# Enhanced Photovoltaic Defect Detection Using Perceptual Loss in DCGAN and VGG16-Integrated Models on Electroluminescence Images

Nadia Drir , Adel Mellit , Maamar Bettayeb , and Mahmoud Dhimish

*Abstract*—Defects in photovoltaic (PV) modules significantly impact their efficiency and reliability, making the accurate detection essential for quality control. This study presents an enhanced generative adversarial network framework, combining deep convolutional generative adversarial network (DCGAN) with visual geometry group 16 (VGG16) and a perceptual loss function, to generate high-quality synthetic defect images and improve defect classification. The proposed model increases classification accuracy from 84% to 90%, demonstrating superior performance over the standard DCGAN. Key improvements include generating more realistic synthetic images, reducing image quality discrepancies, and addressing class imbalances in defect datasets. The enhanced framework performs particularly well in representing rare and complex defects, improving classification outcomes for challenging patterns while maintaining diversity and realism in synthetic data. Despite some limitations in capturing fine defect details, this approach establishes a benchmark for defect detection and synthetic image generation in PV manufacturing. It also offers potential for real-time applications in industrial production lines, ensuring more efficient quality control.

*Index Terms*—Deep convolutional generative adversarial network (DCGAN), DCGAN with VGG16, electroluminescence (EL) image, faults diagnosis, perceptual loss, solar cell.

## I. INTRODUCTION

**T**HE photovoltaic (PV) growing deployment of PV technology is largely fuelled by the rising demand for sustainable and renewable energy sources, particularly in the context to support worldwide initiatives aimed at lowering carbon emissions and shifting toward cleaner energy solutions [1]. This has resulted in the rapid global expansion of solar PV energy in recent years, encompassing both large-scale PV farms and distributed generation systems. Ensuring quality control during the manufacturing process of solar cells is paramount for achieving optimal performance and longevity of PV modules. High-quality production guarantees that each cell operates at peak efficiency and can endure environmental stresses, thereby minimizing the risk of premature failures [2]. Conversely, undetected or poorly managed defects during manufacturing can lead to frequent malfunctions, energy losses, and increased maintenance costs.

Defects, such as cracks, connection failures, or material inconsistencies, can significantly diminish the overall efficiency of the panels, ultimately raising operational expenses and decreasing the financial returns from solar installations [3], [4]. Therefore, early identification and proactive management of defects during the production of solar cells are critical to mitigating these risks and ensuring the long-term profitability and sustainability of PV projects.

One well-known method for looking at PV module internal quality is electroluminescence (EL) imaging, as it provides a high-resolution view of the cells' performance and reveals hidden defects, such as microcracks, delamination, and electrical failures [5], [6]. While EL imaging offers valuable insights, traditional manual inspection methods are often prone to human error, including overlooking minor defects, misinterpretations, or subjective judgments during the analysis process [7]. These inaccuracies can result in undetected defects, which, if left unaddressed, may lead to premature module degradation, reduced energy output, and, ultimately, higher maintenance costs. Furthermore, manual inspections can be time-consuming and labor-intensive, making them less efficient for large-scale production environments, where the need for accurate and rapid assessments is paramount to ensure the reliability of PV systems [8], [9].

Integrating modern image processing techniques powered by artificial intelligence (AI) can substantially improve flaw identification in PV module manufacturing [10]. By applying AI-driven models to EL images, it is possible to automate the inspection process, improving the accuracy and consistency of defect detection. This technological advancement reduces the risk of errors and enhances defect detection capabilities [11]. The association of machine learning into the energy sector, especially in PV systems, presents a promising approach to address the challenges associated with fault detection and diagnosis [12], [13]. Studies conducted in recent years have demonstrated the effectiveness of convolutional neural networks (CNNs) in identifying and diagnosing a range of defects in EL

images, showcasing their potential for accurate fault detection. The progress in this domain is thoroughly recorded, especially in publications [14], [15], [16], which highlight the capability of CNNs for accurate defect detection.

In [17], [18], and [19], researchers applied computer vision techniques, specifically using the You Only Look Once version 3 (YOLOv3), YOLOv5, and YOLOv8 object detection models, to identify faults in EL images. Building upon these advancements, Jha and Babiceanu [20] further explored defect detection methods, incorporating high-performance algorithms, such as faster R-CNN and cascade R-CNN. They also examined the effectiveness of real-time detection techniques, with a particular emphasis on the YOLO series, highlighting their relevance and efficiency in addressing defect detection challenges. The studies presented in [21], [22], [23], and [24] underscore the increasing integration of transformer-based models in computer vision, showcasing their capacity to enhance performance on jobs, including semantic segmentation, object detection, and image classification. This trend represents a significant advancement in the application of transformer architectures to various computer vision challenges. In [25], an enhanced defect detection technique tailored for PV modules was proposed, leveraging the faster R-CNN architecture. This approach incorporates innovative features, such as a lightweight attention module that combines channel and spatial convolutions. Chen et al. [26] combined the convolutional block attention module (CBAM) with MobileNetv3 to further improve model performance.

Innovative deep learning and machine learning techniques have been employed to address critical challenges in defect detection for PV systems. The integration of multiscale linear attention with scale distribution alignment learning has resulted in a framework that delivers highly accurate anomaly detection, as demonstrated in [27]. This approach exhibits exceptional performance on EL datasets and offers significant potential for real-world industrial applications. In [28], the feature map conversion and hypersphere transformation network is introduced. This weakly supervised method leverages prior knowledge and unlabeled data to mitigate the high cost of gathering labeled datasets. Its primary aim is to effectively distinguish between normal and defective PV samples. Meanwhile, Hong and Pula [29] propose a two-step process combining digital twin analysis with a PSO-optimized swarm transformer. This method achieves an impressive fault classification accuracy of 98.55% and demonstrates robust practicality for real-world applications.

A common challenge in many classification images' tasks is the limitation of datasets, particularly when data are scarce or imbalanced, which can hinder model performance. A promising solution to this issue is the implementation of generative adversarial networks (GANs), which produce novel, synthetic images by deriving features from an input image collection, thus augmenting the data and enhancing model training [30]. Various GAN-based models have been proposed, such as the deep convolutional generative adversarial network (DCGAN) 31], Wasserstein generative adversarial network (WGAN) [32], and WGAN with gradient penalty [33]. These models have demonstrated significant advancements in the generation of synthetic data across various domains. An effective solution to the problem of insufficient sample availability in EL images

was presented in [34] through the application of a GAN-based approach. To address image resolution issues, Karnewar and Wang [35] proposed a simple yet effective strategy called a gradient multiscale adversarial network. Undirwade and Das [36] suggested using GANs to transform the input original image into a new synthetic image by applying optimized noise to the original image's latent space representation. A proposed method, f-AnoGAN by Schlegl et al. [37], uses WGAN for anomaly detection following the swift mapping of an image to a designated point in latent space. Although advanced, approaches, such as f-AnoGAN, StyleGAN, BigGAN, and CycleGAN, share significant limitations. In scenarios, such as EL flaw detection, where data are often scarce, their reliance on large datasets and substantial computational resources renders them unsuitable. Furthermore, their high processing times pose challenges for real-time system implementation, particularly when addressing complex flaws.

This research addresses a critical challenge in PV module manufacturing: the classification of defects using EL imaging. EL imaging is a powerful diagnostic tool, but its effectiveness is limited by imbalanced datasets and the inability of traditional methods to accurately capture rare or complex defect features. To address these issues, this study evaluates and compares two generative approaches: a standard DCGAN and a DCGAN enhanced with visual geometry group 16 (VGG16), emphasizing their respective capacities to generate synthetic defect data and improve classification performance. The goal is to generate synthetic defect data to tackle data imbalance and enrich underrepresented defect categories. By adding a perceptual loss function to the generator, a detailed evaluation reveals that the synthetic images generated by DCGAN with VGG16 are of significantly higher quality, particularly for complex defect classes, compared with those produced by the standard DCGAN. These superior synthetic images provide enhanced feature representation, leading to better classification accuracy across all defect types.

