Article

A Novel Method for Localized Typical Blemish Image Data Generation in Substations

Na Zhang *, Jingjing Fan, Gang Yang, Guodong Li, Hong Yang and Yang Bai

State Grid Shanxi Electric Power Research Institute, Taiyuan 030001, China
* Correspondence: zhangna19907@163.com

Abstract: Current mainstream methods for detecting surface blemishes on substation equipment typically rely on extensive sets of blemish images for training. However, the unpredictable nature and infrequent occurrence of such blemishes present significant challenges in data collection. To tackle these issues, this paper proposes a novel approach for generating localized, representative blemish images within substations. Firstly, to mitigate global style variations in images generated by generative adversarial networks (GANs), we developed a YOLO-LRD method focusing on local region detection within equipment. This method enables precise identification of blemish locations in substation equipment images. Secondly, we introduce a SEB-GAN model tailored specifically for generating blemish images within substations. By confining blemish generation to identified regions within equipment images, the authenticity and diversity of the generated defect data are significantly enhanced. Theexperimental results validate that the YOLO-LRD and SEB-GAN techniques effectively create precise datasets depicting flaws in substations.

Keywords: augmentation of blemish images in substations; YOLO; GAN; localized area blemish image generation

MSC: 68T07

1. Introduction

1.1. Background

Electricity is an important factor in guaranteeing the development of the national economy, and a safe and reliable power supply is related to the development of various industries in society. The substation plays a critical role in the power system by facilitating the transmission of electricity from power plants to end-users through voltage conversion between high and low levels. Due to exposure to outdoor environments, important electrical components such as insulators, high-voltage lines, and fittings are subjected to prolonged sunlight, wind, rain, snow, hail, acid and alkali corrosion, electrical arcing, foreign object suspension, and other influences. Compounded by limited production processes for electrical components, materials are prone to aging and sustained tension effects. Consequently, various apparent blemishes may occur in different electrical equipment, such as insulator stringing and contamination, foreign objects hanging on towers, and metal corrosion. If not detected and effectively dealt with in a timely manner, these apparent blemishes may lead to failures in the power grid system, causing interruptions in power transmission and resulting in large-scale power outage accidents, which can inflict significant economic losses on society. Nowadays, identifying issues like metal rust, oil leaks, and damaged insulators in substations usually require manual inspections. These inspections are slow and demand a lot of manpower [1]. With the advancement of robotics technology, intelligent inspection methods are gradually integrating into the process of power grid inspections [2,3]. Inspection robots first follow predefined routes to inspect power equipment, during which onboard devices collect information from the substations [4,5]. Through image processing techniques, the inspection images are analyzed

to detect faults in the power equipment. During the detection process, a large amount of image data is collected. If these were to be observed by human eyes, there could be issues of misjudgment or oversight, making it difficult to efficiently identify safety hazards in power equipment and significantly increasing maintenance costs. Therefore, intelligent surface blemish detection methods based on image processing technology are essential for inspecting power equipment [6,7].

1.2. Related Works

1.2.1. Target Detection Methods

In recent years, as processor performance continues to improve, machine learning methods have also been evolving [8]. Deep learning-based object detection methods have begun to advance at an unprecedented pace [9]. Nowadays, there are mainly two main approaches to deep learning-based object detection. The first approach is detection algorithms based on target region proposal, such as Fast R-CNN, Faster R-CNN, and Mask R-CNN [10–12]. The second approach is detection algorithms based on integrated regression networks, such as YOLOv1-v8 [13]. These methods have become popular technologies applied in defect detection.

The existing blemish detection methods for substation equipment utilizing the above algorithms necessitate a substantial training dataset to ensure accuracy. Nevertheless, establishing datasets for surface blemishes encounters several challenges: Firstly, during the collection of image data, there exists a notable disparity in the quantity of normal images of substation equipment compared with images depicting blemishes [14]. This places a substantial burden on the human review and categorization of defective images. Secondly, differences in manual expertise lead to varying levels of accuracy in the collection, classification, and annotation of blemish images, presenting challenges in maintaining dataset quality. Thirdly, current image augmentation algorithms primarily rely on techniques such as rotation often neglecting super-pixel characteristics. Consequently, the dataset exhibits considerable redundancy and fails to substantially improve the recognition and detection performance of deep learning models.

1.2.2. Deep Learning-Based Blemish Image Generation Methods

To tackle the challenges mentioned above, numerous researchers have utilized image augmentation techniques in a broad spectrum of applications, such as cross-modality insulator augmentation for multidomain insulator defect detection [15], enhanced detection of subway insulator blemishes [16], and so on. These researchers have explored image generation techniques utilizing superpixel characteristics. Generative models such as generative adversarial networks (GAN) [17] play a crucial role in creating new images featuring intricate backgrounds and diverse attributes. This technique aids in developing training datasets that are more comprehensive, thus improving the accuracy and efficiency of blemish diagnosis algorithms. Yang et al. [18] proposed a perceptual capsule cycle generative adversarial network (PreCaCycleGAN) for industrial defect sample augmentation, generating realistic and diverse defect samples from defect-free real samples. Di Maggio et al. [19] introduced an innovative approach that leverages CycleGAN [20] to synthesize data for faulty industrial equipment. Through the training of the CycleGAN model, wavelet images mimicking vibration signals are converted into authentic data representing mechanically impaired bearings. Liu et al. [21] utilized DCGAN [22] to create numerous synthetic samples based on pre-existing concrete crack images. These artificial samples were carefully curated to be diverse and information-rich. Zhuang et al. [23] utilized Defect-GAN [24] to enlarge a limited dataset of plastic blemish images. In response to the scarcity of abnormal blemish images on train surfaces, Liu et al. [25] introduced an innovative approach. Their Anomaly-GAN effectively maintains the reality of the created blemish images at both overall and specific levels.

While the image enhancement techniques utilizing adversarial deep learning have shown promise in generating new industrial blemish images, there are specific challenges when adapt-

ing them to create substation equipment blemish images. Firstly, the real-world environments in which substation equipment images are captured often feature intricate backgrounds, posing a hurdle for existing adversarial deep learning methods that tend to diminish such complexities. These methods generally alter the overall image style rather than targeting specific regions of interest, which is crucial for substation equipment blemish generation. Secondly, substation equipment blemishes, like rust and oil leakage, exhibit diverse and intricate features that extend beyond simple imperfections such as scratches. Existing GAN approaches, focusing on simpler blemishes, may struggle to accurately replicate the rich textures and color variations present in substation equipment blemishes, resulting in lower success rates. Lastly, the quality of generated images requires enhancement. Common issues such as mode collapse during GAN training can obscure blemish details, while unrealistic or unstable generated images may disrupt the coherence between the blemishes and the overall image, diminishing the natural appearance of the blemishes.

1.3. Main Contributions

To solve the challenge of scarce blemish image data for substation equipment, which does not meet the supervised deep learning training data requirements, a method for generating localized typical blemish images in substations is proposed. The key contributions and innovations are as follows:

(1) To address the issue of global style variation in images generated by GAN methods, we have proposed a method for local region detection on equipment. We utilize an improved YOLOv7 [26] method to accurately detect potential blemish locations in substation equipment images.

