Low-Light Image Enhancement Method Based on Optimized Generative Adversarial Network

Zhaokang Ji
School of Computer and Information Engineering
Anhui Polytechnic University
Wuhu, China
zhaokangji@mail.ahpu.edu.cn

Tao Liu*
School of Computer and Information Engineering
Anhui Polytechnic University
Wuhu, China
zhaokangji@mail.ahpu.edu.cn
*Corresponding Author

Ping Zhang
School of Computer and Information Engineering
Anhui Polytechnic University
Wuhu, China
zhaokangji@mail.ahpu.edu.cn

2024 IEEE 4th International Conference on Data Science and Computer Application (ICDSCA) | 979-8-3503-6823-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICDSCA63855.2024.10859773

***Abstract*—This paper proposes a low-light image brightness restoration method based on an improved GAN network. The approach leverages CycleGAN, which does not require paired datasets, simplifying implementation. PatchGAN is used as the discriminator to capture fine-grained image details and accelerate network convergence. A brightness equalization loss function is introduced to ensure balanced brightness in the enhanced images, while a CBAM (Convolutional Block Attention Module) is incorporated to improve feature extraction and pooling effects. Experimental results demonstrate the effectiveness of the proposed algorithm, achieving superior image enhancement in terms of both visual quality and objective metrics. Future work will explore integrating denoising modules to further enhance image quality.**

Keywords—Low-Light Image Enhancement, CycleGAN, PatchGAN, Brightness Equalization, CBAM, Image Processing, GAN Network

I. Introduction

Image acquisition can be achieved through various methods, including camera photography, webcam capture, and thermal imaging. However, due to the limitations of lighting conditions, the resulting images often suffer from issues such as low resolution, poor contrast, and low signal-to-noise ratio [1]. These shortcomings reduce the usability of the images, making them difficult to recognize or analyze. In particular, low-quality images present significant challenges for extracting meaningful features, which in turn hampers further analysis. This makes image enhancement a crucial task [2].

Over the years, image enhancement has remained a key area of research in the field of image processing. By enhancing images, more detailed information can be extracted, which facilitates subsequent tasks such as semantic segmentation and object detection. One of the earliest methods for image enhancement involved adjusting the distribution of the image histogram, known as Histogram Equalization (HE) [3]. This approach is simple and efficient, making it widely used. Subsequently, Land and colleagues introduced the Retinex theory, which posits that human perception of an object's brightness and color is determined by the object's inherent reflectance rather than the intensity of ambient illumination. Building on this concept, researchers developed various Retinex-based algorithms, such as Path-Based Retinex and Center-Surround Retinex. Notable examples include Single Scale Retinex (SSR), Multi-Scale Retinex (MSR), and Multi-Scale Retinex with Color Restoration (MSRCR) [4]. These methods effectively enhance images captured under low-light conditions while preserving color information. However, they often lead to issues such as blurred edges and amplified noise [5].

Despite Retinex theory having been proposed decades ago, efforts to improve its applications continue unabated. With the rapid advancement of deep learning, training models on large datasets has demonstrated exceptional potential for image enhancement. Convolutional Neural Networks (CNNs), which have achieved remarkable success in areas such as object tracking and image classification, are increasingly being employed for image enhancement [6]. For example, Lv et al. proposed a Multi-Branch Low-Light Enhancement Network (MBLLEN) that employs multiple subnetworks to enhance various features. These features are then fused to generate the final output image. Similarly, Tao et al. introduced LLCNN, an end-to-end image enhancement network [7].

Most of these deep learning-based methods rely on paired datasets for training. However, in real-world scenarios, obtaining paired datasets is often challenging, and the difficulty in constructing such datasets significantly impacts the quality of image enhancement. In recent years, Generative Adversarial Networks (GANs) have emerged as a powerful tool for image generation and have been widely adopted for image enhancement. CycleGAN, a type of GAN architecture designed for unpaired learning, enables image transformation between different domains without requiring paired datasets [8].

In light of these advancements, this work proposes an improved GAN-based model for image enhancement. By leveraging the strengths of CycleGAN and integrating novel modifications, the proposed model aims to achieve superior performance in enhancing low-quality images, addressing issues such as edge preservation, noise suppression, and color fidelity [9].

