully accounted for in every case. This noise source is the likely culprit behind unintended short-circuits in between IC structures in SEM images.

Die warpage: This source of error is associated with the deprocessing of the IC. Deprocessing requires delayering the IC die with a fixed cross-sectional thickness. Incremental removal of the material, especially the bulk of the silicon substrate supporting the IC structure, results in mechanical stress accumulating on the die and causing it to warp [35], [36]. This phenomenon, shown in Figure 5, results in a perspective distortion on the features in the imaged region. Although this issue is not resolved yet, knowing the curvature of the die can, potentially, help resolve this distortion. A small Field-of-View can also help alleviate this issue by flattening the region under focus from the observer’s perspective.

Feature dimensions and proximity: These errors are a direct result of the layout synthesis and so-called design rules. Complex geometry of structures can only be imaged if they are within the resolution capability of the imaging modality. Similarly, structures placed in close proximity with each other may, also, not be resolved effectively. In simpler terms, these features may be truncated by the SEM unless a small Field-of-View or high magnification is used. This in turn significantly inflates the resource requirements and cost associated with performing RE or RE-based hardware assurance.

Flicker noise: The concept of flicker noise is common knowledge in the field of semiconductor physics. This parameter is coupled with the performance characteristics of semiconductor devices. This source of noise occurs during normal device operating conditions when electrons are trapped inside electron holes and released after a short time delay. The time delay is modelled using the power law. This noise source is not considered during SEM imaging because the device is assumed to be not operational. However, the principle behind SEM imaging using electrons is similar to the device under operating conditions. In simpler terms, the electrons in the scanning beam of the SEM is akin to the electrons flowing

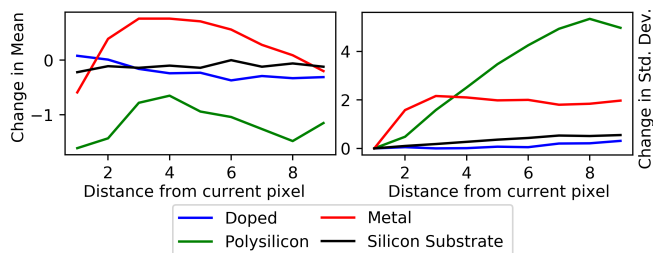


FIGURE 6: Plots indicating the influence of flicker noise interactions on SEM images. The change in pixel intensity values due to flicker noise is inconsequential as compared to the standard deviation in electron beam interaction responses from the materials.

through the device during operation, albeit with different energy levels. The influence of flicker noise on SEM images of ICs have not been studied in literature.

To address this issue, a set of experiments were conducted to study the influence of flicker noise in SEM images. Our experiment protocol for these experiments are straightforward. In theory, the electrons captured by electron holes during irradiation are released in accordance with the power law after a time delay. The SEM, with its reliance on these electrons for image formation, will cause an observable increase or decrease in pixel intensity values. Assuming the imaging modality to scan in a raster mode, the experiment evaluated the change in pixel intensity values when the current material is same as the material observed earlier in the scan, i.e. the scanning beam is on the same material in sequence. The experiment considers the impact of past ten pixel values on the current pixel value if they belong to the same material. Four materials commonly used in the manufacture of ICs were considered: silicon substrate, doped silicon, polysilicon and metal. The samples under study were collected with high dwelling time per pixel and verified to be devoid of any other noise sources. The results are shown in Figure 6. As expected, the materials under SEM observation exhibits flicker noise but at inconsequential levels. The mean change observed was around one intensity level on the entire pixel intensity scale for SEM images (0.004% of the pixel intensity scale: 0→255).

B. RANDOM INTERACTIONS

Some of the random interactions stated below can be analytically modelled but they are still considered as random because their influence on the acquired SEM images cannot be effectively modelled. For example, the process of electromigration is well-known in semiconductor physics and can be modelled analytically but the current operating state of the IC under test may be unknown.

Radiation damage: Damage to the sample under study happens when the sample is irradiated for an extended time period. Sample damage also happens if the radiation power is too high. i.e. high-dose radiation. The radiated region sat-

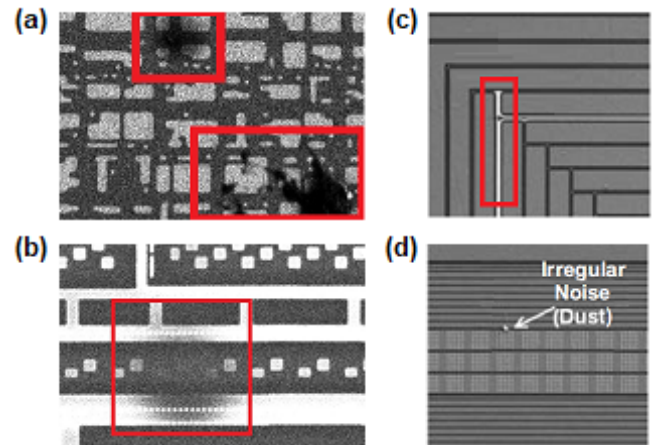


FIGURE 7: Exemplary cases of deprocessing errors. (a) Residue leftovers from the etching process. (b) Missing structures as a consequences of uneven delayering. (c) Conduction in the active layer. (d) Corruption in the SEM image due to a dust particle.

urates to form a contamination layer and suppresses further emission of electrons [37]. Due to the risk of sample damage, there is a persisting interest in the hardware assurance community for low-dose imaging [38].

Residue: Any foreign remnants on the surface of the die that prevents observation and resolution of surface-level features on the IC die can be called a residue. Typically, the etching process used for delayering may leave some residue on the sample. They can be prevented by cleaning thoroughly. For instance, using an ultrasonic bath. An example is shown in Figure 7(a).

Uneven delayering: This specific noise source is a direct consequence of surface-level imperfections on the IC die. Areas with high roughness or edges between materials in the IC have larger escape areas for the secondary electrons [37], [39]. These cause pixel intensity variations in the same material measured in different locations on the same IC die. Further, if the IC die is not mounted properly, the delayering process may unevenly delayer the surface of the die. This may result in features belonging to two layers getting merged together as shown in Figure 7(b). These issues can be resolved by polishing the deprocessed IC die for a planar surface before imaging.

Conduction: Insulating materials may charge positively and suppress the electrons required for obtaining a proper SEM image [39]. This leads to localized pockets of bright and dark regions in the image as shown in Figure 7(c). It is also suggested that dynamic charging of the sample deflects the electron beam from its intended position and cause the intensity of the induced signal to vary uncontrollably [40]. This issue can be resolved by depositing thin layers of conductive materials such as carbon or platinum on the IC die surface [41], [42].

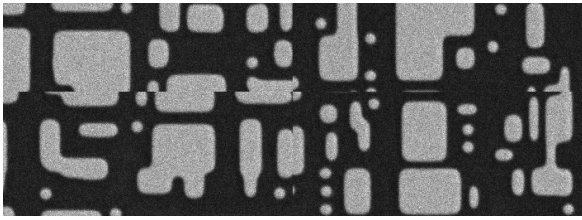


FIGURE 8: Exemplary case of stitching error on both axes.

External contaminants: This class of errors represent a wide selection of external factors that affects the RE workflow in a detrimental way. For example, dust particles are a common source of corruption as shown in Figure 7(d) [43]. Vibrations, even those that are barely perceptible, along with thermal expansions in the sample caused by slight temperature fluctuations in the environment, are examples of this noise source and can significantly affect image quality [33]. This issue can be resolved by isolating the RE device from the environment.

Stitching: In most cases, the Field-of-View provided by SEM does not cover the entire region of interest. This typically results in multiple images to be collected and stitched together to form a complete image. In typical applications, the stitching process involves taking two images with a certain overlap. The degree of overlap is decided by the operator and remains fixed for the entire image acquisition phase. Stitching is usually an error-prone process since it involves finding key points in two images corresponding to the points of highest similarity based on which the two images are merged together [44], [45]. In contemporary nano-scale node technology, the features are very much similar and repetitive resulting in false key point detection and faulty merging of images [46], [47]. An example of a stitching error is shown in Figure 8. To emphasize, if the IC design layout is known beforehand, the likelihood of stitching error can be predicted based on the similarity shared between the features in the overlapping field-of-views. In such situations, stitching errors can be classified as a predictable interaction. With no prior heuristics on the IC design layout, stitching errors are random and are resolved manually by a Subject Matter Expert (SME).

Vertical alignment: Owing to the vertical layered construction of ICs, the individual images taken from multiple layers have to be stacked on top of each other to reconstruct the features from the IC. The alignment is done using correlation matching, typically utilizing vertical interconnects (vias) [48]. With correlation involved, the issues related to stitching are also experienced in vertical alignment. At present, design rule checks and manual operator intervention are used to validate the vertical alignment of the image stack.

Oxidation: Delayering exposes the metallic structures in the IC to the atmosphere [49]. Oxidation of metallic surfaces causes fluctuations in the pixel intensity responses obtained for the same material at various points in the IC. A possible

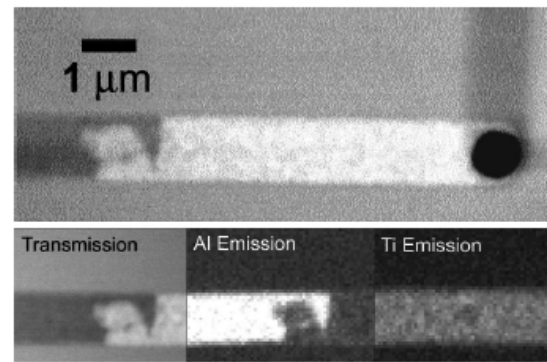


FIGURE 9: Example of electron-migration in the metal layer when subjected to accelerated life conditions.

solution is to perform the deprocessing tasks in an inert atmosphere.

Electromigration: If the IC has been under use, there are chances of having electromigration and changes in the physical structure of the material [50]. This type of defect is usually found in metal interconnects through which high-density currents flow. As shown in Figure 9, electromigration causes change in the pixel intensity values belonging to same material depending on the degree of migration. This issue is predominantly studied in device failure analysis in estimating the lifetime of an IC [51].

Process variations: Due to the high resolution capability of the imaging modality, any small variation in the manufacturing process would cause changes in the acquired image. The degree of influence of these variations depend on the precision/tolerance of the manufacturing process and the resolution of the imaging modality. These variations, being naturally stochastic processes, may not necessarily be modelled parametrically. Therefore, modeling these variations using statistical models is not recommended [18]. In terms of image analysis, the physical realization of the IC design may not be the same as the synthesized layout but, typically, very similar.

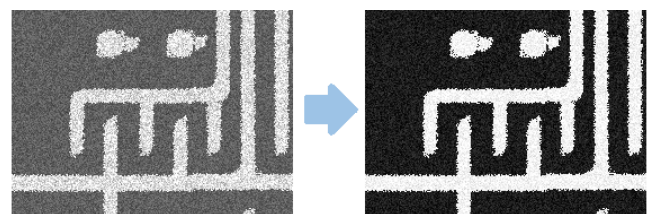


FIGURE 10: Example of operator interaction. Contrast enhancing modifications are applied to the SEM image on the left to obtain the image on the right.

Operator interactions: The operator can perform several modifications to the SEM. Other than basic imaging parameters, such as Field-of-View and dwelling time per pixel, there are several other parameters that can be adjusted by the operator. An exemplary situation involving contrast enhance-

ment is shown in Figure 10. These changes may not be the same across all images acquired by the same operator and constitutes a source of randomness in the image acquisition phase.

Incorrect information: SMEs play a significant role in the RE workflow. Even with their expertise, they are still subject to errors in decision making. The influence of human factors in decision making was investigated with respect to hardware assurance in an earlier work [52]. For instance, SMEs help populate the standard cell library for extracting the gate-level netlist in the RE workflow. With layout-level and gate-level obfuscation applied, it is possible for the SME to incorrectly identify a logic gate. The consequence of incorrect logic gate assignment was demonstrated in the example in Figure 2 and objectively assessed in a RE case study where the authors reported that the time frame required for error resolution was larger than the time frame required for imaging the IC [19]. Limiting human interaction with RE workflow, especially that of inexperienced SMEs, can address this issue to a great extent. In addition to identifying the standard cell library, the role of SMEs in RE also extends to identifying anomalous data, such as changes introduced by stitching errors, and validating compliance with design rules. Without the introduction of robust data-driven algorithms that can effectively identify and resolve anomalies in RE data, active input from SMEs will be required in the RE workflow.

As demonstrated through the taxonomy, some of the noise sources can be modelled and their influence on the RE workflow controlled to a large extent. The random noise sources are more problematic. As discussed, some of these issues can be prevented using simple precautionary measures. For addressing the other noise sources, suitable data-driven approaches needs to be incorporated to detect and curb their influence in the RE workflow.