(2) We use a GAN model for generating blemish images in substations. By generating blemishes on localized images of detected substation equipment, the effective generation of blemish data is achieved.

(3) The above method preserves the features of the original images to a great extent, while also generating different types of blemishes on multiple devices within the same image. This addresses the limitations of original images having a single blemish type and few blemishes. Experimental validation has shown that the dataset generated by this method effectively enhances the precision of mainstream surface blemish detection methods.

The following sections are organized as follows: Section 2 outlines the method used to detect specific areas of substation equipment. Section 3 offers a detailed explanation of the GAN model employed to generate defective images. Section 4 presents experimental findings and discussions. Finally, the Section 4 summarizes this study's findings.

2. Methodologies

2.1. Local Area Detection of Substation Equipment

In the realm of substations, the presence of common imperfections like surface oil contamination, oil leakage, metal corrosion, and cracked insulator sleeves has long been a cause for concern. These flaws not only jeopardize the well-being of substation equipment but also pose a threat to electrical safety. The visual representation of these imperfections can be seen in Figure 1. Given the haphazard nature of blemish occurrences, capturing images of these blemishes is no easy feat, let alone capturing multiple blemishes in a single image. The efficacy of supervised deep learning-based blemish detection hinges on having a substantial number of images featuring singular blemish categories. Furthermore, the variety of blemish types depicted in these images plays a crucial role in refining the accuracy of blemish detection models.

To address the data requirements for training blemish detection models, we introduce a novel approach based on an enhanced YOLOv7 method for detecting local device regions (YOLO-LRD) to generate multiclass blemish images overlaid on normal equipment images, as depicted in Figure 2. This method identifies potential locations for blemish formation

on equipment, confines the area where blemishes may appear, and ensures the creation of blemish images without compromising the overall image aesthetics.
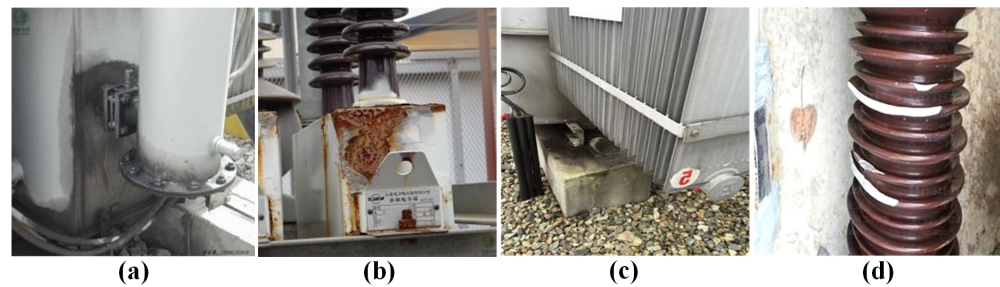

Figure 1. Typical blemishes of substation equipment: (**a**) Surface oil contamination. (**b**) Metal corrosion. (**c**) Oil leakage on the ground. (**d**) Cracked insulator sleeves.
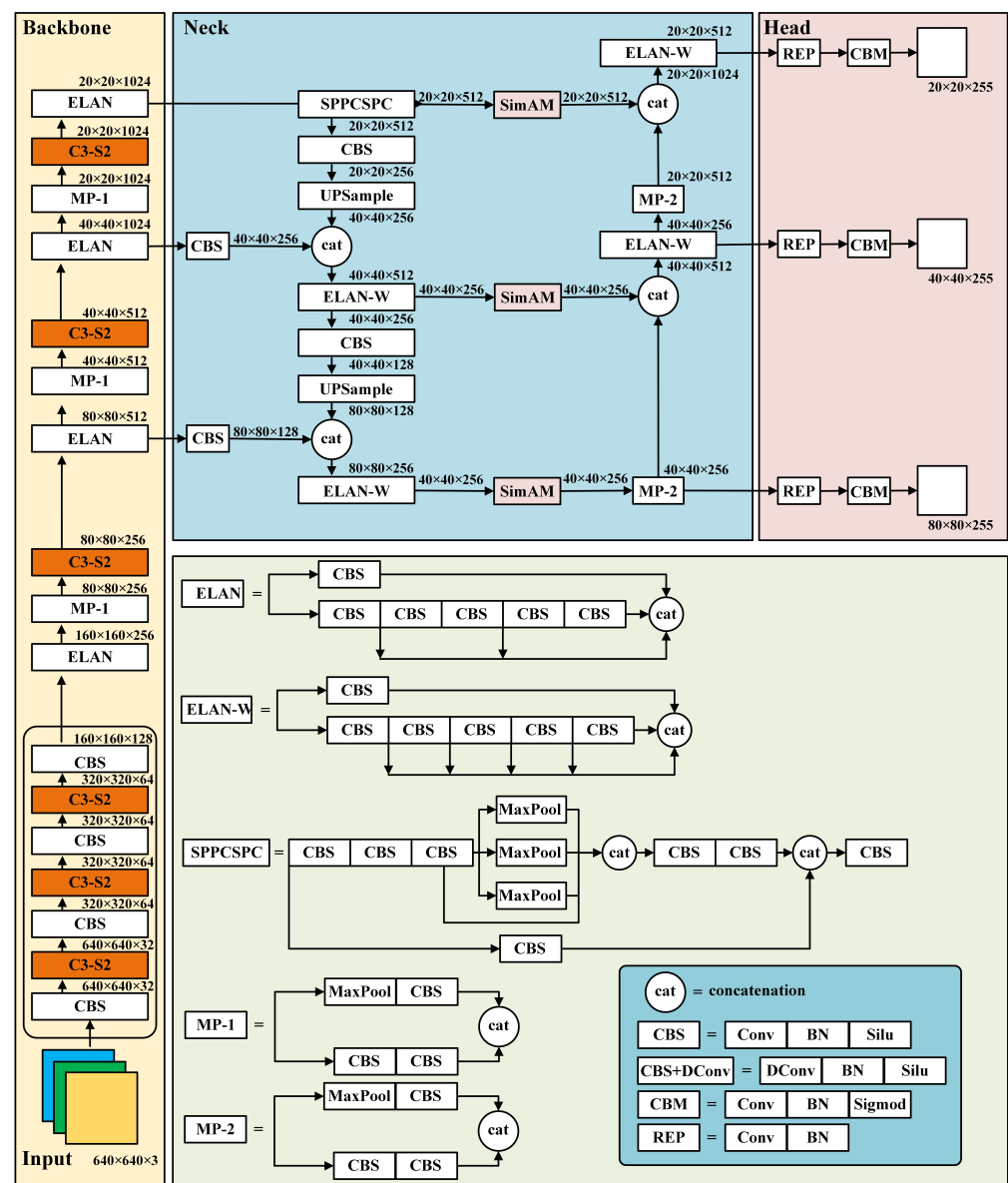

Figure 2. Structure diagram of YOLO-LRD model.