II. GAN Network Model Overview

A. Network Model

Generative Adversarial Networks (GANs) consist of two primary components: a generator (G) and a discriminator (D). These components learn through an adversarial process. The generator, utilizing machine learning, attempts to produce synthetic samples that closely resemble real samples, while the discriminator works to distinguish between real and generated images. During this adversarial training, the discriminator continually guides the generator to produce increasingly realistic outputs. The process eventually reaches a dynamic equilibrium where the generator achieves optimal performance [10].

In this experiment, the generator is primarily implemented using the CycleGAN framework. CycleGAN combines two one-way GANs: one for transforming domain X→Y and the other for Y→X. These two GANs share two generators G and F and each has an associated discriminator. By designing two symmetrical network structures, CycleGAN addresses the challenge of requiring paired datasets for image transformation algorithms. Using the two generators G and F and their corresponding discriminators, CycleGAN maps between original images and generated images, enabling unpaired images from different domains to be transformed into each other. The schematic structure of CycleGAN is shown in Figure 1.

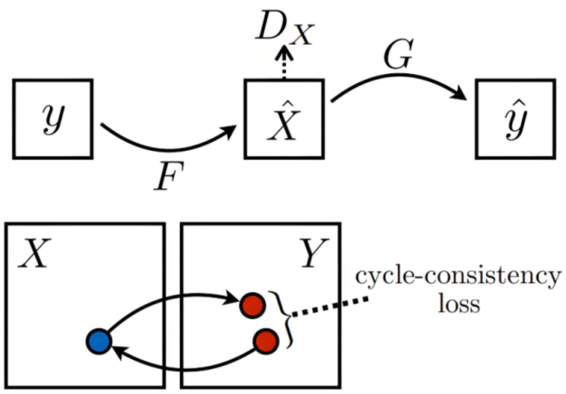

Figure 1. CycleGAN Structural Schematic Diagram

The generative network undergoes an encoding and decoding process. It primarily consists of 8 convolutional blocks, which form multi-scale convolutional modules using kernels of sizes 1×1, 3×3, 5×5, and 7×7, with a stride of 2. The encoder employs a downsampling method to compress and extract feature information from low-light images. Through three downsampling operations, it compresses the input feature map from 512×384 to 64×48. Each downsampling module consists of a 3×3 convolutional kernel followed by an instance normalization layer and the LeakyReLU activation function, with the number of channels set to 64. The decoder utilizes upsampling to gradually restore the shallow features of the image [11]. The upsampling layers use bilinear interpolation to further restore the shallow features of the generated image. Each upsampling layer consists of a deconvolution module, the LeakyReLU activation function, and an instance normalization layer, aiming to reconstruct the generated image to the same size as the input image. Finally, the decoding module applies a 7×7 convolution and removes the instance normalization layer, replacing it with the tanh activation function, which effectively addresses symmetry issues around the origin and ensures faster convergence [12].

B. Objective Function

CycleGAN utilizes unpaired datasets for training to achieve the effects of training with paired datasets. As a type of GAN framework, its objective loss primarily consists of adversarial loss, cycleconsistency loss, and identity mapping loss. The adversarial loss ensures that the generated images are as realistic as possible, while the cycle-consistency loss ensures that the transformations between domains are consistent [13].

Assuming there is a low-light domainXand a normal-light domainY, with a mapping relationship between them, the generator$G: X \rightarrow Y$ and $F: Y \rightarrow X$are used for transformations.

The adversarial loss measures how well the discriminator distinguishes between real and generated images, encouraging the generator to produce increasingly realistic results. The adversarial loss for the discriminator is defined as:

$$L_{GAN}(G, D_Y, X, Y) = E_{x\sim p_{data}(x)}\left[\ln\left(1 - D_Y\left(G(x)\right)\right)\right] + E_{y\sim p_{data}(y)}[\ln D_Y(y)] \quad (1)$$

Through the two generators, the model enables transformations between different lighting domains [14]. However, adversarial loss alone cannot ensure the output remains consistent with the input after mapping. To further constrain the mapping space, CycleGAN introduces the cycle-consistency loss, ensuring that the mappingsGandFare inverses of each other:

$$L_{cyc}(G, F) = E_{x\sim p_{data}(x)}\left[\| F\left(G(x)\right) - x \|_1\right] + \left[\| G\left(F(y)\right) - y \|_1\right] \quad (2)$$