III. GENERATING SYNTHETIC SEM IMAGES FOR RE

The process of image formation in SEM can be simulated. Existing studies in electron microscopy imaging and works in the fault analysis community support this statement. The electron microscopy community introduced an SEM image simulator called ARTIMAGEN, an initiative supported by the National Institute for Standards and Technology (NIST) [33], [53]. This simulator can generate images with varying influences of drift, vibration, thermal expansion, and noise profiles (Gaussian/Poisson). However, the selection of materials, noise profiles, and shape contours are limited and not suitable for IC RE and hardware assurance. The fault analysis community extensively uses simulated images for benchmarking Line Edge/Width Roughness (LER/LWR) algorithms [34]. A recent DL approach generates synthetic SEM images based on layout data for mask optimization and virtual meteorology [54]. Although the data used for these approaches are not publicly available, it does provide a path towards integrating data-driven approaches for RE in the imaging phase [48].

An SEM image generator that can support the inclusion of several noise sources for data diversity is key in adapting data-driven approaches into the hardware assurance community. Being synthetic and simulatable, a large quantity of good quality data can be generated without legal consequences and fear of compromising sensitive design data. Even if an attempt is made on collecting real SEM image data, it will be considerably hard to induce errors on demand. Further, the synthetic image generator can be used to augment real RE case studies. Performing RE on an IC does not yield enough data for training data-hungry techniques like DL. However, designers with access to the layout data can synthetically generate more data from a specific IC to make the model more robust on the chosen IC. Also, data-driven approaches, such as image in-painting, are extensively used to recover corrupted data when exemplary corrupted data with ground truth (GT) labels are available for training [55]. Therefore, synthetic SEM image generation will resolve two key challenges in the assimilation of data-driven approaches for hardware assurance and trust: **the sensitive nature associated with the sharing of design data and incorporating fault-tolerance into individual RE modules**. In Section III-A, the generation process behind synthetic SEM images is discussed followed by the approach on addressing the final challenge, **the ad-hoc nature of the RE workflow**, in Section III-B.

A. SIMULATING THE SEM IMAGE GENERATION PROCESS FOR RE

The initial requirement for generating a synthetic SEM image is to have a context. The context, in this case, is the layout-level design file for an IC. With strict control on the availability of industry-use standard libraries, open-source educational standard cell libraries were used to synthesize an open-source² Advanced Encryption Standard (AES) design into approximately 10,000 standard cells from which four cardinal layers were extracted: doping, polysilicon, contacts and the metal layer. The standard cells were acquired from 32/28nm Educational Design Kit and Synopsys open educational design kit containing 350 and 340 standard cells, respectively [56], [57]. The layout files were split into 250×250 patches and fed into the image synthesis workflow, as described in Figure 11, along with the image synthesis parameters.

There are two sets of input parameters for image synthesis. The first set corresponds to the imaging settings in the SEM. In our case, Field-of-View/Magnification and dwelling time per pixel (3.2 $\mu\text{sec}/\text{pixel}$ and 10 $\mu\text{sec}/\text{pixel}$). With the layouts synthesized to maintain a 1:1 relative scaling for effective comparison between node technologies, the features in the 32nm layout is smaller than the 90nm layout. Consequently, Field-of-View/Magnification was setup differently for the node technologies. The 32nm layout was up-scaled (1×, 2×,

²The AES designs used in this paper are licensed from Synopsys for open-source usage under Creative Commons Attribution licensing

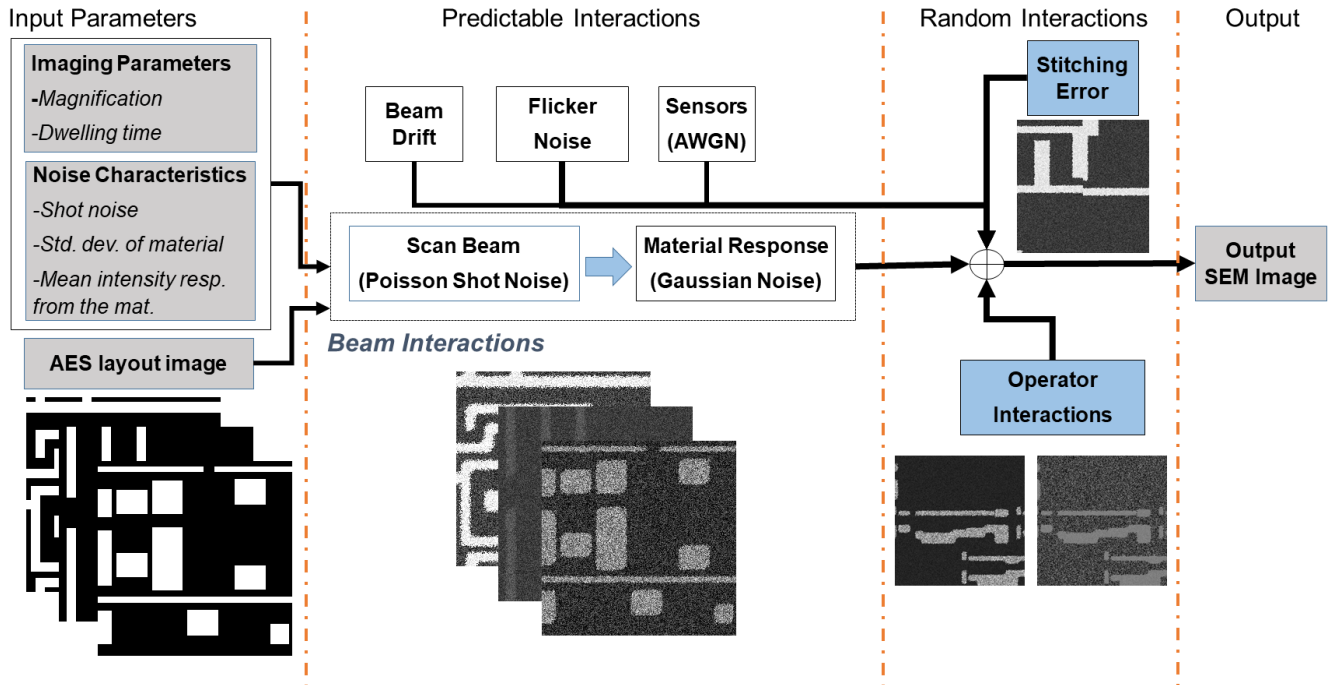


FIGURE 11: Workflow for generating a synthetic SEM image for the REFICS dataset.

$3\times$ and $4\times$ the original standard cell dimensions) and the 90nm layout was down-scaled ($1\times$, $2\times$, $4\times$ and $6\times$ the original standard cell dimensions). The second set of parameters corresponds to the noise characteristics: the shot noise (λ_{shot}) parameter for the primary scanning beam and the expected mean (μ_{mat}) and standard deviation (σ_{mat}) for the pixel intensity response from the material under study. The shot noise distorts the scanning beam intensity at 2%, 5%, 10% and 20% with an excitation potential of 5kV. Every μsec spent on a pixel is equivalent to 1000 samples acquired from the simulation. So, a single pixel acquired at $10 \mu\text{sec}/\text{pixel}$ setting would simulate 10,000 samples from a Monte-Carlo simulation using the beam interaction model. A Poisson-Gaussian model was used to model the beam interactions in the workflow. The model for obtaining a sample using the primary scanning beam (PE) and the corresponding secondary response (SE) from the material is shown in Equations 1 and 2. The average electron count is dependent on the current flowing through the beam emitter but it is assumed to be constant for our model. Typically, the beam current is not modified in real-world experimental setups. The beam drift was simulated using a 5×5 kernel where the beam drift probabilities from the center of the kernel to the periphery was determined by a Gaussian distribution. Finally, the model was augmented using flicker noise and sensor noise. The mean pixel intensity response for each material was acquired from real SEM images of that layer. This concludes contributions from the predictable interactions of the noise taxonomy in the image synthesis workflow.

$$PE = Poisson(\lambda_{shot}) \quad (1)$$

$$SE = Gaussian(\mu_{mat} \pm k \times 2.5 \times \sigma_{mat}, \sigma_{mat}) \quad (2)$$

$$\text{where, } k = \frac{PE}{\max(PE) - \min(PE)}$$

The random interactions are added to the image after the SEM image is generated. Currently, the synthesis workflow applies two random interaction to the SEM image: stitching errors and operator interactions. Stitching errors are taken as random in the workflow and applied randomly on either the vertical, horizontal or both axes at the same time. The operator interactions are limited to contrast adjustments in the image. This modification is applied by introducing random variation to the material's mean pixel intensity response. For instance, the mean pixel intensity response for the doped region is 160. Operator interaction randomly samples a Gaussian distribution with the mean set at 160 and a standard deviation of 15 to change the mean pixel intensity response from the materials. This modification can increase or decrease contrast in the image. Finally, a standalone modification, not listed in the taxonomy, is applied to the AES layout image at the input stage. The corners in the original layout are converted to simple curves in the SEM images as shown in Figure 12. This is to capture the variation introduced by the mask generation process for etching the IC layout onto the wafer during manufacturing [58]. The radius of the curve is randomly sampled from one to five. The chosen magnification parameter also scales the radius. The final SEM image is generated at the end of this stage. Along with

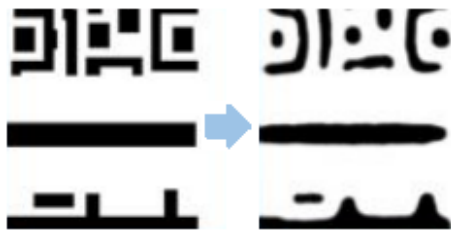


FIGURE 12: Edge distortion caused by limitations in the mask generation process in transition from the GDSII layout (left) to the silicon wafer (right) [54]

noise interactions that can be prevented through reasonable precautions, this SEM image synthesis workflow constitutes the closest substitute to a real SEM image obtained from an IC through RE.

The claim for the closest substitute to a real SEM image can be verified through extensive experimentation. The experiment protocol applied for verification consists of acquiring a large number of SEM images of an IC at a fixed set of imaging parameters and comparing them against synthetic SEM images generated using an exhaustive set of imaging parameters through Monte Carlo simulations of the beam interactions and the other predictable interactions. If the model is valid, then the statistical similarity of the synthetic and the real SEM image should be highest at the same imaging parameters.

A smart card IC was deprocessed to satisfy the real SEM image data requirement for the experiment. A $250\mu\text{m}$ window was opened on the flip-side of the IC using a FIB and images were acquired using a $25\mu\text{m}$ Field-of-View and dwelling times of 10 and $3.2\mu\text{sec}/\text{pixel}$. With a fixed resolution of 1024×1024 pixels, 25 SEM images of the diffusion layer were captured for each dwelling time setting. These images were hand-labelled as pixels belonging to either the silicon substrate or the doped region. For the Monte Carlo simulation counterpart in the experiment, the imaging parameters used were the dwelling time per pixel, shot noise and the standard deviation of the material. The simulation generated 64,000 pixels for every possible combination of the parameters listed earlier. The comparison between the real and synthetic SEM image can now be performed.

In image processing and computer vision, the similarity between images are assessed using two distinct characteristics of the image: the image histogram and texture. In both cases, the data preparation follows the same process. Initially, a hand-labelled ground-truth image is taken and the labelled pixels are filled in by sampling the pixel values generated by the Monte Carlo simulation to produce a synthetic image representing the particular set of image simulation parameters. This is repeated for every possible combination of simulation parameters. The mean pixel intensity values for the silicon substrate and the doped region were used to offset the histogram for the simulated images. For instance, silicon substrate and the doped region had a mean pixel intensity

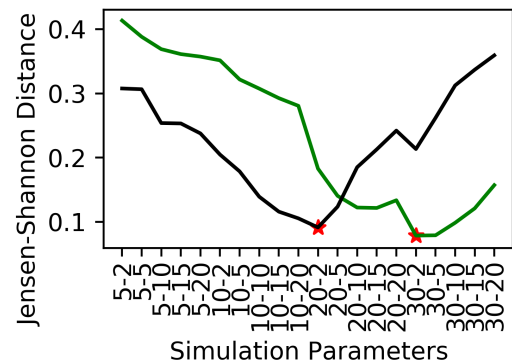


FIGURE 13: Plot indicating the similarity between the real and synthetic SEM image. Simulation parameter format: Standard deviation of the material-Shot noise parameter. The black and green trends indicate $10\mu\text{sec}/\text{pixel}$ and $3.2\mu\text{sec}/\text{pixel}$ dwelling times with the real parameters estimated at 22-2 and 38-2 respectively.

value of 60 and 161 respectively in our real SEM image data. For assessing the similarity in image histogram, the image histogram (pixel intensity frequency distribution) of the real SEM image and the corresponding synthetic SEM images representing every combination of simulation parameters are taken. The similarity between image histograms were assessed using the Jensen-Shannon distance. A distance value of zero indicates that both the histograms are the same. The results of the experiments are shown in Figure 13. The plot shows that, even with all the possible parameter combinations for the simulated images, the points of highest similarity correspond to the true parameters estimated from the image. Similarly, for assessing the similarity in image texture, the real SEM images and the synthetic SEM images were decomposed using the Fourier transform. The magnitude spectrum for both images were rearranged into vectors and the similarity was assessed using the cosine distance between the two vectors. Both the experiments produced identical results. Similarity in histogram of the images suggest that the distribution from which the pixels are sampled are identical. Further, with the Fourier domain representation of the images being identical, the relationship of a pixel with its neighbouring pixels (i.e. the texture) is preserved as well. These observations, along with the model validations reported in literature, suggests that the real and synthetic SEM images are very similar.