YOLO-LRD incorporates the C3-S2 block, an advanced feature extraction unit that enhances both local detail focus and overall processing. This improvement is crucial for

accurately identifying key features in substation images. Additionally, the integration of SimAM into the neck network of YOLO-LRD improves the network's learning ability and supports adaptive attention weight tuning for diverse input scenarios. The new WD-CIoU fusion loss function also addresses challenges related to size and shape variations in substation equipment images. Together, these innovations highlight YOLO-LRD's significant advancements in local region detection for substation imagery.

2.1.1. YOLOv7 Model

The YOLOv7 model is structured into three main components: Backbone, Neck, and Head. In the Backbone, images are initially resized to 640×640 pixels, processed through four convolutional layers, and then processed by an ELAN module along with three MP1-ELAN modules. The ELAN module improves learning efficiency in deep networks by managing gradient flow. The Neck module handles feature extraction and detection output generation, utilizing components such as CBS, SPPCSPC, UpSampling, MPConv, and ELAN-W for enhanced feature processing. The Head segment, featuring REP and CBM blocks, refines the final output by fusing feature maps at various downsampling levels using upsampling to produce the network's results.

2.1.2. C3-S2 Block

The C3 block plays a crucial role in YOLOv5 [27], enhancing the network's depth and perceptive range for better feature extraction. It consists of three convolutional (Conv) blocks, where the first block uses a stride of 2 to decrease the feature map dimensions by 50%, expanding the perceptive range while controlling computational burden. The subsequent two Conv blocks use a stride of 1, which helps maintain detail resolution and preserve fine details of local targets. Each Conv in the C3 block employs 3×3 kernels, and utilizes Batch Normalization (BN) layers and LeakyReLU function to stabilize the block and optimize its effectiveness.

To optimize the YOLOv7 backbone with greater depth and a broader perceptive range, as well as to improve feature processing efficiency, the upgraded C3-S2 block is presented, illustrated in Figure 3. This module incorporates principles from the Swin Transformer and Switchable Atrous Convolution (SAC), facilitating smooth feature transfer between various windows in the network. It effectively maintains detailed depth information, extends the network's perceptual range, and employs attentional strategies to further explore feature information. Leveraging a residual block architecture and advanced feature extraction components, the C3-S2 block significantly boosts the YOLO-LRD backbone's ability.
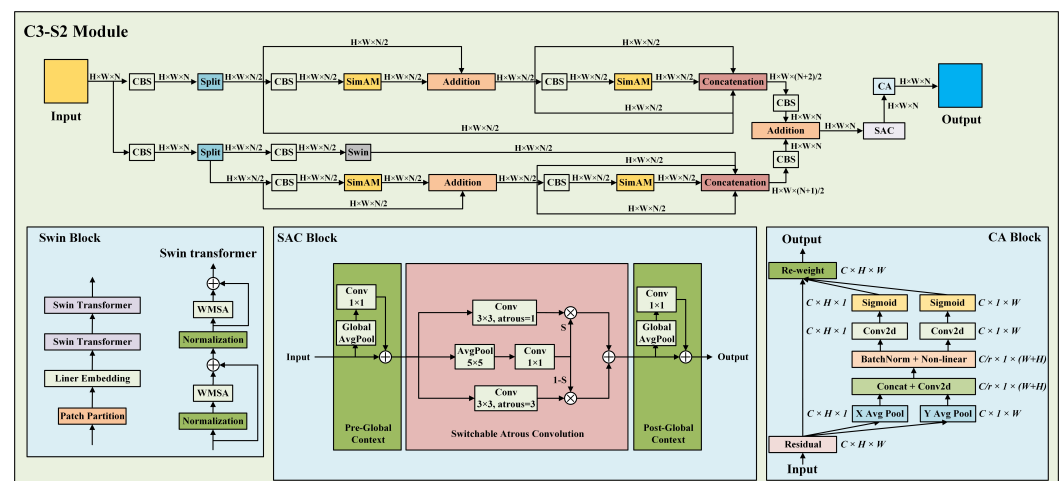

Figure 3. Structure diagram of C3-S2.

The Swin block leverages key aspects of the Swin Transformer. Our study combines the Subregion Partitioning and Dimensional Reduction techniques into this block. Firstly,

the Subregion Partitioning technique divides the input activation map into uniform size, disjoint segments, facilitating the independent analysis of each segment. This division enhances the model's ability to efficiently acquire detailed regional attributes, thereby boosting processing effectiveness. After the Subregion Partitioning, the Dimensional Reduction process reduces the feature complexity, converting the input information into a more compact form. The resultant feature data is subsequently channeled into the Swin Transformer module. This structured approach supports effective information modification and engagement, which is essential for thorough feature extraction and aggregation.

The Switchable Atrous Convolution (SAC) block combines PreGlobal Awareness and PostGlobal Awareness mechanisms, alongside a switching block. This set-up enables adaptive modulation of the expansion parameter and switch parameter within the SAC framework, enhancing the capture of multiscale target details. This strategy alleviates information degradation and improves the network's performance in processing image features. Importantly, a flexible switching structure is utilized to manage the equilibrium between Typical Conv and Atrous Conv, thereby seamlessly integrating Atrous Conv with the switching structure to detail the SAC block.

In this study, the Coordinate Attention (CA) structure is incorporated subsequent to the C3-S2 block after SAC. Its main goal is to mitigate the loss of overall spatial weights, improve the backbone emphasis during characteristic selection, and instantly provide accurate attribute specifics to the subsequent network layer to minimize needless repetition. The central concept of CA is to embed spatial data in the feature channel axis, splitting channel-focused attention into feature consolidation throughout two directional dimensions. One addresses distant relationships, while the other maintains exact spatial details. Ultimately, these aspects are combined to assess the channel relevance weights, improving direct attention to intricate target characteristics while minimizing superfluous or interference-prone channels. The process involves two primary phases: integrating positional data and creating spatial focus, as shown in Figure 3.

2.1.3. SimAM in the Neck

The Simple, Parameter-Free Attention Module (SimAM) [28] refines network focus by fine-tuning the energy evaluation to measure the impact of single neurons, modifying focus coefficients in real time. A key benefit of SimAM is that it delivers these advancements without adding extra parameters to the detection model. Rather, it utilizes its efficient parameter set to compute 3D attention weight allocations for feature layers.

Due to its straightforward nature, parameter efficiency, and flexibility, we incorporate SimAM into the Neck of YOLO-LRD. This inclusion boosts the Neck network's feature extraction ability by adjusting attention coefficient allocations for diverse input cases, such as complicated object features, small details from far-off scales, and large attributes from close ranges. This approach enables the Neck to concentrate on comprehensive substation device features while eliminating unrelated data, thus refining feature integration and enhancing YOLO-LRD's detection precision.