To make the generated images as close as possible to real images, CycleGAN also includes the identity mapping loss:

$$L_{id}(G, F) = E_{x\sim p_{data}(x)}[\| F(x) - x \|_1] + E_{y\sim p_{data}(y)}[\| G(y) - y \|_1] \quad (3)$$

Finally, the overall transfer loss for the generative adversarial network is defined as:

$$L_{transfer}(G, F, D_X, D_Y) = L_{GAN} + \lambda_1 L_{cyc} + \lambda_2 L_{id} \quad (4)$$

In the equation,$\lambda_1$and$\lambda_2$are weighting coefficients for the cycle-consistency loss and identity mapping loss, respectively, with$\lambda_1 = 10$and$\lambda_2 = 5$.

III. IMPROVED GAN NETWORK MODEL

A. Discriminator Selection

The GAN network utilizes game theory to learn the discrepancies between real and generated images. It consists of two independent networks: the generator and the discriminator, working collaboratively. The generator's task is to produce realistic images, while the discriminator's task is to classify real and generated images, determining whether the generator's outputs align with the data distribution. Therefore, selecting an appropriate discriminator is crucial.

Traditional discriminators are often binary classifiers that focus on overall features but fail to capture fine-grained details in images. As a result, they perform poorly in tasks requiring high levels of detail.

For this experiment, the discriminator model references the PatchGAN discriminator proposed by Isola et al. [15]. PatchGAN is a fully convolutional network whose output is an N×N matrix. The matrix's mean value is calculated to determine whether an image is real or generated. Using PatchGAN as the discriminator has several advantages: The mean value of the matrix fully considers the influence of different regions of the image, better representing the likelihood that the receptive field [16] is real. It accelerates the convergence of small-scale feature network patches, making the network easier to train and reducing training time. This paper designs a discriminator model with a 5-layer fully convolutional network, where the convolutional kernel size is 4×4. The network structure is illustrated in Figure 2. The first four convolutional blocks include the LeakyReLU activation function, with a stride of 2. The final convolutional block contains a convolutional layer and an activation function with a stride of 1. Batch normalization (BatchNorm) is applied to the middle three layers to normalize the data and prevent overfitting. The final layer uses the Sigmoid function as the activation function, with an output range of (0, 1). This prevents the network from diverging during training.

In binary classification tasks, the Sigmoid function is commonly used to map model outputs to a range between 0 and 1, representing the probability that the sample belongs to the positive class. For this experiment, the receptive field of PatchGAN was set by configuring the input image patches to 70×70. This ensures that PatchGAN is lighter and faster than traditional discriminators without compromising its ability to guide the generator to produce more realistic results.

Finally, the discriminator outputs a matrix with a single channel, where each output pixel represents the probability that the receptive field region of the input image is real or fake. This is used to determine the authenticity of the generated image.

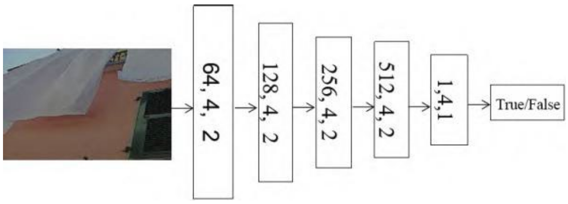

Figure 2. Discriminator Network Architecture

B. Loss Function

Although CycleGAN has proven effective in achieving style transfer between different image domains, it comes with significant drawbacks. These issues primarily arise due to the flexibility of GAN networks, which can occasionally yield unstable results. For instance, generated images may suffer from uneven brightness, color distortions, or halo effects that undermine their visual quality. The loss functions commonly employed in GANs—such as adversarial loss, cycle consistency loss, and identity mapping loss—are not sufficient by themselves to address these problems, particularly when it comes to capturing crucial low-level image features, such as lighting consistency and detail preservation.

One of the primary challenges in image generation, especially for low-light or high-contrast images, is the accurate representation of brightness across different image regions. Without additional constraints, it is difficult for the generator networks G and F to ensure that the brightness is uniformly distributed throughout the image. In particular, overly dark regions, which are typically characterized by a lack of detailed information, often become problematic in generating realistic images. To address these challenges and improve the quality of the generated images, we introduce a brightness equalization loss function designed to ensure that the brightness of the generated images closely matches that of the real images, reducing the issue of uneven brightness.