B. IMPACT OF AD-HOC PROCESSES ON HARDWARE ASSURANCE

There are several works that perform RE on a particular IC, typically a smart card, as a case study for reporting on the challenges and resources cost incurred during the work [6], [19], [58]–[61]. The modules in the RE workflow shown in Figure 1 mostly acts as a placeholder. The actual technique used to execute vary in between ICs and researchers. For

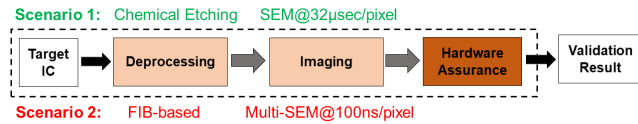


FIGURE 14: Example indicating the ambiguity introduced by ad-hoc processes in RE and hardware assurance.

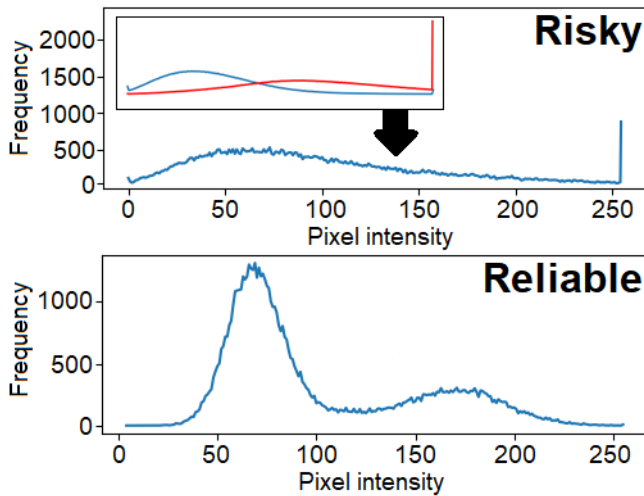


FIGURE 15: Exemplary image histograms for the doping layer. The top and bottom image histograms are acquired at 3.2 and 10 $\mu\text{sec}/\text{pixel}$ dwelling times respectively. The resolved individual distributions are shown for the top plot. There is more overlap in pixel intensities for the individual materials in the top than in the bottom making it hard to resolve pixel membership.

instance, some perform delayering through FIB while others use chemical or mechanical etching. FIB provides convenience at an increased cost and may not be available to all researchers. Some works have access to state-of-the-art SEM machines like the multi-SEM [62] while others are restricted to regular SEMs. The cost incurred and the challenges faced in each of these cases are different. In contemporary nanoscale node technology, like the 14nm finFET, the challenges faced may be completely different and require some other sophisticated technology. A demonstration of this situation is shown in Figure 14. The questions raised here are: *Is the design validation obtained through Scenario 2 better than Scenario 1? How can it be quantified? Does the investment in specialized equipment produce an equivalent boost to the reliability of the hardware assurance process?* The complexity in answering these questions are comparable in scale to the diverse options available for executing the modules in the RE workflow. Currently, there are no common grounds for comparison between these works. With the goal of hardware assurance being trust in the ICs design and functionality, the results obtained through various hardware assurance processes should be comparable and repeatable.

The IC design and RE workflow requires several decisions

to be made from the beginning to its conclusion. Every decision has an impact on the successive steps in the workflow. For instance, if the IC is designed using dielectric materials that does not provide contrast under SEM observation, it will be extremely challenging to validate the design of the IC even with access to state-of-the-art equipment and expert assistance. Therefore, the impact of each decision has to be measured in terms of the uncertainty/risk it introduces to the process. A suitable demonstration can be made using the simulated workflow developed in this paper. The purpose of RE-assisted hardware assurance is to acquire design information, both layout-level and netlist-level, to verify the functionality of the IC and, consequently, enable trust in the device. With the sequential nature of the RE workflow, this in turn condenses to acquiring reliable images of the IC from which the features can be accurately extracted. The reliability of an image lies in effectively differentiating a pixel belonging to the silicon substrate from that of the polysilicon or metal pixels. In image analysis, the process of assigning pixel membership is accomplished through the image histogram, the combined pixel intensity distribution of the materials. Every decision made in the RE workflow, such as choosing the dwelling time per pixel parameter or the quality of the SEM device used to acquire the image, affects the resolvability of pixel membership by increasing or decreasing the overlap between the distributions of the individual materials. A graphical illustration of this situation is shown in Figure 15. This can also be observed in Figure 4(a) with the same region of interest on an IC acquired at different imaging settings exhibiting a change in the standard deviation of the pixel intensity values belonging to the material. The ratio of the area of the overlapping region in the distribution to the entire area under both the distributions can be taken as the risk or uncertainty introduced by the decisions in the workflow. An ideal situation requires having zero overlap between constituent material distributions, a highly unlikely case.

The change in the chosen SEM imaging parameters affects the histogram and, consequently, the overlap between distributions of the individual materials. Hence, the risk can be calculated as a function of the imaging parameters. To facilitate this, the synthetic workflow can be used to generate image histograms with all possible combinations of shot noise and dwelling time representing the quality of the SEM device and the imaging parameters respectively. In addition to the imaging parameters, the materials chosen to manufacture the IC also has an associated risk value with it. The quality of the materials is represented through the sensitivity index (D) as defined in

$$D(N_1, N_2) = \frac{|\mu_1 - \mu_2|}{\frac{1}{2}\sqrt{\sigma_1^2 + \sigma_2^2}}, \quad N \sim \text{Gaussian}(\mu, \sigma). \quad (3)$$

The mean (μ) and standard deviation (σ) in Equation 3 is calculated from the individual pixel intensity distributions from the materials under study. The sensitivity index is a proxy for contrast provided by the material under SEM

observation. Since contrast is relative, at least two materials are required to obtain the index. A higher index indicates higher contrast. Now, the risk/uncertainty can be defined as a conditional distribution over all the parameters that can be modified/chosen during the execution of the RE workflow.

The result of this approach is shown in Figure 16. The plot demonstrates an exemplary study with the choice in imaging parameters. Some observations in this plot is intuitive. For instance, the risk decay is higher for high dwelling time images than for the lower dwelling time images for increasing sensitivity index. In addition, although counterintuitive, the plot suggests using noisier images when the contrast is close to non-existent. This can be attributed to the fact that noisier images increase the standard deviation of pixel intensities belonging to a certain material making it likelier to find regions in the image histogram where the material intensities do not overlap. Using a high quality image with no contrast provides no additional information. All these qualitative observation can be drawn from the plot with the added possibility to quantitatively reason the benefit of choosing a particular set of parameters in the RE workflow.

One of the major concerns at this point is the calculation of the parameters that cannot be directly observed. For instance, the shot noise or the pixel intensity distribution for the individual materials without access to the ground truth labels. There are approaches in literature that can address these issues. The shot noise parameter can be directly estimated from a single SEM image [26]. LASRE, an approach developed for segmenting SEM images, captures the pixel intensity distribution for the individual materials without the ground truth labels or layout images [63]. This framework can be extended to other interchangeable steps in the RE workflow as well. The effect of every interchangeable step is connected to the overall risk to the RE process and, therefore, provides a common frame of reference for every interchangeable step in the hardware assurance process. This helps address the challenges introduced by **the ad-hoc nature of the RE workflow**.

IV. PERFORMANCE ANALYSIS

The primary bottleneck for RE, in terms of computation and human resource requirement, is the *annotation module*. Therefore, this module will benefit the most from the integration of data-driven approaches. The expansion of the workflow associated with the annotation module is shown in Figure 17. With a reliable annotation module built, the gate-level netlist can be obtained using the standard cell library, including using available design data, manual template matching [19], [45], [64], [65] or other automated approaches [46]. There are several approaches that are available for hardware assurance at the *netlist extraction* module. Interested readers are referred to a recent survey [5].

A. DATA-DRIVEN APPROACHES IN RE

The annotation module expects image data from the previous step to be noise-free, free of deprocessing errors, well

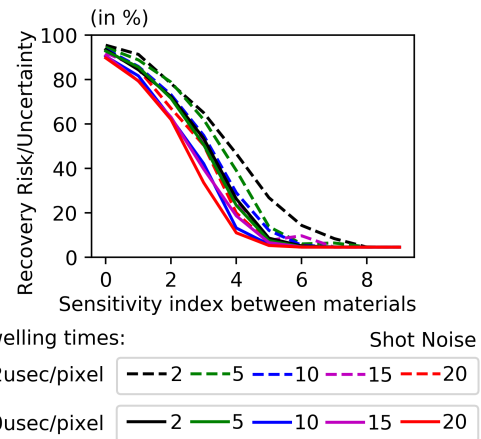


FIGURE 16: Representing risk/uncertainty in terms of controllable imaging and design parameters. As expected, a lower dwelling time increases risk in the RE workflow.

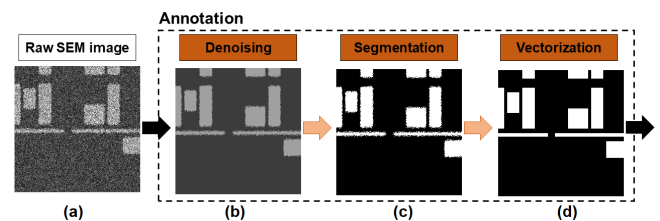


FIGURE 17: Expanded representation for the Annotation module in RE workflow.

stitched, and aligned in all layers. However, in most cases, these are not fully satisfied. Currently, there are no approaches to detect corrupted images and, hence, this module is fault intolerant. Inclusion of data-driven approaches focusing on detecting and resolving errors will assist in addressing this issue. The annotation module can be divided into three sub-modules: denoising, segmentation, and vectorization. An overview of the annotation block can be seen in Figure 17.

Denoising: The denoising sub-module is responsible for ensuring that the noise component in the image is suppressed. There are several approaches currently employed in literature for processing noisy SEM images. They include spatial filtering approaches, including Gaussian, median, curvature, anisotropic diffusion, wavelet, adaptive wiener filter, and hysteresis smoothing [66]–[69]. Simple high-frequency filtering and DL-based denoising approaches have also been used on SEM images [38]. These techniques are mostly naive image processing techniques and do not take the semantics of structures in the image into account. ML-based denoising approaches, such as image inpainting, super-resolution and dictionary-based sparse reconstruction, have also been explored for SEM images [34], [55], [70], [71]. Simple measures like Peak Signal-to-Noise Ratio (PSNR) and SSIM can be used to evaluate SEM image quality [72]–[75].

Segmentation: This step involves the separation of struc-

tures in the IC image based on some qualifying fact. In SEM images, this would be the material represented by grayscale pixel intensity values. Segmentation algorithms can be supervised, unsupervised or interactive. Supervised segmentation approaches require massive amounts of manually ground-truthed image data for learning representative models to discriminate between classes. Support Vector Machines (SVM) and Convolutional Neural Network (CNN) are examples of supervised segmentation methods [43], [76]. The unsupervised approaches are based on generalizable features that can be found in the same IC or across ICs. For instance, the technique developed by [76]–[78] relies on the fact that polysilicon structures and metal layer traces can be generated by simple Manhattan geometry contours. Interactive approaches, such as [79], require the operator to guide the segmentation. K-means and Fuzzy C-means are some simpler unsupervised segmentation approaches [77]. LASRE is another technique that relies on using frequency-based texture signatures for different materials to segment out IC structure across multiple layers [63]. Simple image processing techniques, such as Otsu’s binarization, have also been explored for segmenting SEM images [80]–[84].

Segmentation accuracy is measured in pixel accuracy/Mean Square Error (MSE), SSIM, F-measure and IoU. These measures rely on the similarity between shapes to evaluate segmentation accuracy and does not take electrical connectivity into account. Hence, a simple metric is devised to evaluate electrical connectivity in the segmented image. A common approach from image analysis, called connected components, is used to analyze the connectivity between shapes with respect to the ground truth image. In this method, a component in the image refers to a useful structural feature i.e. a trace in the metal layer or any white region in ground truth shown in Figure 18. If a short-circuit is present, then two separate components in the ground truth image will be represented as one component in the segmented image. This situation, also referred to as under-segmentation in image analysis, can be measured as the ratio of short-circuited components to all the components present in the segmented image – the custom metric (CC-US). Similarly, an open-circuit can happen when one component is split into multiple components due to a segmentation error. Such a situation, also referred to as over-segmentation in image analysis, can be measured as the ratio of open-circuited components to all the components present in the segmented image -the custom metric (CC-OS). In both cases, a value of zero indicates perfect segmentation in terms of electrical connectivity.

Vectorization: The vectorization stage converts the segmented image into a bunch of polygons. The idea behind this module is to recover the design files as close to the original die layout as possible. This specific step enables the use of commercial off-the-shelf tools allowing smoother transition between the annotation and the netlist extraction modules. This step further serves in suppressing edge noise between materials and compressing the amount of data in

the image, where the former is discussed in detail with Edge/Line Width Roughness (EWR/LWR) for fault analysis in ICs [85]–[87]. Vectorization can be achieved through simple edge following algorithms and custom tools like GDS-X [41], [88]. Traditionally, vectorization was well-utilized in compressing images, but, with the large capacity and inconsequential cost associated with data storage for present day computers, vectorization is not considered critical for most RE case studies. Most case studies conclude the annotation module with segmentation. In addition to reducing the memory footprint, a possible use for vectorization is in detecting corruption in segmentation like short-circuits and routing errors. Therefore, this sub-module will play a critical part in making the RE workflow more reliable and fault tolerant.