The key contributions of this research are as follows.

1) *Enhanced synthetic data generation*: The integration of VGG16 within the DCGAN framework introduces perceptual loss, enabling the generation of synthetic defect images that exhibit finer details and higher visual fidelity, especially for complex and subtle defects.

2) *Impactful comparison of generative models*: The side-by-side comparison of DCGAN and DCGAN with VGG16 highlights the superiority of the latter in producing targeted synthetic images that effectively address dataset imbalances.

3) *Class-specific improvements*: The study demonstrates that DCGAN with VGG16 significantly improves classification performance for all defect categories, with notable advancements in detecting underrepresented and challenging defects. This is attributed to the ability of the enhanced framework to generate synthetic data that closely resemble real-world defect characteristics.

4) *Evaluation across real and synthetic datasets*: By combining real and synthetic EL images, the approach ensures robust evaluation and highlights the practical applicability of using advanced generative models for PV defect detection.
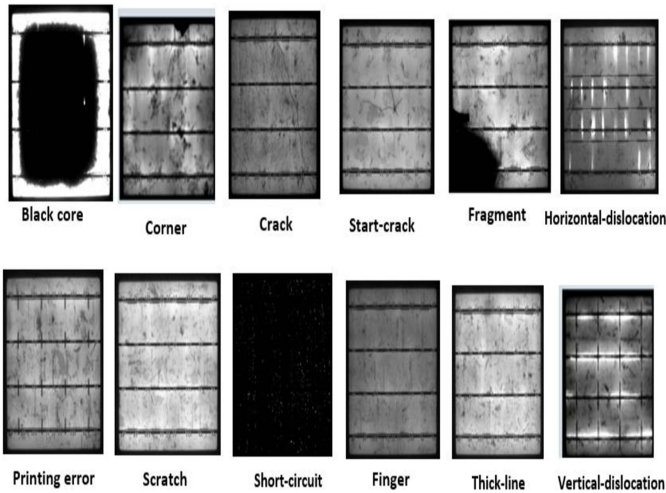
Fig. 1. Dataset of PV defects categories.

TABLE I
DEFECTS TYPE, THEIR LABELS, AND NUMBER OF IMAGES FOR THE
12 FAULTS' CATEGORIES

| Faults | Labels | Number images |
|---|---|---|
| Black_core | 0 | 1028 |
| Corner | 1 | 9 |
| Crack | 2 | 1260 |
| Star_crack | 3 | 135 |
| Fragment | 4 | 7 |
| Horizontal_dislocation | 5 | 798 |
| Printing_error | 6 | 32 |
| Scratch | 7 | 5 |
| Short circuit | 8 | 492 |
| Finger | 9 | 2958 |
| Thick_line | 10 | 981 |
| Vertical_dislocation | 11 | 137 |
| Total | | 7842 |

## II. DATASET DESCRIPTION

This study utilized a curated collection of pictures with excellent resolution depicting faults in PV cells, obtained from a publicly accessible global dataset of EL images. This dataset was jointly compiled by the Hebei University of Technology and Beijing University of Aeronautics and Astronautics [38]. It constitutes a thorough compilation distinguished by superior resolution, variety, and comprehensiveness in the representation of anomalies. The online repository has 12 unique defect kinds. Fig. 1 shows the 12 categories of defects included in the dataset. As an example, the dataset's imbalance is illustrated by the disparity in images across categories, where 'Finger' comprises 37.7% of all images, and underrepresented classes, such as "Scratch" and "Corner," collectively account for less than 0.2%.

The dataset reveals significant class imbalance across the 12 defect categories, as seen in Table I. Some classes contain a large number of images, while others have very few. For example, "finger" has the highest count with 2958 images, and "crack"

follows closely with 1260 images. In contrast, classes, such as "corner" and "scratch," are severely underrepresented, with only nine and five images, respectively. Other categories, such as "printing error" and "fragment," also suffer from a limited number of samples, making up only a small fraction of the dataset. This uneven distribution poses a challenge for defect classification, as the model may become biased toward the more frequent classes, which could lead to poor performance in detecting defects from the rarer categories. The unequal representation of defect types is likely to affect the model's ability to generalize well, particularly for underrepresented classes, thus highlighting the need for techniques that address this imbalance during model training.

## III. METHODOLOGY

This study presents a new DCGAN model that incorporates perceptual loss and VGG16 to address the issues of data scarcity and class imbalance in PV module fault classification. The research investigates and compares two generative approaches, namely DCGAN and DCGAN integrated with VGG16, to evaluate their ability to generate synthetic defect images and improve defect classification accuracy (see Fig. 2).

This approach addresses data scarcity and class imbalance in EL image datasets by generating synthetic images through key steps: capturing complex EL defect features, preserving defect textures, and reproducing structural details. However, challenges include maintaining fine defect details, accurately replicating textures, and ensuring structural integrity for reliable classification or detection.

To address these challenges, the VGG16 network is incorporated as a feature extraction module. Leveraging its deep convolution and pooling layers, VGG16 captures rich semantic representations of EL images, guiding the generation process toward more realistic outputs that remain faithful to the original data. In addition, integrating this feature extraction capability with perspective loss in the generator's update function (G Loss) significantly enhances the quality of the synthetic images. Perspective loss enables a direct comparison of features between generated and real images, ensuring that fine details and complex structures are accurately captured. This combination not only improves the visual fidelity of synthetic images but also boosts their utility in training machine learning models for more effective defect detection and classification.

However, unlike advanced approaches, such as StyleGAN, BigGAN, or CycleGAN—renowned for their ability to generate high-quality images—these models have significant limitations. They require extensive datasets and substantial computational resources, making them impractical for applications, such as EL defect detection, where data availability is limited, and computational efficiency is critical.

To address these challenges, a structured approach was undertaken. First, synthetic images were generated using the standard DCGAN model, which effectively captured certain defect classes but struggled with more complex or subtle defects, lacking the necessary detail and feature representation. To overcome this limitation, a perceptual loss function,
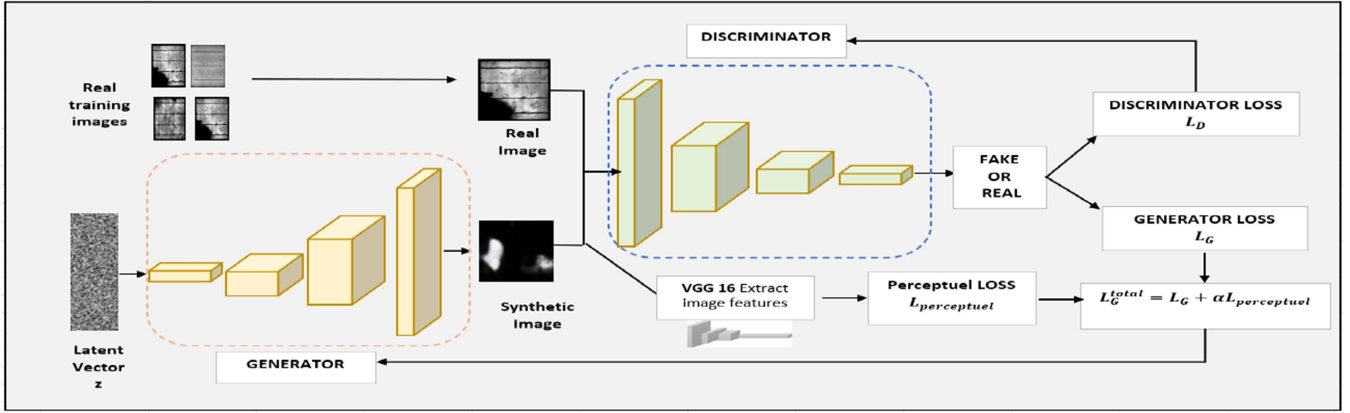
Fig. 2.    DCGAN-VGG16 architecture with perceptual loss for synthetic image generation.

augmented with VGG16, was integrated into the DCGAN framework. This enhancement, referred to as DCGAN with VGG16, significantly improved the generator's ability to produce more realistic and detailed defect images. A comparative analysis was then conducted between the standard DCGAN and the VGG16-enhanced version to assess improvements in image quality and defect representation. Subsequently, two classification models were trained: one using real defect images combined with synthetic data from the standard DCGAN and the other leveraging real images alongside synthetic samples from the enhanced DCGAN-VGG16 framework. The evaluation demonstrated that the DCGAN-VGG16 model outperformed the standard approach in defect classification tasks, providing a more robust and effective solution for identifying and categorizing defects in PV modules. It successfully mitigated dataset imbalances and enhanced the detection of intricate defect patterns, making it a viable method for improving defect analysis in PV applications.