The goal of the brightness equalization loss is to enhance the uniformity of overall image brightness, ensuring that local regions of the image are consistent with the global brightness distribution. By incorporating this loss function, we provide an additional constraint that helps guide the generator to produce images with more consistent brightness and reduced halo effects, which are often seen in low-light image enhancement tasks.

To implement this, the loss function is defined as follows:

Calculate the Overall Average Intensity: The first step is to compute the overall average intensity I of the image. This can be expressed as the mean pixel value across the entire image:

$$\mathrm{I} = \frac{1}{\mathrm{H \times W}} \sum_{\mathrm{h=1}}^{\mathrm{H}} \sum_{\mathrm{w=1}}^{\mathrm{W}} \mathrm{I}\,(\mathrm{h,w}) \quad (5)$$

Where h and w are the height and width of the image, respectively, and $\mathrm{I(h,w)}$ is the intensity of the pixel at position $\mathrm{(h,w)}$.

Divide the Image into Blocks: Next, the image is divided into small blocks of size $\mathrm{n \times m}$ (where n and m are the height and width of each block). For each block, we compute the average intensity:

$$\mathrm{D_{i,j}} = \frac{1}{\mathrm{n \times m}} \sum_{\mathrm{x=1}}^{\mathrm{n}} \sum_{\mathrm{y=1}}^{\mathrm{m}} \mathrm{I}\,(\mathrm{x,y}) \quad (6)$$

Where $\mathrm{D_{i,j}}$ represents the average intensity of the block located at position $\mathrm{(i,j)}$.

Compute the illumination Difference: To ensure that the local regions' brightness matches the global brightness, we subtract the overall average intensity I from each element in the illumination matrix D to form the illumination difference matrix E:

$$\mathrm{E_{i,j}} = \mathrm{D_{i,j}} - \mathrm{I} \quad (7)$$

This matrix E reflects the difference between the local brightness and the overall image brightness. If the difference is large, the image likely has regions that are too dark or too bright compared to the global average.

C. Adding Attention Mechanism

To further improve image quality, this paper introduces a Convolutional Block Attention Module (CBAM) [12] into the backbone feature network. CBAM is a novel approach that enhances the network's representational ability, and its structure is shown in Figure 3.

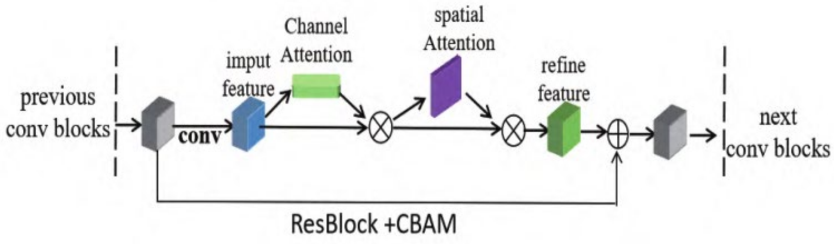

Figure 3. CBAM Structural Diagram

CBAM consists of two main components. The first is a channel attention module, which aims to produce better pooling effects. The second is a spatial attention module, designed to further enhance network performance. By integrating this module into the backbone feature extraction network, CBAM applies global average pooling to the channels, enabling the network to perform both global average and global max pooling simultaneously. This allows the network to generate better pooling effects and recover more image detail information.

IV. EXPERIMENT PROCESS AND RESULTS

A. Experimental Environment and Dataset Description

The experiments in this paper were conducted on a PC with the following hardware configuration: GPU (Geforce 940MX), and the operating system is Windows 10 Home Edition. The software setup includes Anaconda and PyCharm, and the programming language used is Python with the deep learning framework Torch.

The dataset used in this study is a combination of images from the Raise dataset and the LOL dataset, which includes a total of 434 low-light images and 656 normal-light mixed images. These were assembled into an unpaired training set. For testing, 148 low-light images were selected and adjusted to a uniform format of 512×384 pixels. The batch size for the network was set to 2, and the Adam optimizer was used. The model was trained for a total of 100 epochs with a learning rate of 0.0001.