The annotation block is critical to several hardware assurance approaches. Being in the middle of the RE workflow, the amount of errors accumulated at this stage is far lower than further down the line. Several approaches use segmented images of the IC layers for detecting hardware Trojans [6], [84], [89]–[93]. Being segmented, hardware assurance measures performed at this stage can also be free of the influence of localized variation in pixel intensity and noise. It should also be noted that variants in Trojans, such as the parametric Trojans, can only be detected at this stage in RE [94]. The only disadvantage at this stage is the overhead in storing and processing full-scale images.

B. EXPERIMENT PROTOCOLS

The REFICS³ dataset consists of 800,000 synthetic SEM images. These are generated from a 32nm and 90nm AES design for the doping, polysilicon and metal layers. Each layer has 100,000 SEM images. A few example images are presented in Figure 18. In addition, every SEM image has a corresponding segmented ground truth and a layout-level mask. The intended usage of these images are defined by four protocols. These protocols work in tandem with the expanded workflow, shown in Figure 17, and are described below in detail.

- **Denosing Protocol:** This protocol transforms the raw SEM image into the denoised version of the image (Figure 17(a)→(b)). To obtain the ground truth denoised image, apply the mean intensity response of the materials in the image to the segmented ground truth available in the dataset. This information is provided in the dataset. PSNR and SSIM is used to evaluate the efficacy of denoising approaches.
- **Segmentation Protocol:** The segmentation protocol can be conducted in two ways. In the first approach, the raw SEM image is segmented directly (Figure 17(a)→(c)). In the second approach, the raw SEM image is denoised before it is segmented (Figure 17(a)→(b)→(c)). Denoising SEM images before segmentation typically yields better results. SSIM and IoU is used to evaluate

³Hosted on Trust-hub. Link: <https://trust-hub.org>

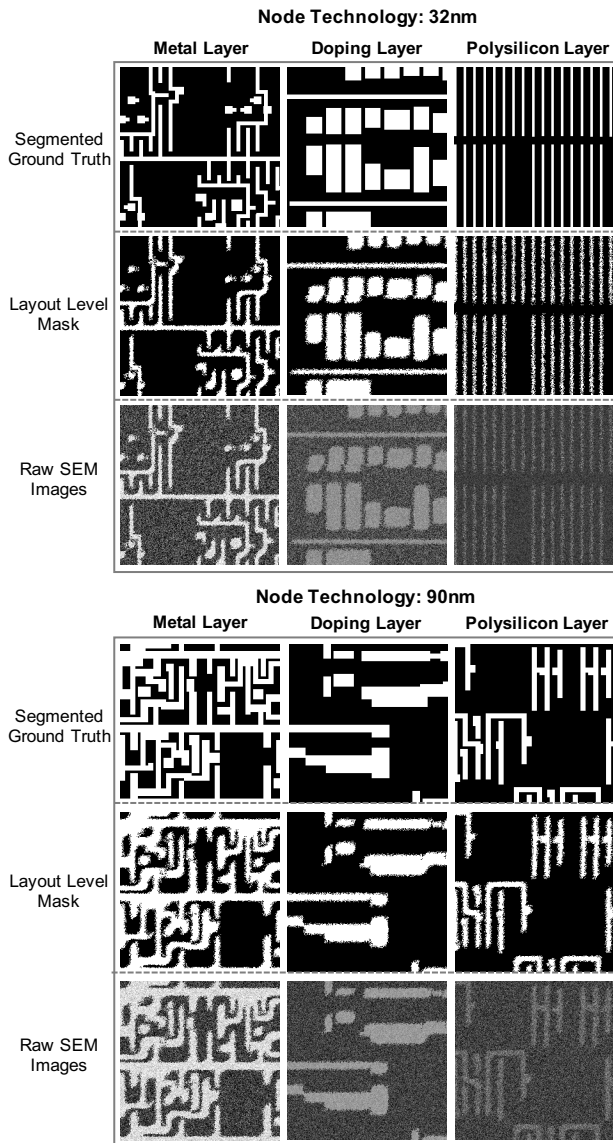


FIGURE 18: Examples of segmented ground truth, layout-level mask, and raw SEM images from the REFICS dataset.

the efficacy of segmentation approaches. We also suggest the use of the CC metric for evaluating connectivity between structures in the segmented image.

- **Vectorization Protocol:** This protocol utilizes the segmented ground truth image as input and produces the layout image (Figure 17(c)→(d)). Although no metrics are suggested in literature to evaluate the quality of the vectorized image, experience suggests that SSIM, IoU and CC can be utilized in this scenario.
- **End-to-end Protocol:** This protocol is typically utilized in DL approaches to achieve all the above steps in a unified architecture. The input is the raw SEM image and the output is expected to be the layout-like image (Figure 17(a)→(d)). Metrics used in segmentation protocol are applied.

The data provided in the dataset is intended for benchmarking novel image processing algorithms and developing neural network architectures. For approaches that involve very complex DL strategies or a directed purpose like handling stitching errors, a tool⁴ is made available for generating more SEM image samples. The tool assists in generating more SEM images and can be modified to generate SEM images with a particular error or set of errors depending on the user's intended application.

C. EXPERIMENT RESULTS

In this section, the performance of the image processing and ML methods are evaluated, scored using the metrics discussed earlier, on the dataset following the denoising, segmentation and end-to-end protocols. The key characteristic of a good algorithm is to score high on the chosen metrics and maintain stable scores across different layers and node technologies. The ability of the algorithm to maintain stable scores in these conditions is called cross-generalizability. In terms of ML, this is quintessentially required for supervised approaches that maintain some sort of heuristic on the design data from the labelled ground truth data provided to the algorithm. Consequently, the cross-generalizability of DL methods between layers and nodes is also investigated and discussed in detail.

A performance benchmark for various algorithms used in the image processing pipeline is critical for the introduction of data-driven paradigms. This can be attributed to the fact that every algorithm currently available in literature is evaluated on a small private dataset and their performance cannot be compared to each other. Further, concepts such as cross-generalizability of supervised algorithms have not been discussed in literature till date. Consequently, the benchmark results provides a quantitative baseline over which better data-driven algorithms can be built and tested. The benchmarks provided for various algorithms on the datasets use the parameters suggested in the source papers. Parameter fine-tuning was only performed for algorithms that require it in the original work. Exhaustive parameter fine-tuning and optimization for every algorithm is out of the scope of this paper.

1) Denoising Benchmark

Denoising assists in removing noise artifacts from the image. This entails differentiating between noise and signal. To actualize this, the denoising approaches use some assumption on the noise characteristics. The filters that make the closest assumption to the noise characteristic perform the best. Table 1 presents the performance of the aforementioned denoising methods, including five filtering and two ML methods. The key observation from the presented data is that the denoising performance reduces in the order: *Polysilicon layer* → *Doping layer* → *Metal layer*. In most cases, the metal layer shows reduction in image quality after denoising. This

⁴Also made available with the dataset

TABLE 1: Denoising algorithms used on SEM images. The reported numbers are improvement (in %) over original raw SEM images for that specific quality metric. The results are represented as PSNR / SSIM with higher values indicating better results. Negative values indicate degradation in image quality after denoising. The highest improvement in metrics for each layer and node technology is highlighted in bold.

Algorithm (↓) / Layers(→) Node (→)	Metal Layer		Doping Layer		Polysilicon Layer	
	32nm	90nm	32nm	90nm	32nm	90nm
Gaussian filter [68]	8.11 / 22.53	9.46 / 22.24	15.50 / 30.58	15.93 / 32.52	15.84 / 27.99	17.27 / 36.75
Anisotropic diffusion filter [68]	1.60 / 45.67	6.37 / 44.79	26.08 / 72.73	28.64 / 79.16	29.44 / 62.28	37.82 / 91.62
Curvature filter [68]	-28.56 / 29.55	-23.27 / 29.93	-14.65 / 50.02	1.36 / 67.16	18.41 / 52.47	32.46 / 84.95
Median filter	0.30 / 48.73	7.01 / 46.83	25.83 / 75.55	30.02 / 82.41	26.82 / 59.91	42.06 / 94.74
Adaptive Weiner filter [69]	-27.20 / -29.05	-21.09 / -12.42	-9.73 / 12.05	3.84 / 33.63	9.69 / 30.03	22.81 / 83.45
BM3D [70]	10.20 / 17.99	12.86 / 18.14	11.21 / 22.56	13.54 / 21.22	7.5 / 14.25	12.93 / 19.66
K-SVD [70]	12.52 / 37.95	15.32 / 64.73	23.07 / 49.35	16.00 / 64.27	22.47 / -9.65	23.73 / 65.88

TABLE 2: Segmentation algorithms utilized for segmenting IC SEM images. The results are represented as SSIM (↑) / IoU (↑) / CC-US (↓) / CC-OS (↓) scores. The direction of increasing quality for the metrics is indicated with ↑ or ↓. Apart from SVM, all other methods are unsupervised. K-means, Fuzzy C-means and HAS use a 5×5 kernel and SVM uses a 10×10 kernel. The highest improvement in metrics for each layer and node technology is highlighted in bold.

Algorithm (↓) / Layers(→) Node (→)	Metal Layer		Doping Layer		Polysilicon Layer	
	32nm	90nm	32nm	90nm	32nm	90nm
Otsu's thresholding [84]	0.77 / 0.88 / 0.11 / 0.91	0.79 / 0.91 / 0.13 / 0.69	0.55 / 0.73 / 0.38 / 0.77	0.27 / 0.49 / 0.29 / 0.61	0.40 / 0.52 / 0.64 / 0.69	0.12 / 0.29 / 0.80 / 0.53
Fuzzy C-means [76]	0.75 / 0.86 / 0.11 / 0.91	0.78 / 0.90 / 0.14 / 0.68	0.53 / 0.72 / 0.38 / 0.77	0.27 / 0.49 / 0.30 / 0.60	0.39 / 0.51 / 0.65 / 0.68	0.11 / 0.28 / 0.83 / 0.52
K-means [76]	0.77 / 0.88 / 0.11 / 0.91	0.79 / 0.91 / 0.13 / 0.69	0.55 / 0.73 / 0.38 / 0.77	0.27 / 0.49 / 0.29 / 0.60	0.40 / 0.52 / 0.64 / 0.69	0.12 / 0.29 / 0.81 / 0.53
HAS [81]	0.85 / 0.78 / 0.41 / 0.70	0.76 / 0.82 / 0.36 / 0.17	0.85 / 0.81 / 0.43 / 0.78	0.81 / 0.80 / 0.17 / 0.18	0.67 / 0.52 / 0.52 / 0.76	0.56 / 0.46 / 0.33 / 0.60
LASRE [63]	0.75 / 0.70 / 0.15 / 0.14	0.72 / 0.76 / 0.28 / 0.22	0.78 / 0.79 / 0.09 / 0.20	0.72 / 0.73 / 0.12 / 0.28	0.46 / 0.58 / 0.30 / 0.38	0.22 / 0.44 / 0.39 / 0.42
SVM-10 [76]	0.76 / 0.73 / 0.26 / 0.78	0.67 / 0.79 / 0.20 / 0.15	0.74 / 0.78 / 0.27 / 0.86	0.85 / 0.85 / 0.05 / 0.11	0.34 / 0.44 / 0.60 / 0.78	0.32 / 0.37 / 0.61 / 0.47

can be attributed to the fact that the contrast in the metal layer is much higher than those of other layers. The contrast is the lowest in the polysilicon layer and, hence, benefits the most from denoising. Anisotropic diffusion filters performs the best. This filter smooths the image while preserving the edges. With the hardware design layout being produced by straight edges, the performance metrics behind this filter can be intuitively understood. The Gaussian filter and Median filter performed relatively well. This is the reason behind several works in RE and hardware assurance supporting the utility of these filters. The ML-based denoising approaches performed poorly as compared to regular methods. Note that, Gaussian filter and BM3D performed consistently across all layers and node technologies.

2) Segmentation Benchmark

The baseline segmentation results shown in Table 2 were obtained by comparing the original and segmented SEM images (Figure 17 (a)→(c)). Denoising was not performed on the raw SEM image before segmentation. The key observation from the table is that the results are similar to the observations from the denoising experiments. A simple image binarization method like Otsu's thresholding has performance equivalent to that of ML approaches in the metal layer. Otsu's thresholding along with K-means and Fuzzy C-means also demonstrate stable performance across all layers and node technologies. HAS conserves more of the shape information in the segmented image while losing connectivity information. LASRE, on the other hand, preserves more connectivity information over shape information. The interesting observation in the table is that a supervised segmentation approach based on SVM performs similar to unsupervised methods despite having access to labelled ground truth data. The SVM was trained on 90,000 images and tested on 10,000

images of a single layer and node technology. Since one GT in the dataset may correspond to a couple of noisy raw SEM images, this splitting is chosen to guarantee the test set is independent from the trained models. The parameter for the SVM classifier was obtained from an earlier work [76]. As suggested by the author, cascading different classifiers or using a committee of classifiers will possibly yield better results than using an individual classifier. The performance metrics obtained on the metal layer and the doping layer succinctly explains the existence of several well-accepted hardware assurance approaches on these layers.