### A. DCGAN-Based Synthetic Image Generation for Addressing Class Imbalance

The proposed architecture begins by using a DCGAN to generate synthetic defect data, addressing the issue of class imbalance in defect detection. Many datasets, especially in PV module inspection, suffer from underrepresented defect classes. DCGANs are used to create high-quality synthetic images that augment the dataset with additional samples from these rare classes, improving model performance and ensuring more balanced and accurate defect detection [39]. The DCGAN architecture comprises two primary elements: the generator and the discriminator. The generator intends to produce synthetic images that closely resemble real defect data by conditioning on class labels. It accepts a stochastic noise vector $z$ and a categorical label $y$ as input and generates an image $G(z, y)$ that appears indistinguishable from real defect images.

The discriminator, on the other hand, is a binary classifier that seeks to differentiate between authentic photos $x$ from the dataset and counterfeit images produced by the generator. It

outputs a probability that indicates whether an image is real or not conditioned on the same label $y$.

The training strategy of the DCGAN entails an adversarial conflict between the generator and the discriminator. The generator aims to minimize the detection of synthetic images by the discriminator, while the discriminator strives to distinguish real from synthetic images. The loss functions for both networks are defined as follows:

$$L_D = -E_{p(x)}[\log\left(D\left(x, y\right)\right] - E_{p(z)}\left[\log(1 - D\left(G\left(z, y\right), y\right))\right] \tag{1}$$

$$L_G = E_{p(z)}\left[\log D\left(G\left(z, y\right), y\right)\right] \tag{2}$$

where $D(x, y)$ represents the discriminator's output for a real image with label $y$, and $G(z, y)$ is the image generated by the generator with label $y$. $E_{p(x)}$ is the expected value over the distribution of real data, and $E_{p(z)}$ represents the expected value over the distribution of the noise vector $z$.

This adversarial training helps the generator improve its image generation capabilities, resulting in synthetic defect images that are realistic and contextually relevant to each defect class. The synthetic images created using this approach augment the dataset, allowing the model to be trained on a more balanced and comprehensive set of images. Thus, the DCGANs ability to produce synthetic images conditioned on defect class labels not only provides a means to augment the dataset but also ensures that the model is exposed to a balanced representation of all defect classes, including those that may be rare or difficult to capture in real-world data.

### B. DCGAN With VGG16

Given the moderate quality of images produced by the standard DCGAN, particularly for complex defect classes, we enhanced the framework by introducing a perceptual loss function derived from the pretrained VGG16 network. This perceptual loss captures high-level visual features of real images, ensuring that the synthetic images generated by the model replicate these features more accurately.

Technically, the initial convolutional layers of VGG16, which extract features, such as edges and textures, were employed to

compute this loss. The features of synthetic images $f_{\text{fake}}$ and real images $f_{\text{real}}$ were extracted using VGG16, and the perceptual loss was defined as the mean squared error (MSE) between the two feature sets

$$L_{\text{perceptual}} = \frac{1}{N} \sum_{i=1}^{N} (f_{\text{fake}}^i - f_{\text{real}}^i)^2 \tag{3}$$

where $N$ denotes the total number of extracted features. This perceptual loss was then weighted and added to the generator's standard adversarial loss to form the total generator loss

$$L_G^{\text{total}} = L_G + \alpha L_{\text{perceptual}} \tag{4}$$

Here, $L_G$ represents the adversarial loss, and $\alpha$ is a hyperparameter that controls the relative importance of the perceptual loss. This enhanced DCGAN framework, incorporating the perceptual loss function, demonstrated a significant improvement in the quality of synthetic images compared with the standard DCGAN, particularly for classes with intricate features. These improvements contribute to a more comprehensive data representation and subsequently lead to better performance in defect classification tasks.

After the standard DCGAN and DCGAN with VGG16 generates synthetic images, they are integrated into the training set for a CNN classifier. CNNs are utilized to categorize images according to defect classifications. The CNN design often consists of several convolutional layers, pooling layers, and fully connected layers, enabling the model to acquire hierarchical characteristics of pictures. In addition, the CNN classifier is trained on both real and synthetic data, allowing it to proficiently identify flaws across all classes, including those that are underrepresented in the real dataset.

## IV. EVALUATION METRICS

The evaluation of synthetic image generation models, specifically DCGAN and DCGAN with VGG16, is conducted on the test dataset using key performance metrics: accuracy, precision, recall, $F1$-score, MSE, peak signal-to-noise ratio (PSNR), and Fréchet inception distance (FID). MSE calculates the average squared difference between synthetic and real images, with lower values indicating superior performance. PSNR assesses the quality of generated images, where higher values correspond to clearer and less noisy images. FID evaluates the similarity between the distributions of real and synthetic images, with lower values indicating greater similarity between them. These metrics are formally defined in (5) through (11)

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} (I_{\text{real}}(i) - I_{\text{gen}}(i)) \tag{5}$$

$$\text{PSNR} = 10 \log \frac{(\text{MAX}i)^2}{\text{MSE}} \tag{6}$$

$$\text{FID} = [\mu_r - \mu_g]^2 + T_r \left( \sum_r + \sum_g - 2 \left( \sum_r \sum_{rg} \right)^{\frac{1}{2}} \right). \tag{7}$$
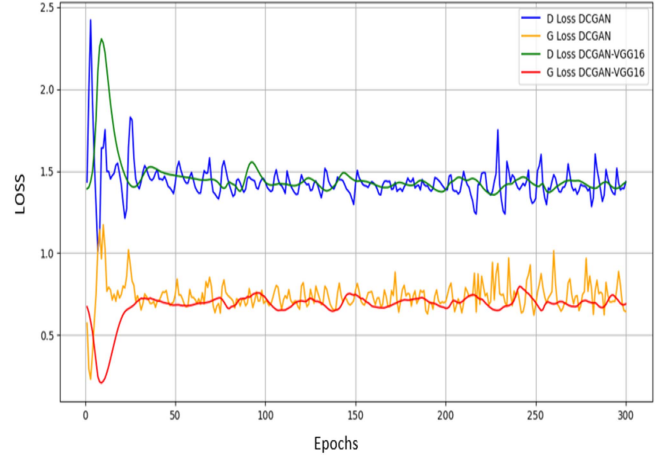


Fig. 3. G loss and D loss across DCGAN and DCGAN-VGG16 models during training.