2.1.4. WD-CIoU Loss

To accurately extract unique attributes, it is crucial to balance target and nontarget instances. Although the detection of objects across different sizes is widespread in actual substation environments, current YOLO models are chiefly optimized for targets of regular scales. The Intersection over Union (IoU) metric, which is size-sensitive, causes significant variations in IoU results for targets of different sizes. This issue is especially noticeable with tiny objects, where even slight location changes can lead to substantial IoU fluctuations. Conversely, comparable adjustments for standard-scaled targets have a negligible effect on IoU scores. This size dependence results in sudden shifts in bounding box positions, eventually affecting arrangement accuracy.

To tackle this challenge, an innovative assessment method utilizing Wasserstein Distance for detecting tiny targets has been proposed [29]. This approach models bounding

boxes as two-dimensional normal profiles and employs the standardized Wasserstein Distance (WD) to evaluate the correspondence between these profiles. WD can be smoothly incorporated into the allocation, nonmaximal value elimination, and loss criteria of anchor-based models, acting as an alternative to the typical IoU function. This method provides two key benefits: it precisely assesses the distributional resemblance for tiny objects even without any intersection, and its unresponsiveness to size enhances its effectiveness in evaluating tiny object similarities.

For tiny targets, nonobject pixels often infiltrate the bounding box since ground truth targets rarely align perfectly inside a bounding box. Generally, ground truth pixels cluster in the center, with background pixels being more frequent near the periphery. To more precisely apply weights to each pixel within the rectangular region, one effective strategy is to frame it as a 2D Gaussian distribution $N(\epsilon, \zeta)$, as illustrated below:

$$\epsilon = \begin{bmatrix} c_x \\ c_y \end{bmatrix}, \zeta = \begin{bmatrix} \frac{w^2}{4} & 0 \\ 0 & \frac{h^2}{4} \end{bmatrix}, \tag{1}$$

where c_x, c_y denotes the central point coordinates, and w and h demonstrate the dimensions of the bounding box. For two 2D Gaussian allocations $\epsilon_1 = N(n_1, \zeta_1)$ and $\epsilon_2 = N(n_2, \zeta_2)$, the second-stage WD is as follows:

$$WD_2^2(\epsilon_1, \epsilon_2) = ||n_1 - n_2||_2^2 + ||\zeta_1^{1/2} - \zeta_2^{1/2}||_F^2, \tag{2}$$

where $|| \cdot ||$ denotes the F-norm. By employing Gaussian allocations N_p and N_r, where N_p depicts the forecasted bounding box $B_p = (cx_p, cy_p, w_p, h_p)$ and N_r illustrates the real data bounding box $B_r = (cx_r, cy_r, w_r, h_r)$, the second-stage WD within these two bounding boxes can be expressed as follows:

$$WD_2^2(N_p, N_r) = ([cx_p, cy_p, \frac{w_p}{2}, \frac{h_p}{2}]^T, [cx_r, cy_r, \frac{w_r}{2}, \frac{h_r}{2}]^T)_2^2, \tag{3}$$

Because $WD_2^2(N_p, N_r)$ is measured in distance metrics, while the value for bounding box correspondence should lie within the $(0, 1)$ range, it is essential to normalize $WD_2^2(N_p, N_r)$ as follows:

$$WD(N_p, N_r) = exp[-\frac{WD_2^2(N_p, N_r)}{CLASS}]. \tag{4}$$

Finally, WD loss is shown as follows.

$$L_{WD} = 1 - WD(N_p, N_r). \tag{5}$$

The WD loss function is effective for tiny-size objects. Nevertheless, because substation device images vary in scale, the CIoU loss [30] is included as follows:

$$IoU = \frac{R_p \cap R_t}{R_p \cup R_t}, \tag{6}$$

$$v = \frac{4}{\pi^2}(arctan\frac{w_t}{h_t} - arctan\frac{w_p}{h_p})^2, \tag{7}$$

$$\alpha = \frac{v}{(1 - IoU) + v}, \tag{8}$$

$$L_{CIoU} = 1 - IoU + \frac{ed^2}{c^2} + \alpha v, \tag{9}$$

where R_p denotes the forecasted bounding box and R_t indicates the real box. The parameter v evaluates the alignment of the width and height proportions. w_t and h_t denote the scale

of the real bounding box, while w_p and h_p denote the scale of the forecasted bounding box. Moreover, α assesses the alignment of the dimension ratios between the forecasted and real bounding boxes. Moreover, *ed* illustrates the geometric distance between the centroids of the forecasted and real bounding boxes, and *c* represents the diagonal distance of the minimal surrounding rectangle that encompasses both the boxes.

By allocating a suitable merging weight Λ to L_{WD} and L_{CIoU}, we propose the WD-CIoU loss function as follows:

$$L_{WD-CIoU} = \Lambda \cdot L_{CIoU} + (1 - \Lambda) \cdot L_{WD}. \tag{10}$$

where Λ is between 0 and 1. Through multiple experiments, we verified that the final $L_{WD-CIoU}$ can be minimized when Λ is set to 0.7.

2.2. Blemish Generation Algorithm

After successfully detecting local areas of substation equipment, we employ a GAN method to generate blemishes within these detected regions. This section introduces two key concepts and their underlying principles: the local area blemish generation model for substation equipment and the objective loss function. These concepts are utilized to create corresponding blemish images for the equipment.

2.2.1. Algorithm Principle of GAN

The GAN model is a commonly adopted algorithm for generating images, trained on a dataset to produce resembling samples [31]. A GAN is composed of two essential components: the generator and the discriminator. The generator learns the dispersion of the trial data and employs noise *z* from a specified dispersion to create data that mirrors actual input data. Its objective is to generate progressively realistic samples. Meanwhile, the discriminator acts as a two-class decision model, assessing the likelihood that a given sample originates from the real data rather than being created.

Currently, most GAN algorithms can only transform the global style of an image and cannot be confined to a specific area, which can lead to distortion of the image's style. For substation equipment blemish images with complex backgrounds, overall image distortion can affect the quality of the generated dataset, thereby reducing the precision of detection models. To overcome this limitation, a substation equipment blemish generation model (SEB-GAN) is proposed, based on the local area detection of substation equipment obtained using the YOLO-LRD model, depicted in Figure 4. SEB-GAN is designed to convert images from one category to another, using two data sets: *N* (normal equipment images) and *B* (blemish images). The primary goal is to map samples from *N* to *B*. SEB-GAN achieves this by learning a mapping function *G* that transforms *N* into *B*. Here, *G* serves as the generator in the SEB-GAN model, producing an image $F(n)$ in set *B* from a sample *n* in set *N*. The discriminator D_B in SEB-GAN assesses whether the generated image $F(n)$ is true, structuring a GAN model. SEB-GAN enables the reversible conversion of normal and blemished images. In this paper, we concentrate on converting typical images of substation devices to blemished ones.