3) End-to-End Benchmarks

The end-to-end RE annotation module aims to achieve denoising, segmentation, and vectorization in one model. This fits into two computer vision tasks, namely, *image-to-image translation* and *blind denoising*. In this work, two deep neural networks from each task, specifically pix2pix [98], cycleGAN [97], DnCNN [95], and CBDNet [96], are evaluated. Their simplified network architectures are presented in Figure 19.

The *image-to-image translation* is used to convert an image from one representation into another [97]. In the RE domain, the raw SEM image and its corresponding layout GT or a noise-free image can be considered as two representations of one image. Under this assumption, the pix2pix network was used for SEM image quality enhancement [99], and cycleGAN was adopted to obtain translated SEM images that can serve as GT [54]. Pix2pix and cycleGAN are generative adversarial network (GAN), which consists of a pair/pairs of generator and discriminator. This architecture allows the network to train the image mapping by learning and minimizing a loss function simultaneously. These two GANs share the same network architecture, while the cycleGAN can be

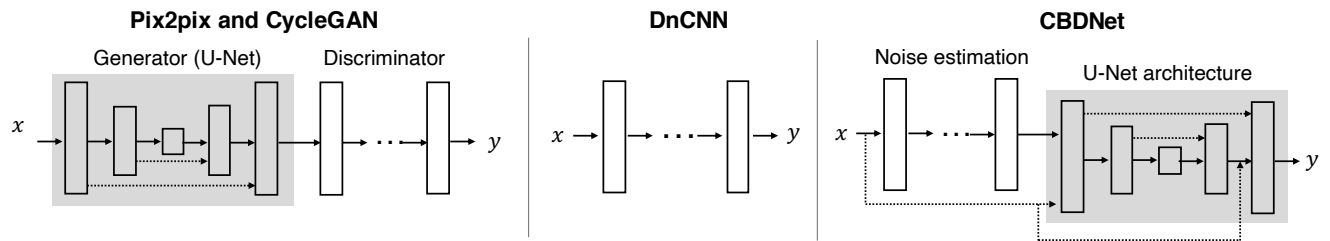


FIGURE 19: Simplified network architectures for DnCNN, CBDNet, pix2pix and cycleGAN. x and y represent the input and output, and each block represents a convolution layer with ReLU (or Leaky ReLU) or a convolution layer with batch normalization and ReLU.

trained using unpaired images and the pix2pix only learns the mapping from the paired ones. Since the cycleGAN follows this “unsupervised” training setting [98], it typically less accurate than pix2pix.

DL-based *blind denoising* was proposed for removing real-world image noise from photographs. DnCNN is one of the well-known discriminative models that can remove AWGN with an unknown noise level [95]. This network assumes that the noise mapping is easier to learn than the image details. Therefore, it learns the image residual (noise) by subtracting the latent clean images from the noisy input. This architecture has been successfully leveraged for EWR/LWR estimation on images with Gaussian-Poisson mixture noise [34], [38]. However, DnCNN is often criticized for easily overfitting to a specific noise model and cannot maintain the same performance on real noisy images. CBDNet was proposed to improve the generalizability by adopting noise estimation and non-blind denoising sub-networks [96]. The regularizer of the noise estimator prevents the over-smoothing of the features. The non-blind denoiser is an U-Net with skip connection, which is used to explore multi-scale features and generate clean images. Later proposed networks inspired by CBDNet achieved higher performance by increasing the network’s receptive field [100], [101]. However, to the best of our knowledge, such blind denoising networks were not evaluated in the context of RE and hardware assurance. Additionally, CBDNet was reported to be adept at preserving sharp edges -a preferable characteristic in RE applications. Thus, the results of CBDNet will serve as a blind image restoration benchmark for future improvements in the RE domain.

The baseline performance for the chosen networks are presented in Table 3. These results are obtained by training and testing on the same subset of one node technology and one layer in a 9:1 split ratio as done in the case of the SVM.

The training parameters for each network are adjusted for the best performance. DnCNN was trained with 70 epochs with a learning rate (α) of 10^{-5} and the batch size of 32. CBDNet was trained with 70 epochs with an α of 10^{-4} without batch normalization. Pix2pix and CycleGAN were trained with 50 epochs with an α of 2×10^{-4} and the batch size of 16. All networks use the Adam optimizer. Note that, network convergence was observed with the set number of epochs through learning curves. Since neural networks tend to produce non-binary intensity values (values other than 0 and 255) for the output image, they are binarized using threshold value 127.

The key observation from the baseline experiments is that most deep neural networks perform consistently on different layers. CBDNet performs the best with consistency. This can be attributed to multiple losses and the U-Net architecture, which was designed for segmentation. Additionally, comparing to other three networks, CBDNet adds connections between input and other layers, which provides more image details for reconstruction. Performance of DnCNN decreases on the polysilicon layer as compared to other layers. Since DnCNN learns noise solely by the differences between noisy and clean images, low contrast images like with the polysilicon layer, having noise pixels similar to the pixels representing clean images could cause denoising difficulties for DnCNN. Meanwhile, the CC-OS scores of DnCNN increases significantly indicating the network tends to remove pixels belonging to structural patterns. The two image translation networks also perform consistently. However, they show lower IoU scores. This may be due to the stitching errors getting transferred to the reconstructed images. Example output from baseline experiments is presented in Figure 20 (a) at the end of this section. Although failure or imperfect cases still exist, as shown in Figure 20 (b) and (c), they are mainly two types of errors that are less likely to affect overall

TABLE 3: Baseline performance of the end-to-end deep neural networks. The results are represented as SSIM (\uparrow) / IoU (\uparrow) / CC-US (\downarrow) / CC-OS (\downarrow) scores. The highest improvement in metrics for each layer and node technology is highlighted in bold.

Networks (\downarrow) / Layers (\rightarrow) Node (\rightarrow)	Metal Layer		Doping Layer		Polysilicon Layer		Averaged
	32nm	90nm	32nm	90nm	32nm	90nm	
DnCNN [95]	0.94 / 0.90 / 0.00 / 0.03	0.92 / 0.92 / 0.00 / 0.10	0.96 / 0.95 / 0.00 / 0.02	0.94 / 0.91 / 0.00 / 0.07	0.83 / 0.67 / 0.00 / 0.48	0.88 / 0.63 / 0.02 / 0.17	0.91 / 0.79 / 0.00 / 0.15
CBDNet [96]	0.96 / 0.94 / 0.00 / 0.03	0.96 / 0.95 / 0.01 / 0.04	0.98 / 0.97 / 0.00 / 0.00	0.98 / 0.96 / 0.00 / 0.01	0.95 / 0.93 / 0.00 / 0.02	0.96 / 0.87 / 0.00 / 0.03	0.97 / 0.94 / 0.00 / 0.02
CycleGAN [97]	0.88 / 0.85 / 0.00 / 0.01	0.72 / 0.70 / 0.01 / 0.04	0.86 / 0.84 / 0.00 / 0.03	0.76 / 0.71 / 0.00 / 0.01	0.90 / 0.85 / 0.00 / 0.07	0.67 / 0.74 / 0.01 / 0.04	0.80 / 0.74 / 0.01 / 0.04
Pix2pix [98]	0.93 / 0.89 / 0.00 / 0.01	0.78 / 0.74 / 0.01 / 0.02	0.96 / 0.90 / 0.00 / 0.03	0.95 / 0.72 / 0.00 / 0.01	0.90 / 0.87 / 0.00 / 0.07	0.91 / 0.62 / 0.02 / 0.04	0.90 / 0.79 / 0.01 / 0.03

* The numbers shown as 0.00 are rounded to two decimal places, they are not numerically equal to zero. This is also applied in the following tables.

TABLE 4: Cross-node generalizability results. The listed node technology represents the test set with the network trained on the other node. The results are represented as SSIM (\uparrow) / IoU (\uparrow) / CC-US (\downarrow) / CC-OS (\downarrow) scores. The highest improvement in metrics is highlighted in bold.

Networks (\downarrow) / Layers (\rightarrow) Node (\rightarrow)	Metal Layer		Doping Layer		Polysilicon Layer		Averaged
	32nm	90nm	32nm	90nm	32nm	90nm	
DnCNN	0.93 / 0.90 / 0.00 / 0.04	0.91 / 0.91 / 0.00 / 0.09	0.92 / 0.91 / 0.02 / 0.06	0.96 / 0.93 / 0.00 / 0.02	0.80 / 0.75 / 0.04 / 0.06	0.83 / 0.41 / 0.00 / 0.53	0.89 / 0.80 / 0.01 / 0.13
CBDNet	0.90 / 0.87 / 0.01 / 0.05	0.94 / 0.94 / 0.00 / 0.05	0.95 / 0.94 / 0.01 / 0.02	0.95 / 0.92 / 0.00 / 0.06	0.83 / 0.80 / 0.01 / 0.03	0.86 / 0.57 / 0.01 / 0.03	0.91 / 0.84 / 0.01 / 0.09
CycleGAN	0.73 / 0.70 / 0.04 / 0.03	0.75 / 0.70 / 0.00 / 0.05	0.86 / 0.84 / 0.00 / 0.04	0.71 / 0.65 / 0.00 / 0.06	0.66 / 0.61 / 0.03 / 0.25	0.66 / 0.41 / 0.03 / 0.30	0.73 / 0.65 / 0.02 / 0.12
Pix2pix	0.88 / 0.83 / 0.02 / 0.03	0.79 / 0.72 / 0.00 / 0.02	0.90 / 0.82 / 0.00 / 0.05	0.91 / 0.70 / 0.01 / 0.09	0.83 / 0.76 / 0.01 / 0.02	0.59 / 0.41 / 0.05 / 0.28	0.82 / 0.71 / 0.01 / 0.08

TABLE 5: Cross-layer generalizability results. The results are represented as SSIM (\uparrow) / IoU (\uparrow) / CC-US (\downarrow) / CC-OS (\downarrow) scores. The highest improvement in metrics is highlighted in bold.

Train set (\rightarrow)	Metal Layer			
Test set (\rightarrow)	Doping Layer		Polysilicon Layer	
Nodes	32nm	90nm	32nm	90nm
DnCNN	0.91 / 0.82 / 0.00 / 0.01	0.94 / 0.89 / 0.00 / 0.03	0.66 / 0.07 / 0.00 / 0.22	0.76 / 0.01 / 0.00 / 0.02
CBDNet	0.87 / 0.72 / 0.00 / 0.15	0.95 / 0.90 / 0.00 / 0.05	0.66 / 0.06 / 0.00 / 0.09	0.76 / 0.01 / 0.00 / 0.04
CycleGAN	0.85 / 0.83 / 0.00 / 0.09	0.68 / 0.63 / 0.00 / 0.02	0.51 / 0.53 / 0.12 / 0.18	0.26 / 0.22 / 0.20 / 0.29
Pix2pix	0.92 / 0.89 / 0.00 / 0.01	0.78 / 0.73 / 0.00 / 0.01	0.75 / 0.69 / 0.09 / 0.17	0.55 / 0.37 / 0.09 / 0.18

Train set (\rightarrow)	Doping Layer			
Test set (\rightarrow)	Metal Layer		Polysilicon Layer	
Nodes	32nm	90nm	32nm	90nm
DnCNN	0.91 / 0.89 / 0.01 / 0.04	0.85 / 0.87 / 0.03 / 0.32	0.78 / 0.47 / 0.00 / 0.38	0.76 / 0.05 / 0.00 / 0.09
CBDNet	0.88 / 0.82 / 0.01 / 0.13	0.91 / 0.91 / 0.02 / 0.08	0.76 / 0.60 / 0.00 / 0.17	0.78 / 0.17 / 0.00 / 0.14
CycleGAN	0.65 / 0.63 / 0.06 / 0.14	0.69 / 0.67 / 0.04 / 0.06	0.34 / 0.40 / 0.20 / 0.26	0.31 / 0.23 / 0.19 / 0.25
Pix2pix	0.81 / 0.70 / 0.04 / 0.24	0.79 / 0.57 / 0.02 / 0.20	0.43 / 0.32 / 0.31 / 0.15	0.77 / 0.23 / 0.06 / 0.17

Train set (\rightarrow)	Polysilicon Layer			
Test set (\rightarrow)	Metal Layer		Doping Layer	
Nodes	32nm	90nm	32nm	90nm
DnCNN	0.85 / 0.82 / 0.08 / 0.09	0.83 / 0.79 / 0.08 / 0.12	0.89 / 0.87 / 0.08 / 0.03	0.88 / 0.80 / 0.03 / 0.03
CBDNet	0.82 / 0.67 / 0.01 / 0.49	0.90 / 0.87 / 0.04 / 0.12	0.80 / 0.63 / 0.00 / 0.24	0.95 / 0.87 / 0.04 / 0.12
CycleGAN	0.75 / 0.55 / 0.05 / 0.52	0.64 / 0.55 / 0.05 / 0.24	0.68 / 0.46 / 0.01 / 0.60	0.72 / 0.58 / 0.00 / 0.09
Pix2pix	0.76 / 0.59 / 0.03 / 0.44	0.91 / 0.72 / 0.02 / 0.04	0.74 / 0.65 / 0.02 / 0.36	0.93 / 0.69 / 0.00 / 0.02

accuracy for RE and hardware assurance applications. The first is caused by low contrast, where the patterns in the noisy image can barely be seen, and it usually happens in images of the polysilicon layer. In real scenarios, images having such a low contrast are less likely to be collected. Another type is stitching error, and its influence depends on particular applications. For example, it may not affect Trojan detection methods that rely on representative features, but it may affect template matching and accuracy for netlist reconstruction. It maybe possible to address these issues using post-processing.