B. Experimental Results Analysis

To verify the superiority of the improved GAN network for low-light image enhancement, comparison experiments were conducted using Zero-DCE[13] and Enlighten GAN[14]. The evaluation was carried out from both subjective and objective perspectives.

Subjective Analysis: This part of the evaluation mainly focuses on the overall enhancement effect of the image. It includes analysis of aspects such as color, contrast, and noise. Objective Analysis: For objective evaluation, the image quality metrics PSNR (Peak Signal-to-Noise Ratio) and SSIM (Structural Similarity Index) were used to compare the results.

The experiment used a set of images from a complex scene containing many buildings and the sky. The input images and their enhanced results are shown in Figure 4. It can be seen that all three methods—Enlighten GAN, Zero-DCE, and the proposed algorithm—can achieve image enhancement. However, there are noticeable differences in performance:

Zero-DCE: The enhanced images have relatively dark colors overall, and visual artifacts such as noise and color distortion are present. Enlighten GAN: This method results in a significant increase in overall image brightness. However, as the image is enlarged, the details become blurred, and noise amplification is evident. Proposed Algorithm: Although the brightness enhancement in the proposed method is not as prominent as in Enlighten GAN, the brightness is more uniform. The enhanced images appear closer to real-world scenes, with more consistent brightness and finer details. In summary, while all methods achieve some level of enhancement, the proposed algorithm shows a better balance of brightness uniformity and detail preservation, making it more suitable for practical applications in low-light image enhancement.

(a)Input (b)Enlighten CA (c)Zero-DCE (d)ours

Figure 4. Image Enhancement Results Display

To further demonstrate the superiority of the low-light image enhancement algorithm proposed in this paper, the PSNR[15] and SSIM[16] image quality evaluation metrics for various algorithms in the experiment are provided in Figure 4. As shown in Table I, the proposed algorithm achieves the highest values in both SSIM and PSNR (higher values of these metrics indicate better image quality). Compared to Enlighten GAN, the proposed algorithm improves PSNR and SSIM by 2.257 dB and 0.055, respectively. This indicates that the images enhanced by the proposed algorithm are less affected by noise, exhibit less distortion, and have higher image recovery quality, making the enhanced images closer to the real images.

TABLE I. EXPERIMENTAL OBJECTIVE EVALUATION METRICS

Method	PSNR/dB	SSIM
Enlighten GAN	21.202	0.742
Zero-DCE	20.949	0.709
Ours	23.459	0.797

To verify the superiority of the improved GAN network for low-light image enhancement, comparison experiments were conducted using Zero-DCE and Enlighten GAN. The evaluation was carried out from both subjective and objective perspectives.

Subjective Evaluation: This part of the evaluation mainly focuses on the overall enhancement effect of the image, including analysis of aspects such as color, contrast, and noise. It was observed that: Zero-DCE: The enhanced images exhibit relatively dark overall colors, with visual artifacts such as noise and color distortion present, particularly in darker areas. Enlighten GAN: This method significantly increases overall brightness, but the image often becomes blurred when enlarged, and noise amplification becomes evident. Proposed Algorithm: The brightness enhancement achieved by the proposed method is more uniform compared to Enlighten GAN. The images appear closer to real-world scenes, with consistent brightness and fine details well-preserved.

Objective Evaluation: The quantitative results, as shown in Table I, demonstrate that the proposed algorithm achieves the highest PSNR and SSIM values, indicating better image quality and fewer distortions. Specifically, the proposed method improves PSNR by 2.257 dB and SSIM by 0.055 compared to Enlighten GAN.

In addition to the direct comparison between methods, further investigation into the behavior of the proposed algorithm under different low-light conditions was conducted. The results indicate that the proposed method performs consistently well across a variety of complex scenes, including images with mixed light sources (e.g., street lighting and natural light), and it effectively addresses the challenge of enhancing low-light images without introducing significant artifacts.

Furthermore, the brightness equalization loss plays a crucial role in achieving uniform lighting across different regions of the image. This feature helps the algorithm avoid the common issue of over-enhanced bright regions while preserving the details in darker areas, resulting in more natural-looking and visually balanced images. The proposed method demonstrates a strong potential for real-world applications, particularly in scenarios such as autonomous driving, night-time photography, and low-light surveillance, where both clarity and brightness uniformity are essential.