Furthermore, accuracy reflects the ratio of correct predictions to the total predictions. Precision measures the proportion of true positives among the predicted positives, while recall assesses the model's capacity to recognize true positives among all actual positive cases. The $F1$-score, integrating precision and recall, provides a comprehensive assessment of the model's efficacy

$$\text{Precision} = \frac{T_p}{T_p + F_p} \tag{8}$$

$$\text{Accuracy} = \frac{T_p + T_N}{T_p + F_p + T_N + F_N} \tag{9}$$

$$\text{Recall} = \frac{T_p}{T_p + F_N} \tag{10}$$

$$F1 - \text{Score} = 2 * \frac{(\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})} \tag{11}$$

where $I_{\text{real}}(i)$ is the pixel value of $i$ in the real image, $I_{\text{gen}}(i)$ is the pixel value of $i$ in the generated image, $N$ is the total number of pixels in the image, and $\text{MAX}i$ is the maximum possible pixel value in the image. $T_p$, $T_N$, $F_p$, and $F_N$ represent the number of true positives, true negatives, false positives, and false negatives, respectively. $\mu_r$ and $\mu_g$ are the mean feature vectors of the real and generated images, respectively, and $\sum_r$ and $\sum_g$ are the covariance matrices of the real and generated images, respectively. $T_r$ is the trace of the matrix.

## V. RESULTS AND DISCUSSIONS

The analysis of the loss curves—discriminator loss (D Loss) and generator loss (G Loss)—for the DCGAN and DCGAN-VGG16 models (see Fig. 3) highlights contrasting training dynamics that reflect the mechanisms of each architecture. For the standard DCGAN, the D Loss (blue curve) shows pronounced oscillations throughout training, indicating instability in the interaction between the generator and discriminator. This instability arises from imprecise model fits, preventing efficient

TABLE II
HYPERPARAMETERS AND TRAINING CONFIGURATIONS OF DCGAN AND DCGAN-VGG16

| PARAMETER | DCGAN | DCGAN-VGG16 | DESCRIPTION |
|---|---|---|---|
| Learning Rate | 0.0002 | 0.0002 | Specifies the step size for the optimizer to update weights during training. |
| Batch Size | 32 | 32 | The number of samples processed before updating the model parameters. |
| Number of Epochs | 200 | 200 | The total number of iterations over the entire dataset during training. |
| Optimizer | Adam | Adam | The optimization algorithm used to minimize the loss function. |
| Generator Layers | 4 | 4 | Number of layers in the generator network. |
| Discriminator Layers | 4 | 4 | Number of layers in the discriminator network. |
| Activation Function | ReLU | ReLU | Specifies the activation function applied after each layer. |
| Perceptual Loss Weight ($\lambda$) | N/A | 0.1 | Weight assigned to the perceptual loss for DCGAN-VGG16 to balance adversarial and feature loss. |
| Noise Vector Dimension | 100 | 100 | The size of the input noise vector for the generator. |
| Input Image Resolution | 640 × 640 | 640 × 640 | Resolution of the images used for training and generation. |
| Pre-trained Model | N/A | VGG16 | Indicates whether a pre-trained model (e.g., VGG16) is integrated into the framework for feature extraction. |



Fig. 4. Comparison of real and synthetic defect images generated by DCGAN.



Fig. 5. Comparison of real and synthetic defect images generated by DCGAN with VGG16.

convergence. Similarly, the G Loss (orange curve) fluctuates significantly, revealing the generator's struggle to consistently improve the quality of the generated images. The hyperparameters and training configurations of DCGAN and DCGAN-VGG16 are listed in Table II.

In contrast, the DCGAN-VGG16 model demonstrates a more stable and structured learning process. Although the D Loss (green curve) converges more slowly during the initial phase, this period is essential for the model to absorb complex semantic features, guided by VGG16s feature extraction. After this phase, the D Loss stabilizes with minimal oscillations, indicating improved equilibrium between the generator and discriminator. Meanwhile, the G Loss (red curve) steadily declines, reflecting the generator's ability to produce higher quality images with greater detail and realism.

This comparison underscores the advantages of incorporating VGG16, which introduces perceptual loss to guide the generator toward meaningful and consistent improvements. The result is a model that outperforms the standard DCGAN in generating realistic and detailed images.

A detailed comparison of real and synthetic data generated by DCGAN and DCGAN with VGG16 is presented in Figs. 4 and 5. Initially, we observe that the quality of synthetic images generated by the standard DCGAN varies significantly across different defect classes. For classes 1, 11, 8, 4, 5, and 0, the generated images are relatively clear and representative of the corresponding defects, offering a reasonable approximation of the defect patterns. However, despite these satisfactory results, the images exhibit a certain degree of blur and lack fine detail, which affects the overall sharpness and clarity of the generated defects. This limitation is particularly noticeable in the subtle
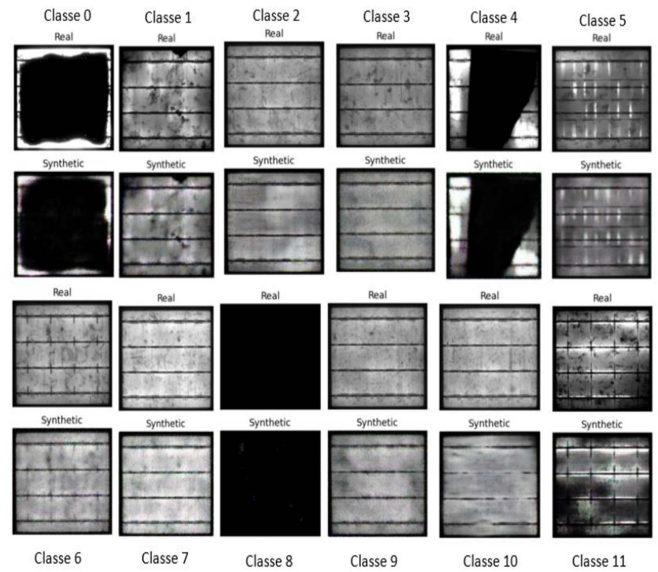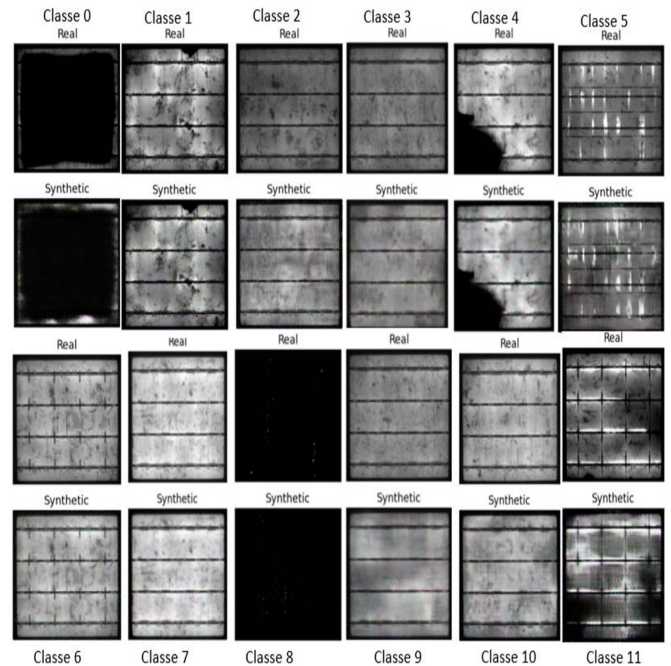
textures and minute features of the defects, which are difficult to discern due to the inherent noise in the generated images.