It is also observed that these four networks outperform conventional methods shown in Table 2. The best SSIM and IoU values are improved by 15% and 10%, respectively, with the best CC-US and CC-OS score approaching zero in all networks. Additionally, DL is more efficient once the network is trained. Using two GeForce 2080 Ti graphic cards, the training takes up to two days, while the testing time for each image only takes around 0.05 seconds. This is much faster than the conventional methods, which take around a second on average for each image. These observations reiterate the necessity of data-driven methods for RE and hardware assurance.

Nevertheless, as mentioned earlier, a major concern of applying deep neural networks is that they are easy to overfit and cannot generalize well. A generalizable network means it is only trained once but can achieve a similar performance on different testing sets i.e. different layers and node technolo-

gies. This saves efforts on data collections, computational resources, and time cost for adapting it on various RE and hardware assurance applications. In this work, cross-node and cross-layer generalizability are evaluated and discussed -a major contribution of this work.

Cross-node generalizability is presented in Table 4. The results are obtained by using the network trained on one node technology to test on another node technology for the same layer. For instance, using the network trained on the 32nm metal layer to test on the 90nm metal layer. The key observation is that four networks present different cross-node generalizability on different layers. Overall, CBDNet still achieves the best evaluation scores on most test sets, while DnCNN performs the most consistently. CBDNet, pix2pix, and cycleGAN cannot maintain a similar performance on cross-node testing for the polysilicon layers. These may be because the network is overfitted to the node-specified features of polysilicon layers. Note that, the four networks still perform better than the conventional methods on most testing sets.

Cross-layer generalizability is presented in Table 5. These results are obtained using the network trained on one layer to test another layer for the same node technology. For instance, using the network trained on the 32nm metal layer to test the 32nm doping layer of the same node technology. The key observation is that the networks trained on metal or doping layers easily fail in testing the polysilicon layers (cannot

generate correct structural patterns). Note that four metrics measure similarity from different aspects; a well-performed network should have good evaluation scores from all aspects. As shown in Table 5, the IoU scores of DnCNN and CBDNet in testing polysilicon layers decrease to lower than 0.1, and the CC-US and CC-OS scores of pix2pix and cycleGAN have increased when comparing to the baseline. The low contrast of the polysilicon layer may be the reason for this behaviour. Once the networks are fitted on images with high contrast, they can hardly recognize features from images having lower contrast. The analysis on failure cases shows that the average intensity difference between foreground and background on polysilicon layers is 40, while it is 160 for the metal layer training set. Similarly, it is 100 in the test set of doping layer, and a moderate decrease in the score is observed, accordingly.

4) Discussion and Insights

The benchmarks serve as a quantitative reminder over the type of algorithms that can be chosen to resolve any directed task in RE and RE-assisted hardware assurance. However, there are some key observations from the presented results that can be leveraged for the development of better algorithms and smoother integration of data-driven paradigms into RE.

The metrics commonly used in evaluating image quality and segmentation accuracy are not stable. There are several instances where the highest score in two metrics evaluating segmentation accuracy, in terms of shape for instance, goes to two different algorithms. Similarly, the methods that can achieve a high score on shape similarity measurement may not perform the same in terms of electrical connectivity. For example, in the metal layer, similar SSIM and IoU values are observed across nodes, while CC exhibits significant differences. To truly evaluate image quality, multiple metrics maybe necessary or a novel metric, specifically designed for RE and RE-based hardware assurance tasks, has to be developed.

A very interesting observation from the results is that most approaches show a lack of stability across node technologies and IC layers. Realizing the fact that the images are generated using the same beam interaction models with varying layouts, it is counter-intuitive for the algorithms to have variations in performance. The effect is compounded for the polysilicon layer. Even supervised methods with access to large quantities of high-quality labelled data shows this trend. This suggest that the approaches are not able to detect the edges in the original layout effectively. This effect can be clearly seen in the performance metrics reported, especially for the polysilicon layer with the lowest contrast among all three layers. An even more noteworthy observation is in Table 5. DL networks trained on the metal layer, a high-contrast image with relatively simple geometry, performed well on other layers especially in terms of separation between structures, i.e. the CC metric. However, when trained on the other two layers with structures having complex geometry,

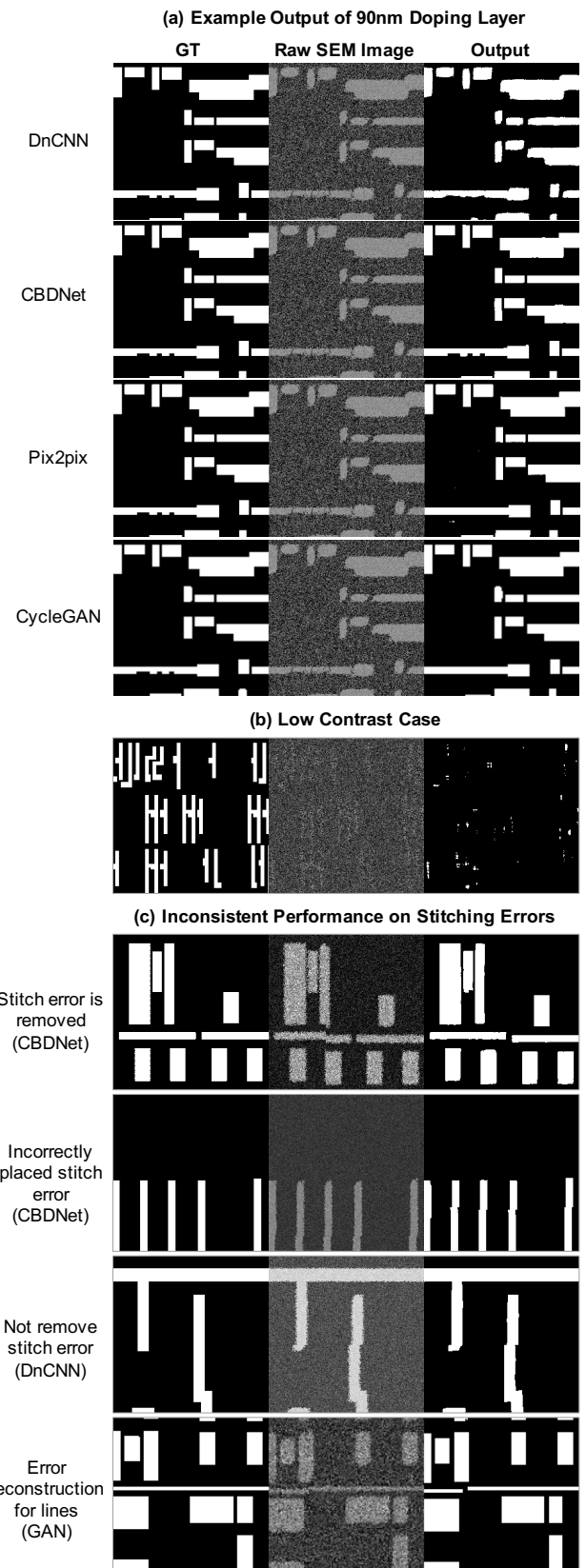


FIGURE 20: Example output from the end-to-end networks. Results in (a) are four outputs for the same testing sample.

they performed better in terms of conserving the shape of the structures, i.e. SSIM and IoU metrics. Although this doesn't affect the state-of-the-art performances provided by the DL models significantly, this does underline the fact that model architectures that are capable of resolving the edges between different materials under low contrast need to be developed. Off-the-shelf complex neural architectures may not be enough for hardware assurance applications. Supporting evidence can be found in a critical work that suggests that neural networks, especially those that work on images, are influenced more by the texture of the image than by the edges themselves [102]. Hence, more directed research is necessary for the development of effective neural network models. Furthermore, these observations do provide credence to the efficacy of template based models on the polysilicon layer [78].

With a correlation observed in between contrast and the efficacy of image analysis algorithms on resolving membership between materials and effective edge detection, the need for risk analysis models discussed in Section III-B is further reinforced. The correlation between the sensitivity index and the metrics, depicted using the Pearson correlation coefficient, is shown in Figure 21. Despite using multiple algorithms, IC layers and node technologies, the metrics consistently exhibited very strong correlation with the sensitivity index between the materials, i.e. the contrast of the image. Therefore, it can be reasoned that the higher the likelihood of correctly resolving the membership of a pixel, the lower the risk to the RE process irrespective of the steps followed. Along with the qualitative discussion in Section III-B, this provides a quantitative reasoning to support the central role of image contrast in feature recovery and risk assessment. Risk evaluation approaches like these are critical in quantifying the benefits of every step, even novel ad-hoc steps, in the RE process and transforming RE into a formalized, generalizable and repeatable process for hardware assurance and trust.

V. CONCLUSION

RE is a great tool for hardware assurance and trust. This is exhibited in its ability to discover well-placed stealthy hardware Trojans and verify the source design to discover IP infringement. The only limitations for the process were its likelihood for use as an attack mechanism and its resource-intensive nature. The first limitation was addressed in existing literature through design obfuscation techniques. A pathway to resolve the latter effectively is introduced in this paper.

Summarizing the paper, a large-scale SEM image dataset is introduced to support the integration of data-driven paradigms into the hardware assurance community. With a detailed taxonomy of challenges in RE and a tool to simulate these challenges, the dataset also provides an avenue for directed research into challenges for RE without the associated cost and resource overhead. Further, this also provides an opportunity for other communities, such as the image processing and computer vision communities, to get involved in the development of robust image processing and

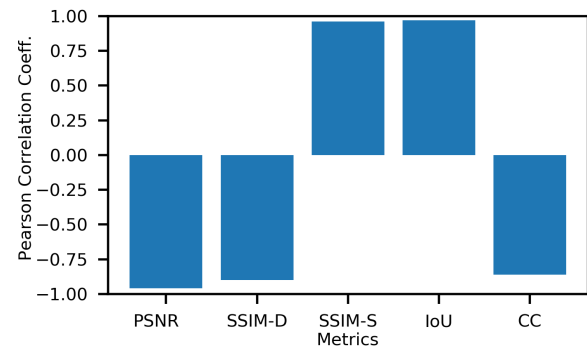


FIGURE 21: Plot indicating strong correlation between the sensitivity index between materials to the metrics irrespective of the algorithm used. The “D” and “S” in SSIM represent denoising and segmentation respectively. The PSNR and SSIM-D correlations were calculated as the improvement over the raw image and not the raw PSNR or SSIM values; hence, the negative correlation. CC represents both CC-OS and CC-US as they have identical values.

DL architectures for use in hardware assurance. The detailed benchmarks, especially the generalizability studies, provides the very insights required to facilitate this purpose.

Finally, the risk assessment framework introduced in the paper forms a common basis for comparison between works executed using ad-hoc steps. This serves as a viable approach for assessing the cost-benefit trends of any ad-hoc steps in the process and generates a basis of comparison for process efficacy. This framework along with the incorporation of data-driven paradigms can transform RE from a technique widely used in small-scale case studies to a versatile tool for effective and expedited hardware assurance.

REFERENCES

- [1] Kan Xiao, Domenic Forte, Yier Jin, Ramesh Karri, Swarup Bhunia, and Mohammad Tehranipoor. Hardware trojans: Lessons learned after one decade of research. *ACM Transactions on Design Automation of Electronic Systems (TODAES)*, 22(1):1–23, 2016.
- [2] Seetharam Narasimhan, Xinmu Wang, Dongdong Du, Rajat Subhra Chakraborty, and Swarup Bhunia. Tesr: A robust temporal self-referencing approach for hardware trojan detection. In *2011 IEEE International Symposium on Hardware-Oriented Security and Trust*, pages 71–74. IEEE, 2011.
- [3] Rajat Subhra Chakraborty, Francis Wolff, Somnath Paul, Christos Papachristou, and Swarup Bhunia. Mero: A statistical approach for hardware trojan detection. In *International Workshop on Cryptographic Hardware and Embedded Systems*, pages 396–410. Springer, 2009.
- [4] Mohammad Tehranipoor and Farinaz Koushanfar. A survey of hardware trojan taxonomy and detection. *IEEE design & test of computers*, 27(1):10–25, 2010.
- [5] Ulbert J Botero, Ronald Wilson, Hangwei Lu, Mir Tanjidur Rahman, Mukhil A Mallaiyan, Fatemeh Ganji, Navid Asadizanjani, Mark M Tehranipoor, Damon L Woodard, and Domenic Forte. Hardware trust and assurance through reverse engineering: A tutorial and outlook from image analysis and machine learning perspectives. *J. Emerg. Technol. Comput. Syst.*, 17(4), 2021.
- [6] Franck Courbon, Philippe Loubet-Moundi, Jacques JA Fournier, and Asia Tria. A high efficiency hardware trojan detection technique based on fast sem imaging. In *Proceedings of the 2015 Design, Automation & Test in Europe Conference & Exhibition*, pages 788–793. EDA Consortium, 2015.