2.2.2. Localized Blemish Generation Model

In this study, a localized blemish generation algorithm is designed using a dual-path network built upon the U-Net structure [32], as illustrated in Figure 4. The U-Net is composed of two primary segments: a feature extraction pathway and a reconstruction pathway. The feature extraction pathway operates as a standard CNN with a recurring design for activation layer processing utilizing ConV-BN-ReLU stages. The reconstruction pathway starts each step with a deconvolution operation, which increases the feature map dimensions and reduces the count of feature channels by 50%. The resulting feature map is then aggregated with the aligned map from the feature extraction pathway and backpropagated. This dual-path mechanism, combined with feature merging, enables the integration of surface-level positional features with in-depth categorical features, enhancing

both feature efficiency and image synthesis. Thus, this study uses the U-Net structure rather than a CNN for the localized blemish generation model.
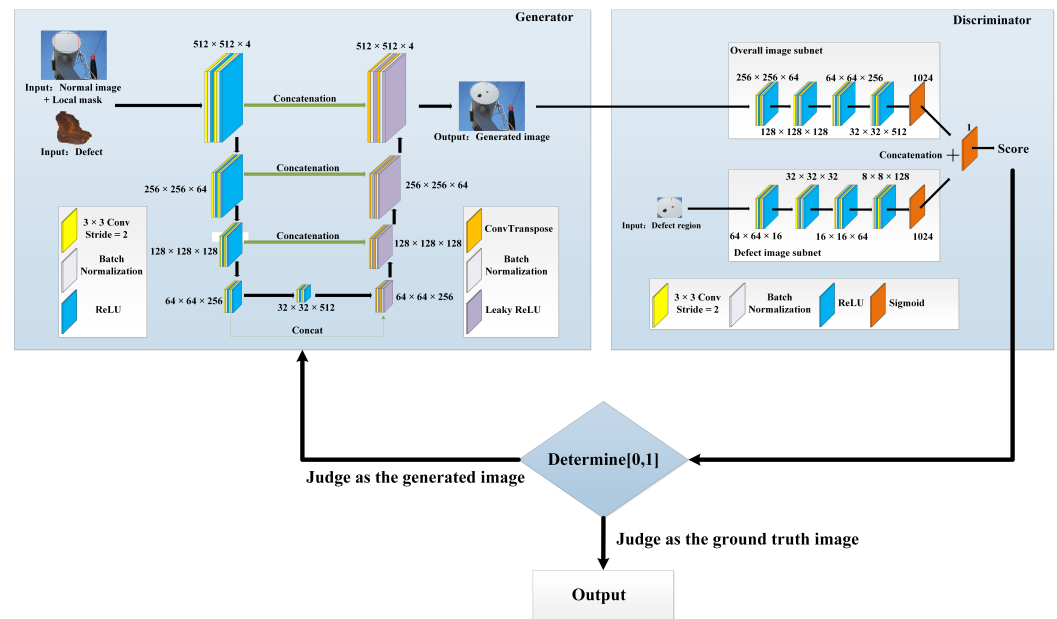

Figure 4. Structure diagram of SEB-GAN model.

2.2.3. Composite Discriminator

The assessment of created images addresses two primary factors. The first factor is whether the image correctly depicts substation devices. The second factor is whether the generated blemishes exhibit characteristics typical of surface blemishes. To achieve these objectives, we introduce a composite discriminator that evaluates both the complete image and the blemish. This discriminator consists of two distinct networks: one dedicated to the blemish region and the other to the entire image. The composite discriminator integrates the results from these networks to determine the realism of the created image, thereby improving blemish quality. The structure of the composite discriminator is shown in Figure 4. The vectors produced by the two distinct networks are concatenated and fed into a Sigmoid layer, yielding a score between 0 and 1 for the generated image. This score indicates the likelihood that the image is real.

2.2.4. Loss Function

We use the convergence of the loss curve of the SEB-GAN model to determine whether the GAN network has converged. When the loss function of the SEB-GAN gradually decreases and converges to a minimum value during training, we consider the GAN network to be converged.

Based on the SEB-GAN model's architecture, the loss function includes three parts [14]. The loss function for the generator $G_{norm2def}$, responsible for converting blemish-free images into defective ones, along with its corresponding discriminator $D1_{def}$, is formulated as follows:

$$
\begin{aligned}
L_{glo}(G_{norm2def}; D1_{def}; I_{def}; MI_{def}; F_{def}; MF_{def}) = \\
E_{I_{norm} \in P_{data}(I_{norm})}[lgD1_{def}(I_{def}, MI_{def})] \\
+ E_{I_{def} \in P_{data}(I_{def})}[lg(1 - D1_{def}(F_{def}, MF_{def}))],
\end{aligned}
\tag{11}
$$

where I_{norm} denotes the blemish-free image, I_{def} represents the collected blemish image, MI_{def} represents the blemish image in YOLO-LRD detected area, F_{def} stands for generated image, and MF_{def} refers to the generated image in the detected area. This equation describes a process aimed at enhancing the discriminator $D1_{def}$'s ability to identify local blemishes,

while simultaneously minimizing the discrepancy between the generated image and its corresponding blemish-free image in terms of local blemish perception, governed by $G_{norm2def}$. As $D1_{def}(I_{def}, MI_{def})$ approaches 1, the discriminator effectively distinguishes authentic samples. Throughout training, the discriminator remains fixed, and improved performance by $G_{norm2def}$ can sometimes lead the discriminator to misclassify generated images as originals, thereby increasing $D1_{def}(I_{def}, MI_{def})$, and reducing overall loss.

Given that surface blemishes on devices typically occupy only a small portion of the surface, we introduce a local defect perceptual loss within the generator section of the network to enhance the quality of defect generation:

$$L_{def}(MI_{norm}; MF_{norm}; MI_{def}; MF_{def}) = L_{semantic}(MF_{norm}; MI_{norm})$$
$$+ L_{semantic}(MF_{def}; MI_{def}), \tag{12}$$

where $L_{semantic}(MF_{norm}; MI_{norm})$ represents the difference loss between the blemish-free area generated within the YOLO-LRD detected area MF_{norm} and the actual blemish-free image MI_{norm}. Similarly, $L_{semantic}(MF_{def}; MI_{def})$ denotes the difference loss between the defective portion generated within the YOLO-LRD detected area MF_{def} and the real defective image MI_{def}. This loss component aims to ensure high image quality in both defect-free and defective regions within the mask during generator training. A lower loss value indicates superior image quality within the mask.

In practical applications, the scarcity of actual defect images compared with the abundant normal images may result in a lack of diversity in the generated defect image features. This presents a challenge when using local perceptual D2 loss to ensure high quality and varied characteristics in generated surface defect images. To address this issue, this study introduces cycle consistency loss [20]. This method aims to prevent this scenario and achieve the goal of generating defect images with diverse features using normal images. The cycle consistency loss assumes the existence of a mapping $H_{def2norm}$ that can transform the generated defect images back into the space of defect-free sample images. By establishing a one-to-one mapping between normal and defect sample image spaces, this approach mitigates the issue of repetitive features in generated defect sample images. Cycle consistency loss can be formulated as follows:

$$L_{cyc}(G_{norm2def}; H_{def2norm}) = E_{I_{norm} \in P_{data}(I_{norm})}||F_{norm} - I_{norm}||$$
$$+ E_{I_{def} \in P_{data}(I_{def})}||F_{def} - I_{def}||. \tag{13}$$

3. Experiments and Discussion

Following the construction of the YOLO-LRD and SEB-GAN models, we proceeded with experiments conducted on an IW4210-8G server running Ubuntu 18.04. Training of both models utilized GPU resources.