In contrast, the synthetic images generated by the standard DCGAN for defect classes 2, 3, 10, and 7 exhibit significantly lower quality. These images are characterized by substantial blurring, which obscures meaningful details crucial for the accurate defect identification. The absence of well-defined structural information, such as crack patterns or localized anomalies, highlights the model's limitations in capturing the intricate features necessary for precise defect detection. This deficiency
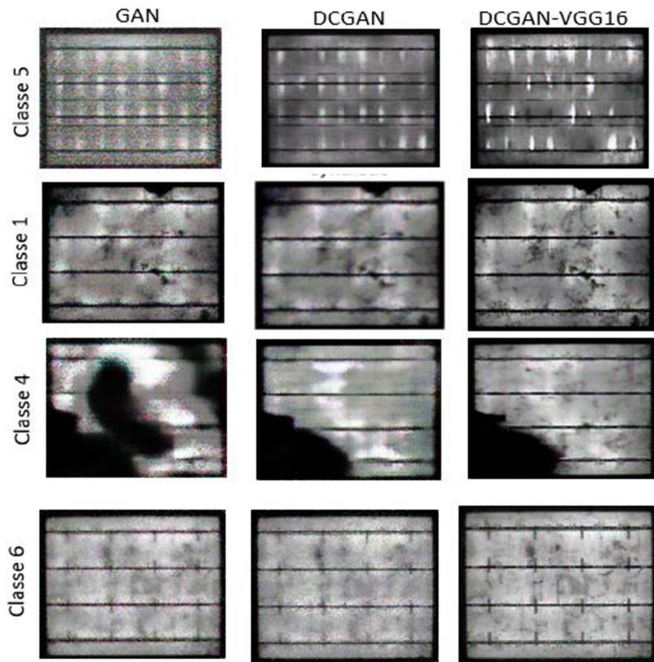
Fig. 6. Comparison of synthetic defect images generated by GAN, DCGAN, and DCGAN-VGG16.

TABLE III
COMPARISON OF IMAGE QUALITY METRICS (MSE, PSNR, AND FID) FOR SYNTHETIC IMAGES GENERATED BY DCGAN AND DCGAN WITH VGG16

| | SYNTHETIC IMAGES DCGAN | | | | SYNTHETIC IMAGES DCGAN WITH VGG 16 | | |
|---|---|---|---|---|---|---|---|
| | *MSE* | *PSNR* | *FID* | | *MSE* | *PSNR* | *FID* |
| 0 | 0.325 | 53.00 | 7.66 | 0 | 0.313 | 52.92 | 9.65 |
| 1 | 0.047 | 61.32 | 209.68 | 1 | 0.044 | 61.18 | 207.32 |
| 2 | 0.342 | 52.78 | 0.83 | 2 | 0.355 | 53.24 | 2.25 |
| 3 | 0.109 | 57.75 | 1.49 | 3 | 0.099 | 58.92 | 0.47 |
| 4 | 0.294 | 53.43 | 11.96 | 4 | 0.262 | 54.14 | 9.39 |
| 5 | 0.249 | 54.16 | 2.93 | 5 | 0.265 | 54.04 | 1.73 |
| 6 | 0.154 | 56.23 | 208.13 | 6 | 0.007 | 59.11 | 208.45 |
| 7 | 0.057 | 60.54 | 210.13 | 7 | 0.067 | 60.29 | 209.36 |
| 8 | 0.339 | 52.85 | 1.63 | 8 | 0.331 | 52.33 | 2.00 |
| 9 | 0.373 | 52.40 | 1.60 | 9 | 0.360 | 52.53 | 0.82 |
| 10 | 0.011 | 67.59 | 5.07 | 10 | 0.011 | 67.29 | 0.49 |
| 11 | 0.289 | 53.17 | 7.27 | 11 | 0.297 | 53.23 | 7.62 |

improves the overall structure of images, making shapes and contours consistent with real data. However, it may still lack precision in fine detail, essential for reproducing the complex characteristics of electroluminescent defects. The addition of VGG16 to DCGAN represents a significant step forward by introducing perceptual loss, which enables the features of synthetic and real images to be compared at a more abstract level of representation. This combination results in more realistic synthetic images, capturing both the overall structure and subtle details of defects, enhancing their usefulness for training detection and classification models.

Table III provides a comparative analysis of image quality metrics (MSE, PSNR, and FID) for synthetic images generated by the DCGAN and DCGAN-VGG16 models, highlighting their performance differences. MSE, a key indicator of similarity between real and synthetic images, shows generally lower values for the DCGAN-VGG16 model. This suggests that integrating VGG16 enhances feature extraction, enabling the generation of synthetic images with greater fidelity to real-world counterparts. Notably, class 6 exhibits a significant improvement, with the DCGAN-VGG16 model achieving a much lower MSE, reflecting its superior ability to replicate fine image details.

For other classes, such as 0 and 1, the differences in MSE are relatively small, indicating that the impact of VGG16 is less pronounced for simpler or less complex image patterns. Interestingly, class 10 shows no measurable improvement, as the MSE values remain consistent across both models. These results imply that while VGG16 contributes to enhanced feature representation, its effectiveness may vary depending on the complexity and characteristics of the defect class.

Overall, the integration of VGG16 offers a consistent, albeit modest, advantage in image quality across most classes. The model's improvements are particularly evident for complex defect categories, underscoring its potential to significantly enhance the performance of generative frameworks, such as DCGAN.

The FID is a prevalent metric for assessing the quality of images produced by GANs. It analyses the distribution of features derived from both actual and generated images, offering insight into the similarity of the two distributions. Reduced FID values signify enhanced similarity between created and

renders the synthetic images unsuitable for high-quality classification or advanced analytical tasks, underscoring the need for more sophisticated generative techniques to address these challenges effectively. To address these challenges and improve the quality of the generated images, a VGG16-based perceptual loss function was integrated into the DCGAN framework. This enhancement significantly elevated the realism and clarity of the synthetic data, particularly by preserving intricate structural details. The synthetic images produced by the DCGAN-VGG16 model demonstrated substantial improvements, offering sharp, high-resolution depictions of defects (see Fig. 5). Importantly, defect features, such as surface anomalies, are distinctly represented, closely mirroring the characteristics of real defect images.

However, despite these improvements, certain classes, particularly 3, 2, and 10, still pose challenges. For these classes, the defects are extremely fine and subtle, making it difficult for the model to generate realistic representations with sufficient clarity. The details of these defects remain faint, and while the images are significantly sharper compared with the DCGAN-generated ones, the very fine structures still elude perfect recovery. This suggests that, while the VGG16-enhanced DCGAN has improved the quality of image generation, further refinement is needed for capturing highly detailed defects that require high spatial resolution.

The quality of the images generated by the different GANs models reveals significant distinctions in terms of realism and consistency, as shown in Fig. 6. Basic GAN, while effective for generating synthetic images, often produces blurred and inconsistent results, particularly in the critical details of PV module defects. By incorporating a convolutional architecture, DCGAN

authentic images, indicating superior quality of synthetic data. In the comparison of DCGAN and DCGAN with VGG16, the results show notable differences in FID scores across various classes. For DCGAN, the FID values are generally higher, with some extreme values, such as 209.68 for class 1 and 208.13 for class 6, indicating that the generated images deviate significantly from the real images in terms of feature distribution. On the other hand, the DCGAN with VGG16 generally achieves lower FID values, especially for classes, such as 3 (0.47), 5 (1.73), and 10 (0.49), suggesting that the addition of VGG16 helps the model to generate images that are closer in feature distribution to the real images.

The integration of VGG16 into the DCGAN framework significantly improves the FID for many classes, indicating enhanced realism and diversity in the generated images. However, certain classes, such as 1 (207.32), 6 (208.45), and 7 (209.36), still show relatively high FID values. These elevated scores highlight persistent discrepancies in feature distributions between real and synthetic images for these specific categories.

While the VGG16 enhancement improves the overall quality of synthetic data, it does not fully align the feature distribution with real data in all cases. The lower FID scores for more complex defect classes reflect the effectiveness of the approach, but the higher values in certain categories emphasize the need for further refinements in training strategies or architectural design to address these gaps and enhance generative performance.

The PSNR is a critical metric for evaluating the quality of generated images, with higher values indicating better image fidelity. In this analysis, the PSNR values for both DCGAN and DCGAN-VGG16 are generally high, with most exceeding 50, indicating good quality synthetic images across generated classes. While the overall PSNR values are relatively close between the two models, the inclusion of VGG16 often results in superior image quality.