V. Conclusion

This paper proposes a method for brightness restoration of low-light images based on an improved GAN network. Enhancing images under low-light conditions not only improves their visual quality but also facilitates better analysis and research of the images. Since CycleGAN does not require paired datasets, it makes the network easier to implement. PatchGAN is selected as the discriminator, as it can identify more image details and accelerate the convergence speed of the network. Additionally, a brightness equalization loss function is introduced to constrain the network, enabling it to generate images with more balanced brightness. Moreover, the CBAM attention mechanism module is integrated into the feature extraction module, allowing the network to focus on more detailed image features during feature extraction and achieve better pooling results. Overall, experimental results confirm the feasibility of the proposed algorithm, which achieves notable improvements in both visual and objective evaluations. The enhanced model significantly improves the effect of image enhancement. However, there is still room for further improvement in the experiments. In the future, incorporating an image denoising module into the network will be considered to improve the quality of the generated images.

References

[1] J. Duan, M. Gao, G. Zhao, et al., "DStokes-CGCP: A low-light color polarization image enhancement method combining chroma spectrum and global contour awareness," Optics and Lasers in Engineering, vol. 185, p. 108712, 2025.

[2] Z. Li, X. Li, J. Shi, et al., "Perceptually-calibrated synergy network for night-time image quality assessment with enhancement booster and knowledge cross-sharing," Displays, vol. 86, p. 102877, 2024.

[3] L. Li, W. Xu, Y. Gao, et al., "Attention-oriented residual block for real-time low-light image enhancement in smart ports," Computers and Electrical Engineering, vol. 120, p. 109634, 2024.

[4] L. Jada, R. Srikanth, K. Bikshalu, "Effective low-exposure color image enhancement based on histogram equalization with spatial contextual information," Engineering Research Express, vol. 6, no. 4, p. 045236, 2024.

[5] Y. Gao, Z. Xu, X. Xu, "Enhanced cigarette pack counting via image enhancement techniques and advanced SAFECount methodology," Traitement du Signal, vol. 40, no. 6, 2023.

[6] H. Wang, C. A. Frery, M. Li, et al., "Underwater image enhancement via histogram similarity-oriented color compensation complemented by multiple attribute adjustment," Intelligent Marine Technology and Systems, vol. 1, no. 1, p. 12, 2023.

[7] F. Lv, F. Lu, J. Wu, et al., "MBLLEN: Low-light image/video enhancement using CNNs," in Proc. British Machine Vision Conference (BMVC), Newcastle, UK, 2018, p. 220.

[8] L. Tao, C. Zhu, G. Xiang, et al., "LLCNN: A convolutional neural network for low-light image enhancement," in Proc. IEEE Visual Communications and Image Processing Conference (VCIP), Piscataway, NJ, USA, 2023, pp. 1–4.

[9] I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., "Generative adversarial nets," [Online]. Available: https://arxiv.org/abs/1406.2661, June 2024.

[10] U. Demir and G. Unal, "Patch-based image inpainting with generative adversarial networks," [Online]. Available: http://arxiv.org/abs/1803.07422, Mar. 2018.

[11] W. Luo, Y. Li, R. Urtasun, et al., "Understanding the effective receptive field in deep convolutional neural networks," in Advances in Neural Information Processing Systems (NeurIPS), vol. 8, pp. 4898–4906, 2016.

[12] J. B. Dai, Research on Image Feature Extraction Algorithm Based on Visual Information, M.S. thesis, Jilin University, Changchun, China, 2023.

[13] Y. Jiang, X. Gong, D. Liu, et al., "EnlightenGAN: Deep light enhancement without paired supervision," IEEE Transactions on Image Processing, vol. 30, pp. 2340–2349, 2021.

[14] C. Guo, C. Li, J. Guo, et al., "Zero-reference deep curve estimation for low-light image enhancement," in Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Piscataway, NJ, USA, 2020, pp. 1780–1789.

[15] X. Huang, J. Shi, J. Yang, et al., "Research on color image quality evaluation based on color difference MSE and PSNR," Acta Photonica Sinica, no. S1, pp. 295–298, 2022.

[16] L. Zhu, Q. Li, T. Zhang, et al., "Image quality evaluation method based on structural similarity," Opto-Electronic Engineering, vol. 34, no. 11, pp. 108–113, 2024.