3.1. Experimental Preparation

In this paper, we generated several typical blemishes, including surface oil contamination on equipment, oil leakage on the ground, metal corrosion, and cracked insulator sleeves. Therefore, in the equipment surface area detection stage, we collected images of the substation equipment surface area, the substation equipment ground, and the insulator sleeves. For each category, 2000 images were collected and annotated, constructing a training dataset for the substation equipment surface images used to train the YOLO-LRD model. The YOLO-LRD model initialization involves setting the parameters detailed in Table 1. The SEB-GAN model initialization involves setting the parameters detailed in Table 2. To achieve peak performance, in YOLO-LRD, input images are resized to 640 × 640 pixels and training is carried out with a batch size of 8. We utilize key parameters from the YOLOv7 model, including momentum, initial learning rate, and decay, due to their established effectiveness. An extensive training regimen of 4000 steps provides a thorough understanding of the process. In SEB-GAN, input images are resized to 256 × 256 pixels

and training is carried out with a batch size of 1. We utilize key parameters from the Cycle-GAN model, including training steps, momentum, and initial learning rate due to their established effectiveness. The learning rate is adjusted dynamically to enhance model performance and convergence. By following these protocols, we train the YOLO-LRD and the SEB-GAN models. The process of generating defects in substation equipment is illustrated in Algorithm 1.

Algorithm 1 Workflow for Generating Defective Images using YOLO-LRD and SEB-GAN

1: **Step 1:** Image Acquisition
2: Collect images of substation equipment surfaces $\{I_1, I_2, \ldots, I_n\}$
3: **Step 2:** Annotation and Dataset Construction
4: Construct annotated dataset $\{(I_i, B_i)\}$ where B_i denotes bounding boxes
5: **Step 3:** YOLO-LRD Model Construction
6: Initialize YOLO-LRD model architecture
7: **Step 4:** YOLO-LRD Model Training
8: Train YOLO-LRD model using dataset $\{(I_i, B_i)\}$
9: **Step 5:** Obtain Trained YOLO-LRD Model
10: Trained YOLO-LRD model: $M_{YOLO-LRD}$
11: **Step 6:** Dataset Construction for Defects and Normal Surfaces
12: Construct datasets: normal surface images $\{I_{normal}\}$ and defect images $\{I_{defect}\}$
13: **Step 7:** SEB-GAN Model Construction
14: Initialize SEB-GAN model architecture
15: **Step 8:** SEB-GAN Model Training
16: Train SEB-GAN model using datasets $\{I_{normal}\}$ and $\{I_{defect}\}$
17: **Step 9:** Generate Defective Images
18: **for** each new image I_i in substation equipment **do**
19:     Detect surface area using $M_{YOLO-LRD}$: B_i
20:     Generate defect in B_i using $M_{SEB-GAN}$
21:     Obtain defective image $I_{defective}$
22: **end for**
23: **return** Generated defective images $\{I_{defective}\}$

Table 1. Training parameters of YOLO-LRD.

Size of input images	640×640
Batch	8
Optimizer	Adam
Momentum	0.999
Learning rate	0.001
Decay	0.0005
Training steps	4000

Table 2. Training parameters of SEB-GAN.

Size of input images	256×256
Batch	1
Optimizer	Adam
Momentum	0.5
Learning rate	0.0002
Training steps	150,000

We then compiled a dataset of 700 images showcasing various surface blemishes on substation equipment from field scenarios: 353 images of surface oil contamination on equipment, 233 of oil leakage on the ground, 347 depicting metal corrosion, and 175 showing cracked insulator sleeves. Furthermore, 1200 images of blemish-free substation equipment were collected for training the SEB-GAN model. We then generate 2000 images of substation equipment blemishes using the SEB-GAN model. In these generated images, there are 500 images for each type of blemish.

This study evaluates the blemish image dataset generated by the SEB-GAN model using the YOLOv7 target detection model. Performance is measured using metrics such as Mean Average Precision (mAP), Precision, Recall, and F1 Score. The specific calculation formulas are as follows:

$$Precision = \frac{TP}{TP + FP}, \tag{14}$$

where TP is True Positives, and FP is False Positives.

$$Recall = \frac{TP}{TP + FN}, \tag{15}$$

where FN is False Negatives.

$$mAP = \frac{1}{n} \sum_{i=1}^{n} AP_i, \tag{16}$$

where n is the number of classes, and AP_i is the average precision for class i. AP_i is the area under the precision–recall curve for class i.

$$AP_i = \int_0^1 P_i(R_i) dR_i, \tag{17}$$

where $P_i(R_i)$ is the precision of class i under recall R_i.

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}, \tag{18}$$

3.2. Complexity of the Proposed Method

Our method consists of two main components: YOLO-LRD for surface area detection of substation equipment and SEB-GAN for generating blemish images on the detected areas. Analyzing the complexity involves several aspects:

1. Model Complexity: YOLO-LRD has a parameter size of 53.7 MB and 49 GFOLPs, making it suitable for real-time detection. SEB-GAN has a parameter size of 43.9 MB and 146 GFOLPs, indicating greater computational demands for generating defect images.
2. Memory Usage: YOLO-LRD uses 6.4 G of GPU memory, while SEB-GAN requires 8.0 G. This difference in memory demand can impact overall performance, especially in resource-limited environments.
3. Training and Inference Time: YOLO-LRD has a faster inference speed during detection, but SEB-GAN's training and generation times are longer, requiring a balance between real-time performance and generation quality.

Overall, our approach creates a complex dependency between real-time detection and image generation, which may present challenges in performance and resource management. Delving into the optimization of interactions between these two models can help reduce overall complexity.

3.3. Effect of Generated Blemishes

After training the YOLO-LRD and SEB-GAN models, we created a test set with blemish-free substation equipment images to validate the blemish generation methods for local areas. The generated images are shown in Figure 5.

Figure 5. Blemish image generated by YOLO-LRD and SEB-GAN models: (**a**–**d**) Blemish-free image. (**a1**–**d1**) Detection results of YOLO-LRD. (**a2**–**d5**) The generated blemish images.

Figure 5a–d shows the original images collected from the substation, which are free of blemishes. Figure 5a1–d1 displays the local areas of substation equipment detected by the YOLO-LRD model, including the equipment surface, ground, and insulator sleeves. Figure 5a2–d5 illustrates the blemishes generated by the SEB-GAN model in the detected local areas, such as metal corrosion, surface oil contamination, oil leakage on the ground, and cracks in the insulator sleeves.

The results demonstrate that the proposed method for generating blemishes on substation equipment effectively maintains the original image's authenticity while creating realistic blemishes in likely locations. Moreover, it can generate various types of blemishes within a single image, facilitating the creation of numerous blemish samples. This capability meets the training data needs for supervised learning-based defect detection models.