For instance, in class 3, DCGAN-VGG16 achieves a PSNR of 58.92, compared with DCGANs 57.75. Similarly, for class 6, DCGAN-VGG16 records a PSNR of 59.11, outperforming DC-GANs 56.23. These improvements highlight the impact of the perceptual loss function introduced by VGG16, which enhances the clarity and fidelity of the generated images in specific cases, further refining the model's output quality.

The histograms comparing the pixel power distributions of DCGAN and DCGAN-VGG16 (see Figs. 7 and 8) reveal notable differences in image diversity and pixel power concentration. Images generated by DCGAN exhibit narrow distributions with sharp peaks, reflecting limited diversity in textures and visual details. This lack of variability leads to artificial and unrealistic images, highlighting the model's inability to capture the complexity of real-world images.

In contrast, DCGAN-VGG16 demonstrates broader distributions with less pronounced peaks, indicating improved representation of pixel power variability. This suggests a better ability to simulate the diversity and complexity of real images. The enhancement is attributed to the integration of VGG16, which leverages perceptual loss to guide the generator in capturing intricate textures and visual details, resulting in more realistic and diverse synthetic images.
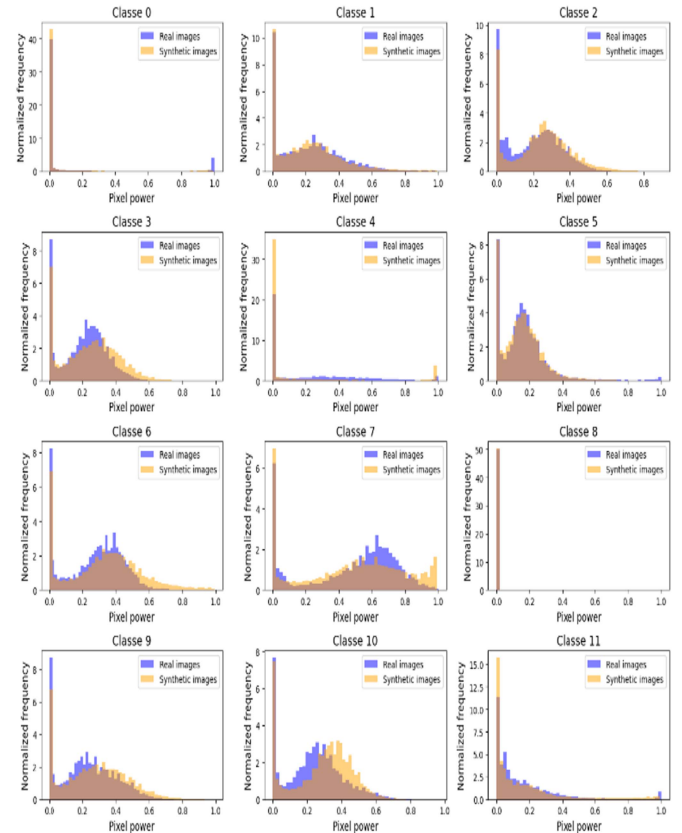


Fig. 7. Histogram analysis of pixel intensity distributions for DCGAN-generated images.
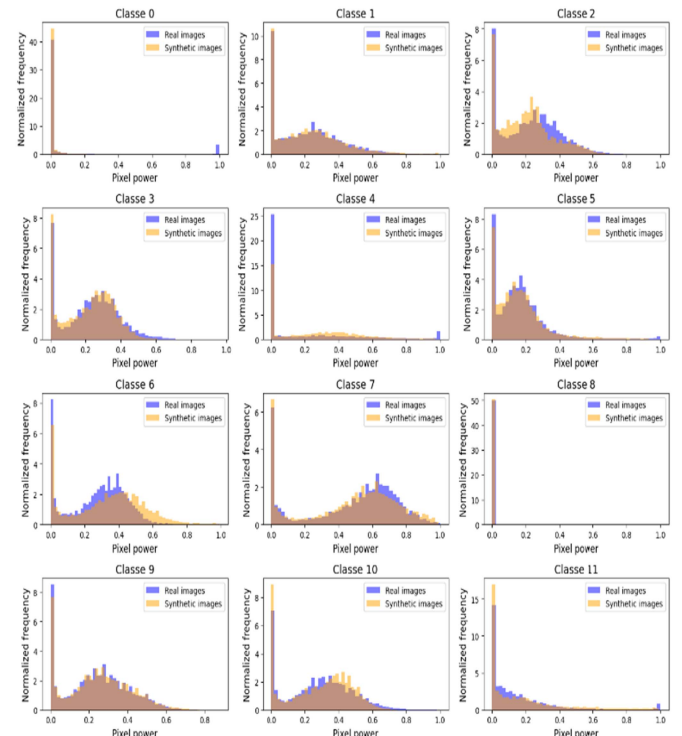


Fig. 8. Histogram analysis of pixel intensity distributions for DCGAN-VGG16-generated images.
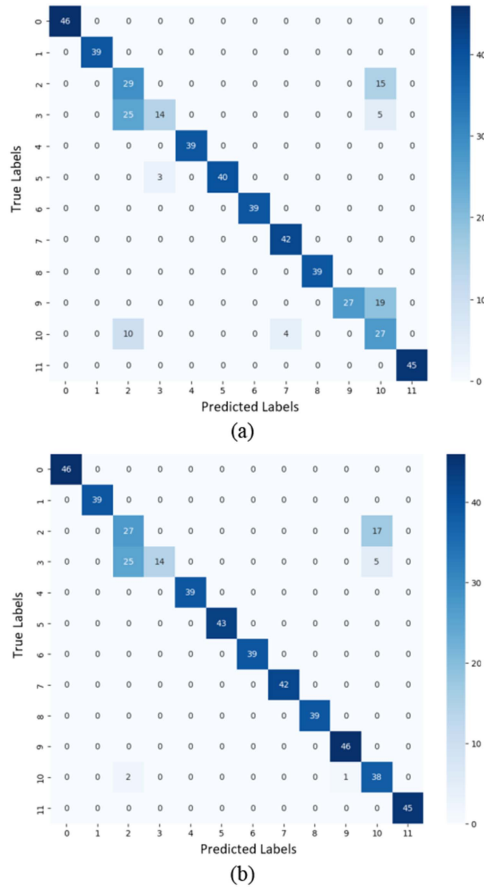
Fig. 9. (a) Confusion matrix of real and DCGAN synthetic data. (b) Confusion matrix of real and DCGAN with VGG16 synthetic data.

TABLE IV
ERROR METRICS FOR REAL AND SYNTHETIC IMAGES

| REAL AND SYNTHETIC DCGAN | | | | REAL AND SYNTHETIC DCGAN WITH VGG 16 | | |
|---|---|---|---|---|---|---|
| | PRES | REC | F1-SCORE | | PRES | REC | F1-SCORE |
| 0 | 1.00 | 1.00 | 1.00 | 0 | 1.00 | 1.00 | 1.00 |
| 1 | 1.00 | 1.00 | 1.00 | 1 | 1.00 | 1.00 | 1.00 |
| 2 | 0.45 | 0.66 | 0.54 | 2 | 0.47 | 0.60 | 0.52 |
| 3 | 0.82 | 0.32 | 0.46 | 3 | 0.67 | 0.34 | 0.51 |
| 4 | 1.00 | 1.00 | 1.00 | 4 | 1.00 | 1.00 | 1.00 |
| 5 | 1.00 | 0.93 | 0.96 | 5 | 1.00 | 1.00 | 1.00 |
| 6 | 1.00 | 1.00 | 1.00 | 6 | 1.00 | 1.00 | 1.00 |
| 7 | 0.91 | 1.00 | 0.95 | 7 | 1.00 | 1.00 | 1.00 |
| 8 | 1.00 | 1.00 | 1.00 | 8 | 1.00 | 1.00 | 1.00 |
| 9 | 1.00 | 0.59 | 0.74 | 9 | 1.00 | 1.00 | 1.00 |
| 10 | 0.41 | 0.66 | 0.50 | 10 | 0.62 | 0.82 | 0.71 |
| 11 | 1.00 | 1.00 | 1.00 | 11 | 1.00 | 1.00 | 1.00 |
| Accuracy | | 0.84 | | Accuracy | | 0.90 | |

The two confusion matrices provide insights into the performance of defect classification models trained using combined datasets of real and synthetic images with the synthetic data generated by DCGAN [see Fig. 9(a)] and DCGAN with VGG16 [see Fig. 9(b)].