- [7] Chongxi Bao, Domenic Forte, and Ankur Srivastava. On reverse engineering-based hardware trojan detection. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 35(1):49–57, 2016.
- [8] Domenic Forte, Swarup Bhunia, and Mark M Tehranipoor. *Hardware protection through obfuscation*. Springer, 2017.
- [9] Muhammad Yasin, Jeyavijayan JV Rajendran, and Ozgur Sinanoglu. *Trustworthy Hardware Design: Combinational Logic Locking Techniques*. Number 1. Springer, 2019.
- [10] Bicky Shakya, Haoting Shen, Mark Tehranipoor, and Domenic Forte. Covert gates: Protecting integrated circuits with undetectable camouflaging. *IACR Transactions on Cryptographic Hardware and Embedded Systems*, pages 86–118, 2019.
- [11] Ronald P Cocchi, James P Baukus, Lap Wai Chow, and Bryan J Wang. Circuit camouflage integration for hardware ip protection. In *Proceedings of the 51st Annual Design Automation Conference*, pages 1–5. ACM, 2014.
- [12] Jeyavijayan Rajendran, Ozgur Sinanoglu, and Ramesh Karri. Vlsi testing based security metric for ic camouflaging. In *2013 IEEE International Test Conference (ITC)*, pages 1–4. IEEE, 2013.
- [13] Hector Gomez, Kkristian Duran, and Elkim Roa. Standard cell camouflage method to counter silicon reverse engineering. In *2018 IEEE International Conference on Consumer Electronics (ICCE)*, pages 1–4. IEEE, 2018.
- [14] Hector Gomez, Kkristian Duran, and Elkim Roa. Defeating silicon reverse engineering using a layout-level standard cell camouflage. *IEEE Transactions on Consumer Electronics*, 65(1):109–118, 2019.
- [15] Randy Torrance and Dick James. The state-of-the-art in ic reverse engineering. In *International Workshop on Cryptographic Hardware and Embedded Systems*, pages 363–381. Springer, 2009.
- [16] Shahed E Quadir, Junlin Chen, Domenic Forte, Navid Asadizanjani, Sina Shahbazmohamadi, Lei Wang, John Chandy, and Mark Tehranipoor. A survey on chip to system reverse engineering. *ACM journal on emerging technologies in computing systems (JETC)*, 13(1):6, 2016.
- [17] Girishma Jain, Sandeep Raghuvanshi, and Gagan Vishwakarma. Hardware trojan: Malware detection using reverse engineering and svm. In *International Conference on Intelligent Systems Design and Applications*, pages 530–539. Springer, 2017.
- [18] Esha Sarkar and Michail Maniatakos. On automating delayed ic analysis for hardware ip protection. In *Proceedings of the International Conference on Omni-Layer Intelligent Systems*, pages 205–210. ACM, 2019.
- [19] Adam Kimura, Jon Scholl, James Schaffranek, Matthew Sutter, Andrew Elliott, Mike Strizich, and Glen David Via. A decomposition workflow for integrated circuit verification and validation. *Journal of Hardware and Systems Security*, pages 1–10, 2020.
- [20] Reza Shokri and Vitaly Shmatikov. Privacy-preserving deep learning. In *Proceedings of the 22nd ACM SIGSAC conference on computer and communications security*, pages 1310–1321, 2015.
- [21] Zekeriya Erkin, Martin Franz, Jorge Guajardo, Stefan Katzenbeisser, Inald Lagendijk, and Tomas Toft. Privacy-preserving face recognition. In *International symposium on privacy enhancing technologies symposium*, pages 235–253. Springer, 2009.
- [22] Rakesh Agrawal and Ramakrishnan Srikant. Privacy-preserving data mining. In *Proceedings of the 2000 ACM SIGMOD international conference on Management of data*, pages 439–450, 2000.
- [23] Yongqin Xian, Christoph H Lampert, Bernt Schiele, and Zeynep Akata. Zero-shot learning—a comprehensive evaluation of the good, the bad and the ugly. *IEEE transactions on pattern analysis and machine intelligence*, 41(9):2251–2265, 2018.
- [24] Elyor Kodirov, Tao Xiang, and Shaogang Gong. Semantic autoencoder for zero-shot learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3174–3183, 2017.
- [25] Bernardino Romera-Paredes and Philip Torr. An embarrassingly simple approach to zero-shot learning. In *International Conference on Machine Learning*, pages 2152–2161, 2015.
- [26] Stephen Kockentiedt, Klaus Tönnies, Erhardt Gierke, Nico Dziurawitz, Carmen Thim, and Sabine Plitzko. Poisson shot noise parameter estimation from a single scanning electron microscopy image. In *Image Processing: Algorithms and Systems XI*, volume 8655, page 86550N. International Society for Optics and Photonics, 2013.
- [27] James Pawley and Heide Schatten. *Biological low-voltage scanning electron microscopy*. Springer, 2007.
- [28] Luděk Frank. Noise in secondary electron emission: the low yield case. *Journal of electron microscopy*, 54(4):361–365, 2005.
- [29] Makoto Sakakibara, Makoto Suzuki, Kenji Tanimoto, Yasunari Sohda, Daisuke Bizen, and Koji Nakamae. Impact of secondary electron emission noise in sem. *Microscopy*, 68(4):279–288, 2019.
- [30] KS Sim, JTL Thong, and JCH Phang. Effect of shot noise and secondary emission noise in scanning electron microscope images. *Scanning: The Journal of Scanning Microscopies*, 26(1):36–40, 2004.
- [31] Christopher F Batten. *Autofocusing and astigmatism correction in the scanning electron microscope*. Mphil thesis, University of Cambridge, 2000.
- [32] F Timischl, M Date, and S Nemoto. A statistical model of signal–noise in scanning electron microscopy. *Scanning*, 34(3):137–144, 2012.
- [33] Petr Cizmar, András E Vladár, Bin Ming, Michael T Postek, National Institute of Standards, and Technology. Simulated sem images for resolution measurement. *Scanning*, 30(5):381–391, 2008.
- [34] Narendra Chaudhary, Serap A Savari, and Sai S Yeddulapalli. Line roughness estimation and poisson denoising in scanning electron microscope images using deep learning. *Journal of Micro/Nanolithography, MEMS, and MOEMS*, 18(2):024001, 2019.
- [35] EL Principe, Navid Asadizanjani, Domenic Forte, Mark Tehranipoor, Robert Chivas, Michael DiBattista, Scott Silverman, Mike Marsh, Nicolas Piche, and John Mastovich. Steps toward automated deprocessing of integrated circuits. In *ISTFA 2017: the 43rd International Symposium for Testing and Failure Analysis*, page 285. ASM International, 2017.
- [36] Sebastian Wallat, Nils Albartus, Steffen Becker, Max Hoffmann, Maik Ender, Marc Fyrbiak, Adrian Drees, Sebastian Maaßen, and Christof Paar. Highway to hal: open-sourcing the first extendable gate-level netlist reverse engineering framework. In *Proceedings of the 16th ACM International Conference on Computing Frontiers*, pages 392–397, 2019.
- [37] H Seiler. Secondary electron emission in the scanning electron microscope. *Journal of Applied Physics*, 54(11):R1–R18, 1983.
- [38] E Giannatou, G Papaveros, V Constantoudis, H Papageorgiou, and E Gogolides. Deep learning denoising of sem images towards noise-reduced ler measurements. *Microelectronic Engineering*, 216:111051, 2019.
- [39] LR Harriott, A Wagner, and F Fritz. Integrated circuit repair using focused ion beam milling. *Journal of Vacuum Science & Technology B: Microelectronics Processing and Phenomena*, 4(1):181–184, 1986.
- [40] Michael T Postek and AE Vladar. Is low accelerating voltage always the best for semiconductor inspection and metrology? *Microscopy and Microanalysis*, 9(S02):978–979, 2003.
- [41] Raul Quijada, Roger Dura, Jofre Pallares, Xavier Formatje, Salvador Hidalgo, and Francisco Serra-Graells. Large-area automated layout extraction methodology for full-ic reverse engineering. *Journal of Hardware and Systems Security*, 2(4):322–332, 2018.
- [42] Jin Won Koh, Gu Teak Hwang, Moon Seop Hyun, Jun-Mo Yang, and Jeoung Woo Kim. Semiconductor layer extraction techniques by sem. In *18th IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA)*, pages 1–3. IEEE, 2011.
- [43] Xuenong Hong, Deruo Cheng, Yiqiong Shi, Tong Lin, and Bah Hwee Gwee. Deep learning for automatic ic image analysis. In *2018 IEEE 23rd International Conference on Digital Signal Processing (DSP)*, pages 1–5. IEEE, 2018.
- [44] Bernhard Lippmann, Michael Werner, Niklas Unverricht, Aayush Singla, Peter Egger, Anja Dübotzky, Horst Gieser, Martin Rasche, Oliver Kellermann, and Helmut Graeb. Integrated flow for reverse engineering of nanoscale technologies. In *Proceedings of the 24th Asia and South Pacific Design Automation Conference*, pages 82–89. ACM, 2019.
- [45] R. Quijada, A. Raventós, F. Tarrés, R. Durà, and S. Hidalgo. The use of digital image processing for ic reverse engineering. In *2014 IEEE 11th International Multi-Conference on Systems, Signals Devices (SSD14)*, pages 1–4, 2014.
- [46] Ronald Wilson, Rabin Y Acharya, Domenic Forte, Navid Asadizanjani, and Damon Woodard. A novel approach to unsupervised automated extraction of standard cell library for reverse engineering and hardware assurance. In *ISTFA 2019: 45th International Symposium for Testing and Failure Analysis*, page 249. ASM International, 2019.
- [47] David Zhang, Gooitzen van der Wal, Phil Miller, David Stoker, Erik Matlin, Naveen Marri, Gary Gan, Joe Zhang, Jane Asmuth, Sek Chai, et al. Fast, full chip image stitching of nanoscale integrated circuits. Technical report, SRI International Princeton United States, 2019.
- [48] Tong Lin, Yiqiong Shi, Na Shu, Deruo Cheng, Xuenong Hong, Jingsi Song, and Bah Hwee Gwee. Deep learning-based image analysis