3.4. Ablation Experiments of YOLO-LRD

YOLO-LRD introduces two primary elements: the C3-S2 block and the SimAM. To determine their effects on performance, a set of ablation tests were carried out. The comprehensive results are presented in Table 3.

The addition of the C3-S2 block to the backbone of YOLO-LRD significantly improves general performance. This improvement is due to the module's capability to enhance focused region awareness and holistic image analysis, enabling the accurate detection of image features pertinent to substation equipment. This novel approach allows the backbone to handle feature data across different regions, extend the perceptive field depth, and maintain crucial parameter details.

Furthermore, incorporating SimAM into the neck of YOLO-LRD results in a significant boost in IoU precision. The SimAM module enhances the network's training ability by independently adjusting focus parameters for different input conditions, including complex object features, fine aspects of faraway objects, and prominent features of proximal objects. As a result, the neck network can effectively focus on capturing intricate features of substation equipment and ignore nonessential information. This improvement significantly enhances YOLO-LRD's general detection performance.

Table 3. Results of the ablation experiments.

C3-S2	SimAM	mAP@0.5 (%)	F1 Score (%)	IoU (%)
		83.7	87.6	89.0
✓		88.1	88.5	89.7
	✓	84.8	87.0	89.9
✓	✓	**90.3**	**90.8**	**92.5**

3.5. Comparison Experiment of GAN Models

To assess the image quality produced by the SEB-GAN algorithm, we conducted comparison experiments. Initially, we applied conventional image enhancement techniques such as rotation, scaling, and brightness adjustment to generate 2000 images of substation equipment blemishes. These enhanced images were then combined with the original blemish images to build a training dataset.

The other five comparison datasets use both reference methods and optimized approaches [19,21,23], to generate 2000 blemish images sequentially. These images are merged with the accumulated blemish images to form distinct datasets. Table 4 provides a comparison of the computational demands among the six GAN techniques.

It can be concluded that the Def-GAN model consumes a significant amount of computational resources, reflecting its superior intricacy compared with the other models. The SEB-GAN and Cycle-GAN algorithms show similar complexity levels. Due to the intricate backgrounds in substation device images, these algorithms often alter the general appearance of the images, resulting in greater distortion compared with authentic scenarios. This complicates image annotation and the training of the blemish detection algorithm.

Table 4. Comparison experiment results.

Image Augmentation Method	mAP/%	F1 Score/%	Params/MB	GPU Memory Usage/G	GFLOPs
Conventional techniques	74.9	84.4	-	-	-
Cycle-GAN	73.7	84.3	39.1	6.3	105
ResNet + Cycle-GAN	74.1	84.7	41.5	6.3	129
DCGAN	71.8	80.9	27.5	4.7	75
CONV + DCGAN	72.4	82.6	27.7	4.7	82
Def-GAN	76.9	86.5	**76.3**	**11.2**	**335**
SEB-GAN	**90.3**	**90.8**	43.9	8.0	146

We evaluated YOLOv7 object detection models trained on datasets generated by different GAN methods using authentic scenarios test datasets, detailed in Table 4. Compared with the current state-of-the-art model Def-GAN, our SEB-GAN model trained on a dataset of images it generated achieved a 13.4% improvement in mAP accuracy and a 4.3% increase in F1 score, while requiring fewer computational resources for training. The results show that SEB-GAN creates more varied images compared with conventional techniques. Unlike other GAN methods, SEB-GAN precisely manages blemish regions, effectively mitigating issues related to image authenticity. This enhancement in object detection performance highlights SEB-GAN's ability to produce blemish images with both intricate detail and enhanced authenticity.

3.6. Limitations of the Proposed Method

During the experiments, we found that using the YOLO-LRD model for substation equipment surface area detection and the SEB-GAN model for defect generation could generate blemishes of the corresponding categories on the equipment surface without altering the overall image style. However, this method still has the following limitations:

1. Compared with real defect images, the defect positions in the images generated by our method are not sufficiently realistic. For instance, oil stains typically occur at seals or other vulnerable joints of oil tanks prone to leakage, whereas our model may generate oil stains on the tank body surface where such stains are less likely to occur.
2. The connection between the generated blemishes and the background image still exhibits a noticeable boundary, lacking naturalness. For example, significant differences in brightness, saturation, and other aspects between the generated blemishes and the background image can make the generated image appear unnatural.
3. Our proposed method is capable of generating defect textures that are relatively simple, but it struggles to generate blemishes with complex textures such as bird nests, foreign objects, or intricate damage patterns like markings. The SEB-GAN method's performance in generating such complex defect features is inadequate.

4. Conclusions

In this paper, we propose a novel method for generating localized typical blemish images in substations. To address the issue of global style variation in GAN-generated images, we introduce a local region detection approach. By utilizing the YOLO-LRD method, we accurately pinpoint potential blemish locations in substation equipment images. YOLO-LRD features the C3-S2 block, a sophisticated unit for feature extraction that enhances both local detail and overall processing efficiency. This enhancement is vital for accurately detecting key features in substation images. Furthermore, integrating SimAM into the neck network boosts the model's learning capacity and facilitates adaptive attention weight adjustment for various input conditions. The introduction of the WD-CIoU fusion loss function also tackles issues related to size and shape variations in images of substation equipment. Collectively, these advancements demonstrate YOLO-LRD's notable progress in local region detection for substation imagery.

We then developed the SEB-GAN model specifically for generating blemish images in substations. SEB-GAN employs a joint discriminator that evaluates both the overall image and the defect image, thereby improving the quality of the generated defect images. This approach targets the creation of defect images in particular areas of substation equipment while maintaining the overall quality of the images. As a result, it prevents distortions that may arise from alterations in the global image style.

Our method excels in producing high-quality equipment blemish images with rich features and high realism. Experimental results show that the YOLOv7 defect detection model trained by the dataset generated by SEB-GAN, achieves an mAP of 90.3%, an F1 score of 90.8%, and an IoU of 92.5%. The results demonstrate that the SEB-GAN model surpasses performance of other conventional GAN-based methods in image generation

quality. Additionally, the training sets produced enhance the detection accuracy of leading object detection algorithms.

Future work may involve conducting application experiments on a wider range of substation equipment data, further validating the practical performance of YOLO-LRD and SEB-GAN, and making additional enhancements to the networks.