The confusion matrix in Fig. 9(a) shows the classification results when combining real defect images with synthetic data generated by DCGAN. While the synthetic data improve the representation of rare classes, challenges remain in differentiating categories with high visual similarity, particularly in underrepresented classes. For instance, rare defects, such as "Scratch" (class 7) and "Corner" (class 1), are well classified, but complex patterns, such as "Star Crack" (class 3), are often confused with other defects due to overlapping features and insufficient differentiation. Misclassifications are also evident in "Crack" (class 2), "Horizontal Dislocation" (class 5), and "Thick Line" (class 10), reflecting limitations in feature diversity within the synthetic dataset.

Fig. 9(b) illustrates the improved classification outcomes when using DCGAN enhanced with VGG16. Most defect classes, such as "Black Core" (class 0), "Scratch" (class 7), and "Corner" (class 1), show higher accuracy, highlighting the effectiveness of VGG16 in generating realistic and representative synthetic samples. High-frequency classes, such as "Finger" (class 9) and "Horizontal Dislocation" (class 5), also exhibit improved performance, indicating enhanced feature representation. However, misclassifications persist in "Crack" (class 2), "Star Crack" (class 3), and "Horizontal Dislocation" (class 5), suggesting that subtle interclass variations remain difficult to resolve.

Overall, the confusion matrices demonstrate the value of synthetic data in addressing dataset imbalances. While the standard DCGAN provides moderate improvements, the DCGAN-VGG16 framework generates higher quality synthetic images that significantly enhance classification accuracy. However, persistent challenges in certain defect classes highlight the need for further advancements in generative processes and classifier architectures to fully capture intricate defect features and interclass distinctions.

Table IV presents a detailed comparison of performance metrics, highlighting key differences between the DCGAN and DCGAN-VGG16 models. Both models achieve perfect precision, recall, and F1-scores (1.00) for most classes, including 0, 1, 4, 6, 8, and 11, demonstrating their strong ability to detect defects in these categories. However, challenges persist in more complex classes, such as 2, 3, and 10, where performance declines for both models. The standard DCGAN struggles significantly with these defect types, as indicated by moderate precision and recall values. While the addition of VGG16 leads to slight improvements in feature representation and classification accuracy, these defect classes remain difficult due to their intricate patterns or overlapping features. This analysis underscores the need for further optimization to address nuanced defect detection, even with advanced models, such as DCGAN-VGG16.

The integration of VGG16 into the DCGAN framework yields notable improvements, particularly in classes 5 and 7, where recall and F1-scores are significantly enhanced. The overall accuracy of the DCGAN-VGG16 framework increases from 0.84 to 0.90, showcasing the effectiveness of the perceptual loss function in generating more realistic and detailed synthetic images. However, fine defect details in classes 2, 3, and 10 remain challenging for both models, highlighting the need for further refinements in the generative architecture or training process to better capture subtle features.

The comparison between the standard DCGAN and DCGAN-VGG16 shows clear advancements in metrics, such as MSE, PSNR, and FID. The integration of VGG16 consistently reduces MSE, reflecting improved image quality across most defect classes. Notable enhancements are observed in classes 1 and 6, with class 6 exhibiting a significant reduction in MSE and a corresponding increase in PSNR. In contrast, simpler or less distinct patterns, such as in classes 0 and 11, show minimal differences, where the impact of VGG16 is less pronounced. These results underline the value of VGG16 in improving image generation, particularly for more complex defect classes.

FID scores, which assess the similarity between real and synthetic image distributions, further highlight the advantages of incorporating VGG16. Lower FID scores for complex classes, such as 3 and 10, indicate superior generative performance, with enhanced realism and diversity in the synthetic data. Overall, while the DCGAN-VGG16 framework outperforms the standard DCGAN in terms of image quality and diversity, the improvements are more pronounced for complex defect classes, underscoring the potential of this approach for addressing challenging generative tasks.

## VI. CONCLUSION AND PERSPECTIVES

This study proposed an enhanced DCGAN framework integrated with VGG16 and a perceptual loss function to address critical challenges in PV defect detection using EL imaging. The enhanced model demonstrated significant improvements in synthetic image quality, achieving higher PSNR values (e.g., 59.11 for class 6) and lower FID scores (e.g., 0.47 for class 3), compared with the standard DCGAN. These advancements underline the model's ability to generate realistic and diverse synthetic images, particularly for complex and underrepresented defect classes.

The classification accuracy of the DCGAN-VGG16 framework improved from 0.84 to 0.90, showcasing its effectiveness in tackling dataset imbalance and improving defect detection. Rare defect classes, such as "Scratch" (class 7) and "Corner" (class 1), were better represented, while challenging categories, such as "Star Crack" (class 3) and "Thick Line" (class 10), benefited from improved feature extraction. However, fine defect details in these challenging classes remain difficult to capture, indicating the need for further refinements in the generative architecture or training process.

While the integration of VGG16 consistently enhanced image realism and fidelity, its impact was more pronounced for complex defects and less significant for simpler patterns. This highlights the dependency of performance gains on the complexity of the defect class, suggesting areas for improvement in future iterations.

Future work could explore advanced generative models, such as transformer-based or hybrid GAN-CNN architectures, to further address limitations in defect detection. In addition, deploying this framework in real-world industrial applications, such as automated PV module inspection lines and extending its evaluation to more complex datasets, could enhance its scalability and practical impact. This study establishes a benchmark for synthetic data generation and highlights the potential of hybrid GAN architectures to revolutionize defect detection in the PV industry.

## REFERENCES

[1] J. Chen, Y. Wen, Y. A. Nanehkaran, D. Zhang, and A. Zeb, "Multiscale attention networks for pavement defect detection," *IEEE Trans. Instrum. Meas.*, vol. 72, Jul. 2023, Art. no. 2522012.

[2] S. Deitsch et al., "Automatic classification of defective photovoltaic module cells in electroluminescence images," *Sol. Energy*, vol. 185, pp. 455–468, 2019.

[3] B. Su, Z. Zhou, and H. Chen, "PVEL-AD: A large-scale open-world dataset for photovoltaic cell anomaly detection," *IEEE Trans. Ind. Inform.*, vol. 19, no. 1, pp. 404–413, Jan. 2023, doi: 10.1109/TII.2022.3162846.

[4] X. Chen et al., "Automatic crack segmentation and feature extraction in electroluminescence images of solar modules," *IEEE J. Photovolt.*, vol. 13, no. 3, pp. 334–342, May 2023, doi: 10.1109/JPHOTOV.2023.3249970.

[5] M. Y. Demirci, N. Beşli, and A. Gümüşçü, "Efficient deep feature extraction and classification for identifying defective photovoltaic module cells in electroluminescence images," *Expert Syst. Appl.*, vol. 175, 2021, Art. no. 114810, doi: 10.1016/j.eswa.2021.114810.

[6] A. M. Karimi et al., "Generalized and mechanistic PV module performance prediction from computer vision and machine learning on electroluminescence images," *IEEE J. Photovolt.*, vol. 10, no. 3, pp. 878–887, May 2020, doi: 10.1109/JPHOTOV.2020.2973448.