- framework for hardware assurance of digital integrated circuits. In 2020 IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA), pages 1–6. IEEE, 2020.
- [49] Marek Sikul, Karel Novotny, Matthias Kemmler, and Andreas Rummel. Sem-based nanoprobing on in-situ delayered advanced 10 nm technology node ic. In 2018 IEEE International Symposium on the Physical and Failure Analysis of Integrated Circuits (IPFA), pages 1–4. IEEE, 2018.
- [50] Sean P Frigo, Zachary H Levine, and Nestor J Zaluzec. Submicron imaging of buried integrated circuit structures using scanning confocal electron microscopy. *Applied Physics Letters*, 81(11):2112–2114, 2002.
- [51] DW Malone and RE Hummel. Electromigration in integrated circuits. *Critical Reviews in Solid State and Material Sciences*, 22(3):199–238, 1997.
- [52] Marc Fyrbiak, Sebastian Strauß, Christian Kison, Sebastian Wallat, Malte Elson, Nikol Rummel, and Christof Paar. Hardware reverse engineering: Overview and open challenges. In 2017 IEEE 2nd International Verification and Security Workshop (IVSW), pages 88–94. IEEE, 2017.
- [53] Petr Cizmar, András E Vladár, and Michael T Postek. Optimization of accurate sem imaging by use of artificial images. In *Scanning Microscopy 2009*, volume 7378, page 737815. International Society for Optics and Photonics, 2009.
- [54] Hao-Chiang Shao, Chao-Yi Peng, Jun-Rei Wu, Chia-Wen Lin, Shao-Yun Fang, Pin-Yen Tsai, and Yan-Hsiu Liu. From ic layout to die photo: A cnn-based data-driven approach. *arXiv preprint arXiv:2002.04967*, 2020.
- [55] Patrick Trampert, Faysal Bourghorbel, Pavel Potocek, Maurice Peemen, Christian Schlinkmann, Tim Dahmen, and Philipp Slusallek. How should a fixed budget of dwell time be spent in scanning electron microscopy to optimize image quality? *Ultramicroscopy*, 191:11–17, 2018.
- [56] R Goldman, K Bartleson, T Wood, K Kranen, V Melikyan, and E Babayan. 32/28nm educational design kit: Capabilities, deployment and future. In 2013 IEEE Asia Pacific Conference on Postgraduate Research in Microelectronics and Electronics (PrimeAsia), pages 284–288. IEEE, 2013.
- [57] Richard Goldman, Karen Bartleson, Troy Wood, Kevin Kranen, C Cao, Vazgen Melikyan, and Gayane Markosyan. Synopsys’ open educational design kit: capabilities, deployment and future. In 2009 IEEE International Conference on Microelectronic Systems Education, pages 20–24. IEEE, 2009.
- [58] Bernhard Lippmann, Niklas Unverricht, Aayush Singla, Matthias Ludwig, Michael Werner, Peter Egger, Anja Duebotzky, Helmut Graeb, Horst Gieser, Martin Rasche, et al. Verification of physical designs using an integrated reverse engineering flow for nanoscale technologies. *Integration*, 71:11–29, 2020.
- [59] Franck Courbon, Sergei Skorobogatov, and Christopher Woods. Reverse engineering flash eeprom memories using scanning electron microscopy. In *International Conference on Smart Card Research and Advanced Applications*, pages 57–72. Springer, 2016.
- [60] Franck Courbon, Philippe Loubet-Moundi, Jacques JA Fournier, and Assia Tria. Samba: A sem based acquisition technique for fast invasive hardware trojan detection. In 2015 European Conference on Circuit Theory and Design (ECCTD), pages 1–4. IEEE, 2015.
- [61] Franck Courbon. Practical partial hardware reverse engineering analysis. *Journal of Hardware and Systems Security*, pages 1–10, 2019.
- [62] Anna Lena Keller, Dirk Zeidler, and Thomas Kemen. High throughput data acquisition with a multi-beam sem. In *Scanning Microscopies 2014*, volume 9236, page 92360B. International Society for Optics and Photonics, 2014.
- [63] Ronald Wilson, Domenic Forte, Navid Asadizanjani, and Damon L Woodard. Lasre: A novel approach to large area accelerated segmentation for reverse engineering on sem images. In *ISTFA 2020: Papers Accepted for the Planned 46th International Symposium for Testing and Failure Analysis*, pages 180–187. ASM International, 2020.
- [64] Karsten Nohl, David Evans, Starbug Starbug, and Henryk Plötz. Reverse-engineering a cryptographic rfid tag. In *USENIX security symposium*, volume 28, 2008.
- [65] Roger Durà, Jofre Pallarès, Raúl Quijada, Xavier Formatjé, Salvador Hidalgo, and Francisco Serra-Graells. fast and robust topology-based logic gate identification for automated ic reverse engineering. In *ISTFA 2017: the 43rd International Symposium for Testing and Failure Analysis*, page 299. ASM International, 2017.
- [66] G Masalskis et al. Reverse engineering of cmos integrated circuits. *Elektronika ir elektrotechnika*, 88(8):25–28, 2008.
- [67] N. Arazm, A. Sahab, and M. F. Kazemi. Noise reduction of sem images using adaptive wiener filter. In 2017 IEEE International Conference on Cybernetics and Computational Intelligence (CyberneticsCom), pages 50–55, 2017.
- [68] Bruno Machado Trindade, Eranga Ukwatta, Mike Spence, and Chris Pawlowicz. Segmentation of integrated circuit layouts from scan electron microscopy images. In 2018 IEEE Canadian Conference on Electrical & Computer Engineering (CCECE), pages 1–4. IEEE, 2018.
- [69] Mohadeseh Mazhari and Reza PR Hasanzadeh. Suppression of noise in sem images using weighted local hysteresis smoothing filter. *Scanning*, 38(6):634–643, 2016.
- [70] A Lazar and Petru S Fodor. Sparsity based noise removal from low dose scanning electron microscopy images. In *Computational Imaging XIII*, volume 9401, page 940105. International Society for Optics and Photonics, 2015.
- [71] Myungjun Lee, Jason Cantone, Ji Xu, Lei Sun, and Ryoung-han Kim. Improving sem image quality using pixel super resolution technique. In *Metrology, Inspection, and Process Control for Microlithography XXVIII*, volume 9050, page 90500U. International Society for Optics and Photonics, 2014.
- [72] KS Sim, ME Nia, and CP Tso. Image noise cross-correlation for signal-to-noise ratio estimation in scanning electron microscope images. *Scanning*, 33(2):82–93, 2011.
- [73] JTL Thong, KS Sim, and JCH Phang. Single-image signal-to-noise ratio estimation. *Scanning*, 23(5):328–336, 2001.
- [74] KS Sim, ME Nia, and Chih Ping Tso. Noise variance estimation using image noise cross-correlation model on sem images. *Scanning*, 35(3):205–212, 2013.
- [75] Nidal S Kamel and KS Sim. Image signal-to-noise ratio and noise variance estimation using autoregressive model. *Scanning: The Journal of Scanning Microscopies*, 26(6):277–281, 2004.
- [76] Deroo Cheng, Yiqiong Shi, Tong Lin, Bah-Hwee Gwee, and Kar-Ann Toh. Hybrid k-means clustering and support vector machine method for via and metal line detections in delayered ic images. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 65(12):1849–1853, 2018.
- [77] Deroo Cheng, Yiqiong Shi, Bah-Hwee Gwee, Kar-Ann Toh, and Tong Lin. A hierarchical multiclassifier system for automated analysis of delayered ic images. *IEEE Intelligent Systems*, 34(2):36–43, 2018.
- [78] Deroo Cheng, Yiqiong Shi, Tong Lin, Bah-Hwee Gwee, and Kar-Ann Toh. Global template projection and matching method for training-free analysis of delayered ic images. In 2019 IEEE International Symposium on Circuits and Systems (ISCAS), pages 1–5. IEEE, 2019.
- [79] Piali Das, Olga Veksler, Vyacheslav Zavadsky, and Yuri Boykov. Semi-automatic segmentation with compact shape prior. *Image and Vision Computing*, 27(1-2):206–219, 2009.
- [80] Alexander Doudkin, Alexander Inyutin, and Maksim Vatin. Objects identification on the color layout images of the integrated circuit layers. In 2005 IEEE Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, pages 610–614. IEEE, 2005.
- [81] Ronald Wilson, Navid Asadizanjani, Domenic Forte, and Damon L Woodard. Histogram-based auto segmentation: A novel approach to segmenting integrated circuit structures from sem images. *arXiv preprint arXiv:2004.13874*, 2020.
- [82] Jang Hee Lee and Suk In Yoo. An effective image segmentation technique for the sem image. In 2008 IEEE international conference on industrial technology, pages 1–5. IEEE, 2008.
- [83] Dmitry Lagunovsky, Sergey Ablameyko, and M Kutas. Recognition of integrated circuit images in reverse engineering. In *Proceedings. Fourteenth International Conference on Pattern Recognition (Cat. No. 98EX170)*, volume 2, pages 1640–1642. IEEE, 1998.
- [84] Nidish Vashistha, M Tanjidur Rahman, Haoting Shen, Damon L Woodard, Navid Asadizanjani, and Mark Tehranipoor. Detecting hardware trojans inserted by untrusted foundry using physical inspection and advanced image processing. *Journal of Hardware and Systems Security*, 2(4):333–344, 2018.
- [85] Jon R Pearce and Duane C Holmes. Noise contributions to feature dimension measurement in a scanning electron microscope (sem). In *Integrated Circuit Metrology, Inspection, & Process Control*, volume 775, pages 153–158. International Society for Optics and Photonics, 1987.
- [86] Vassilios Constantoudis, Vijaya-Kumar Murugesan Kuppuswamy, and Evangelos Gogolides. Effects of image noise on contact edge roughness and critical dimension uniformity measurement in synthesized scanning electron microscope images. *Journal of Micro/Nanolithography, MEMS, and MOEMS*, 12(1):013005, 2013.
- [87] Yoshihiro Midoh, Katsuyoshi Miura, Koji Nakamae, and Hiromu Fujioka. Statistical optimization of canny edge detector for measurement of

- fine line patterns in sem image. *Measurement Science and Technology*, 16(2):477, 2005.
- [88] Simon Blythe, Beatrice Fraboni, Sanjay Lall, Haroon Ahmed, and Ugo de Riu. Layout reconstruction of complex silicon chips. *IEEE journal of solid-state circuits*, 28(2):138–145, 1993.
- [89] Chongxi Bao, Domenic Forte, and Ankur Srivastava. On application of one-class svm to reverse engineering-based hardware trojan detection. In *Fifteenth International Symposium on Quality Electronic Design*, pages 47–54. IEEE, 2014.
- [90] Chongxi Bao, Domenic Forte, and Ankur Srivastava. On reverse engineering-based hardware trojan detection. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 35(1):49–57, 2015.
- [91] Abdurrahman A Nasr and Mohamed Z Abdulmageed. An efficient reverse engineering hardware trojan detector using histogram of oriented gradients. *Journal of Electronic Testing*, 33(1):93–105, 2017.
- [92] Nidish Vashistha, Hangwei Lu, Qihang Shi, M Tanjidur Rahman, Haoting Shen, Damon L Woodard, Navid Asadizanjani, and Mark Tehranipoor. Trojan scanner: Detecting hardware trojans with rapid sem imaging combined with image processing and machine learning. In *ISTFA 2018: the 44th International Symposium for Testing and Failure Analysis*, page 256. ASM International, 2018.
- [93] Qihang Shi, Nidish Vashistha, Hangwei Lu, Haoting Shen, Bahar Tehranipoor, Damon L Woodard, and Navid Asadizanjani. Golden gates: A new hybrid approach for rapid hardware trojan detection using testing and imaging. In *2019 IEEE International Symposium on Hardware Oriented Security and Trust (HOST)*, pages 61–71. IEEE, 2019.
- [94] Yier Jin and Yiorgos Makris. Hardware trojan detection using path delay fingerprint. In *2008 IEEE International workshop on hardware-oriented security and trust*, pages 51–57. IEEE, 2008.
- [95] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, 2017.
- [96] Shi Guo, Zifei Yan, Kai Zhang, Wangmeng Zuo, and Lei Zhang. Toward convolutional blind denoising of real photographs. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1712–1722, 2019.
- [97] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1125–1134, 2017.
- [98] Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pages 2223–2232, 2017.
- [99] Yoshihiro Midoh and Koji Nakamae. Image quality enhancement of a cd-sem image using conditional generative adversarial networks. In *Metrology, Inspection, and Process Control for Microlithography XXXIII*, volume 10959, page 109590B. International Society for Optics and Photonics, 2019.
- [100] Saeed Anwar and Nick Barnes. Real image denoising with feature attention. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 3155–3164, 2019.
- [101] Yoonsik Kim, Jae Woong Soh, Gu Yong Park, and Nam Ik Cho. Transfer learning from synthetic to real-noise denoising with adaptive instance normalization. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3482–3492, 2020.
- [102] Robert Geirhos, Patricia Rubisch, Claudio Michaelis, Matthias Bethge, Felix A Wichmann, and Wieland Brendel. Imagenet-trained cnns are biased towards texture; increasing shape bias improves accuracy and robustness. *arXiv preprint arXiv:1811.12231*, 2018.



RONALD WILSON received his B.E in Aeronautical Engineering from the Riga Technical University, Riga, Latvia and M.S in Electrical and Computer Engineering from the University of Florida, Gainesville, FL, USA in 2015 and 2017 respectively. He is currently a doctoral candidate, under Dr. Damon L. Woodard, with the Electrical and Computer Engineering Department at the University of Florida. His research interests include Image Processing and Computer Vision, Machine Learning, Deep Learning, Biometrics, Reverse Engineering, and, Natural Language Processing for Identity Sciences and Forensics.



HANGWEI LU received her M.S in electrical and computer engineering from the University of Florida in 2017. She is currently pursuing the Ph.D. in electrical and computer engineering in the Florida institute for cybersecurity research at the University of Florida. Her research interests include computer vision and machine learning, hardware assurance, and biometrics identification.



MENGDİ ZHU received the B.S degree in Aeronautical and Astronautical Engineering from Purdue University in 2015 and the M.S. degree in Electrical and Computer Engineering from University of Florida in 2019. She is currently pursuing her PhD degree in Electrical and Computer Engineering department at University of Florida. Her research interest is mainly focus on deep learning and its applications.



DOMENIC FORTE (S'09–M'13–SM'18) received the B.S. degree from the Manhattan College, Riverdale, NY, USA, in 2006, and the M.S. and Ph.D. degrees from the University of Maryland at College Park, College Park, MD, USA, in 2010 and 2013, respectively, all in electrical engineering. He is currently an Associate Professor with the Electrical and Computer Engineering Department, University of Florida, Gainesville, FL, USA. His research interests include the domain of hardware security, including the investigation of hardware security primitives, hardware Trojan detection and prevention, electronics supply-chain security, and anti-reverse engineering. He was a recipient of the Presidential Early Career Award for Scientists and Engineers (PECASE), the Early Career Award for Scientists and Engineers (ECASE) by the Army Research Office (ARO), the NSF Faculty Early Career Development Program (CAREER) Award, and the ARO Young Investigator Award. His research has also been recognized through multiple best paper awards and nominations.



DAMON L. WOODARD (S'01-M'04-SM'11) received his Ph.D. in Computer Science and Engineering from the University of Notre Dame in 2005. He obtained his M.E. from Pennsylvania State University in 1999 and the B.S. in Computer Science and Computer Information Systems from Tulane University in 1997. He is currently an Associate Professor in the Electrical and Computer Engineering Department at the University of Florida. He is also a member of the Florida Institute of Cybersecurity (FICS) Research and a member of NAVY CRANE's Computer Vision and Machine Learning for Hardware Assurance Working Group. His research interests include image analysis/computer vision for hardware assurance, artificial intelligence, machine learning, and pattern recognition. Dr. Woodard is a Senior Member of the Association for Computing Machinery (ACM), a member of the Association for the Advancement of Artificial Intelligence (AAAI), and a member of the National Society of Black Engineers (NSBE).

...