Author Contributions: Conceptualization, N.Z., J.F. and G.Y.; methodology, N.Z. and G.Y.; validation, G.L. and H.Y.; investigation, N.Z.; resources, Y.B. and H.Y.; data curation, G.Y.; writing—original draft preparation, N.Z.; writing—review and editing, N.Z.; supervision, J.F. and H.Y. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Melo, J.V.J.; Lira, G.R.S.; Costa, E.G.; Vilar, P.B.; Andrade, F.L.M.; Marotti, A.C.F.; Costa, A.I.; Leite Neto, A.F.; Santos Júnior, A.C.d. Separation and Classification of Partial Discharge Sources in Substations. *Energies* **2022**, *17*, 3804. [CrossRef]
2. Zhang, H.; Zhou, B.; Tian, Y.; Li, Z. Segmentation and Tracking Based on Equalized Memory Matching Network and Its Application in Electric Substation Inspection. *Algorithms* **2024**, *17*, 203. [CrossRef]
3. Gyrichidi, N.; Romanov, A.M.; Trofimov, O.V.; Eroshenko, S.A.; Matrenin, P.V.; Khalyasmaa, A.I. GNSS-Based Narrow-Angle UV Camera Targeting: Case Study of a Low-Cost MAD Robot. *Sensors* **2024**, *24*, 3494. [CrossRef] [PubMed]
4. de Castro, G.G.R.; Santos, T.M.B.; Andrade, F.A.A.; Lima, J.; Haddad, D.B.; Honório, L.d.M.; Pinto, M.F. Heterogeneous Multi-Robot Collaboration for Coverage Path Planning in Partially Known Dynamic Environments. *Machines* **2024**, *12*, 200. [CrossRef]
5. Lu, X.; Yuan, S.; Nian, Z.; Mu, C.; Li, X. Optimizing Task Offloading for Power Line Inspection in Smart Grid Networks with Edge Computing: A Game Theory Approach. *Information* **2024**, *15*, 441. [CrossRef]
6. Wu, Y.; Xiao, F.; Liu, F.; Sun, Y.; Deng, X.; Lin, L.; Zhu, C. A Visual Fault Detection Algorithm of Substation Equipment Based on Improved YOLOv5. *Appl. Sci.* **2023**, *13*, 11785. [CrossRef]
7. Wang, Q.; Yang, L.; Zhou, B.; Luan, Z.; Zhang, J. YOLO-SS-Large: A Lightweight and High-Performance Model for Defect Detection in Substations. *Sensors* **2023**, *23*, 8080. [CrossRef]
8. Thillaigovindhan, S.K.; Roslee, M.; Mitani, S.M.I.; Osman, A.F.; Ali, F.Z. A Comprehensive Survey on Machine Learning Methods for Handover Optimization in 5G Networks. *Electronics* **2024**, *13*, 3223. [CrossRef]
9. Wang, W.; Chen, J.; Han, G.; Shi, X.; Qian, G. Application of Object Detection Algorithms in Non-Destructive Testing of Pressure Equipment: A Review. *Sensors* **2024**, *24*, 5944. [CrossRef]
10. Girshick, R. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile, 7–13 December 2015.
11. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster r-cnn: Towards real-time object detection with region proposal networks. *Comput. Vis. Pattern Recognit. arXiv* **2015**, arXiv:1506.01497. [CrossRef]
12. He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017.
13. Hussain, M.; Khanam, R. In-Depth Review of YOLOv1 to YOLOv10 Variants for Enhanced Photovoltaic Defect Detection. *Solar* **2024**, *4*, 351–386. [CrossRef]
14. Zhang, N.; Yang, G.; Hu, F.; Yu, H.; Fan, J.; Xu, S. A Novel Adversarial Deep Learning Method for Substation Defect Image Generation. *Sensors* **2024**, *24*, 4512. [CrossRef] [PubMed]
15. Liu, Y.; Huang, X. Efficient Cross-Modality Insulator Augmentation for Multi-Domain Insulator Defect Detection in UAV Images. *Sensors* **2024**, *24*, 428. [CrossRef] [PubMed]
16. Huang, L.; Li, Y.; Wang, W.; He, Z. Enhanced Detection of Subway Insulator blemishes Based on Improved YOLOv5. *Appl. Sci.* **2023**, *13*, 13044. [CrossRef]
17. Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial networks. *Commun. ACM* **2020**, *63*, 139–144. [CrossRef]
18. Yang, J.; Wang, K.; Luan, F.; Yin, Y.; Zhang, H. PreCaCycleGAN: Perceptual Capsule Cyclic Generative Adversarial Network for Industrial Defective Sample Augmentation. *Electronics* **2023**, *12*, 3475. [CrossRef]

19. Di Maggio, L.G.; Brusa, E.; Delprete, C. Zero-Shot Generative AI for Rotating Machinery Fault Diagnosis: Synthesizing Highly Realistic Training Data via Cycle-Consistent Adversarial Networks. *Appl. Sci.* **2023**, *13*, 12458. [CrossRef]

20. Zhu, J.Y.; Park, T.; Isola, P.; Efros, A.A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2223–2232.

21. Liu, Y.; Gao, W.; Zhao, T.; Wang, Z.; Wang, Z. A Rapid Bridge Crack Detection Method Based on Deep Learning. *Appl. Sci.* **2023**, *13*, 9878. [CrossRef]

22. Radford, A.; Metz, L.; Chintala, S. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv* **2015**, arXiv:1511.06434.

23. Zhuang, C.; Li, J. Industrial Defect Detection of Plastic Labels Based on YOLOv5 and Generative Adversarial Networks. *Comput. Meas. Control.* **2023**, *31*, 91–98.

24. Zhang, G.; Cui, K.; Hung, T.Y.; Lu, S. Defect-GAN: High-Fidelity Defect Synthesis for Automated Defect Inspection. *arXiv* **2017**, arXiv:1709.03831.

25. Liu, R.; Liu, W.; Zheng, Z.; Wang, L.; Mao, L.; Qiu, Q. Anomaly-GAN: A data augmentation method for train surface anomaly detection. *Expert Syst. Appl.* **2023**, *228*, 120284. [CrossRef]

26. Wang, C.-Y.; Bochkovskiy, A.; Liao, H.-Y.M. Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *arXiv* **2022**, arXiv:2207.02696.

27. Jocher, G. YOLOV5 [EB/OL]. 2022. Available online: https://github.com/ultralytics/yolov5 (accessed on 5 April 2024).

28. Yang, L.; Zhang, R.-Y.; Li, L.; Xie, X. Simam: A Simple, Parameter-Free Attention Module for Convolutional Neural Networks. In Proceedings of the International Conference on Machine Learning, Online, 18–24 July 2021; pp. 11863–11874.

29. Wang, J.; Xu, C.; Yang, W.; Yu, L. A normalized Gaussian Wasserstein distance for tiny object detection. *arXiv* **2021**, arXiv:2110.13389.

30. Zheng, Z.; Wang, P.; Liu, W.; Li, J.; Ye, R.; Ren, D. Distance-IoU loss: Faster and better learning for bounding box regression. In Proceedings of the AAAI Conference on Artificial Intelligence, New York, NY, USA, 7–12 February 2020; Volume 34, pp. 12993–13000.

31. Ye, R.; Shao, G.; Yang, Z.; Sun, Y.; Gao, Q.; Li, T. Detection Model of Tea Disease Severity under Low Light Intensity Based on YOLOv8 and EnlightenGAN. *Plants* **2024**, *13*, 1377. [CrossRef]

32. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In *Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015, Munich, Germany, 5–9 October 2015* ; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2015; pp. 234–241.