[7] Y. Shi, L. Cui, Z. Qi, F. Meng, and Z. Chen, "Automatic road crack detection using random structured forests," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 12, pp. 3434–3445, Dec. 2016.

[8] M. Dhimish and V. Holmes, "Solar cells micro crack detection technique using state-of-the-art electroluminescence imaging," *J. Sci., Adv. Mater. Devices*, vol. 4, pp. 499–508, 2019, doi: 10.1016/j.jsamd.2019.10.004.

[9] A. S. Rajput et al., "Comparative study of the electrical parameters of individual solar cells in a c-Si module extracted using indoor and outdoor electroluminescence imaging," *IEEE J. Photovolt.*, vol. 10, no. 5, pp. 1396–1402, Sep. 2020.

[10] B. Su, H. Chen, Y. Zhu, W. Liu, and K. Liu, "Classification of manufacturing defects in multicrystalline solar cells with novel feature descriptor," *IEEE Trans. Instrum. Meas.*, vol. 68, no. 12, pp. 4675–4688, Dec. 2019, doi: 10.1109/TIM.2019.2900961.

[11] C. Mantel et al., "Machine learning prediction of defect types for electroluminescence images of photovoltaic panels," *Appl. Mach. Learn.*, vol. 11139, 2019, Art. no. 1113904, doi: 10.1117/12.2528440.

[12] M. U. Ali, H. F. Khan, M. Masud, K. D. Kallu, and A. Zafar, "A machine learning framework to identify the hotspot in photovoltaic module using infrared thermography," *Sol. Energy*, vol. 208, pp. 643–651, 2020, doi: 10.1016/j.solener.2020.08.027.

[13] A. Mellit and S. Kalogirou, *Handbook of Artificial Intelligence Techniques in Photovoltaic Systems: Modeling, Control, Optimization, Forecasting and Fault Diagnosis*. New York, NY, USA: Academic, 2022.

[14] M. Akram et al., "CNN-based automatic detection of photovoltaic cell defects in electroluminescence images," *Energy*, vol. 189, 2019, Art. no. 116319, doi: 10.1016/j.energy.2019.116319.

[15] B. Su, H. Chen, and Z. Zhou, "BAF-detector: An efficient CNN-based detector for photovoltaic cell defect detection," *IEEE Trans. Ind. Electron.*, vol. 69, no. 3, pp. 3161–3171, Mar. 2022, doi: 10.1109/TIE.2021.3070507.

[16] M. Xin and Y. Wang, "Research on image classification model based on deep convolution neural network," *EURASIP J. Image Video Process.*, vol. 2019, 2019, Art. no. 40, doi: 10.1186/s13640-019-0417-8.

[17] A. Di Tommaso et al., "A multi-stage model based on YOLOv3 for defect detection in PV panels based on IR and visible imaging by unmanned aerial vehicle," *Renewable Energy*, vol. 193, pp. 941–962, 2022, doi: 10.1016/j.renene.2022.04.046.

[18] M. Zhang and L. Yin, "Solar cell surface defect detection based on improved YOLO v5," *IEEE Access*, vol. 10, pp. 80804–80815, 2022, doi: 10.1109/ACCESS.2022.3195901.

[19] N. Drir and F. Chekired, "Automatic detection of solar cell surface defects in electroluminescence images based on YOLOv8 algorithm," *Indonesian J. Elect. Eng. Comput. Sci.*, vol. 32, pp. 1392–1404, 2023, doi: 10.11591/ijeecs.v32.i3.pp1392-1404.

[20] S. B. Jha and R. F. Babiceanu, "Deep CNN-based visual defect detection: Survey of current literature," *Comput. Ind.*, vol. 148, 2023, Art. no. 103911.

[21] E. Xie et al., "SegFormer: Simple and efficient design for semantic segmentation with transformers," in *Proc. 35th Int. Conf. Neural Inf. Process. Syst.*, 2021, vol. 34, pp. 12077–12090.

[22] A. Dosovitskiy, "An image is worth 16 × 16 words: Transformers for image recognition at scale," 2020, *arXiv:2010.11929*.

[23] N. Carion et al., "End-to-end object detection with transformers," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 213–229.

[24] D. Korkmaz and H. Acikgoz, "An efficient fault classification method in solar photovoltaic modules using transfer learning and multi-scale convolutional neural network," *Eng. Appl. Artif. Intell.*, vol. 113, 2022, Art. no. 104959, doi: 10.1016/j.engappai.2022.104959.

[25] A. Chen, H. Y. Li, H. Chen, and M. H. Li, "Anomaly detection algorithm for photovoltaic cells based on lightweight multi-channel spatial attention mechanism," *Energies*, vol. 16, 2023, Art. no. 1619, doi: 10.3390/en16041619.

[26] L. Chen, H. Yao, J. Fu, and C. T. Ng, "The classification and localization of crack using lightweight convolutional neural network with CBAM," *Eng. Struct.*, vol. 275, 2023, Art. no. 115291.

[27] Z. Chang, A.-J. Zhang, H. Wang, J. Xu, and T. Han, "Photovoltaic cell anomaly detection enabled by scale distribution alignment learning and multi-scale linear attention framework," *IEEE Internet Things J.*, vol. 11, no. 16, pp. 27816–27827, Aug. 2024.

[28] Z. Chang, K. Jia, T. Han, Y. Zhao, and H. Wang, "Towards more reliable photovoltaic energy conversion systems: A weakly-supervised learning perspective on anomaly detection," *Energy Convers. Manage.*, vol. 316, 2024, Art. no. 118845.

[29] Y. Y. Hong and R. A. Pula, "Diagnosis of photovoltaic faults using digital twin and PSO-optimized shifted window transformer," *Appl. Soft Comput.*, vol. 150, 2024, Art. no. 111092.

[30] I. Goodfellow et al., "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2020, doi: 10.1145/3422622.

[31] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," 2015, *arXiv:1511.06434*.

[32] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein generative adversarial networks," in *Proc. Int. Conf. Mach. Learn.*, 2017, pp. 214–223.

[33] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved training of Wasserstein GANs," in *Proc. 31st Inf. Conf. Neural Inf. Process. Syst.*, 2017, pp. 5769–5779.

[34] W. Tang, Q. Yang, K. Xiong, and W. Yan, "Deep learning-based automatic defect identification of photovoltaic module using electroluminescence images," *Sol. Energy*, vol. 201, pp. 453–460, 2020, doi: 10.1016/j.solener.2020.03.049.

[35] A. Karnewar and O. Wang, "MSG-GAN: Multi-scale gradients for generative adversarial networks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 7796–7805.

[36] A. Undirwade and S. Das, "Image anonymization using deep convolutional generative adversarial network," in *Machine Learning Algorithms and Applications*, M. Srinivas, G. Sucharitha, A. Matta, and P. Chatterjee, Eds. Boca Raton, FL, USA: CRC Press, 2021.

[37] T. Schlegl, P. Seebock, S. M. Waldstein, U. Schmidt-Erfurth, and G. Langs, "Unsupervised anomaly detection with generative adversarial networks to guide marker discovery," in *International Conference on Information Processing in Medical Imaging*, vol. 10265. Berlin, Germany: Springer, 2017, pp. 146–147.

[38] PVEL-AD, Binyisu, GitHub repository. Accessed: Feb. 28, 2024. [Online]. Available: https://github.com/binyisu/PVEL-AD

[39] G. Kasi, S. Abirami, and R. D. Lakshmi, "A deep learning based cross model text to image generation using DC-GAN," in *Proc. 12th Int. Conf. Adv. Comput.*, Chennai, India, 2023, pp. 1–6, doi: 10.1109/ICoAC59537.2023.10250086.