Received: 13 January 2025 | Accepted: 11 June 2025

DOI: 10.1111/mice.70001

COMPUTER-AIDED CIVIL AND INFRASTRUCTURE ENGINEERING WILEY

RESEARCH ARTICLE

Advanced low-light image transformation for accurate nighttime pavement distress detection

Yuanyuan Hu[1] | **Hancheng Zhang**[1] | **Yue Hou**[2] | **Pengfei Liu**[1]

[1]Institute of Highway Engineering, RWTH Aachen University, Aachen, Germany

[2]Department of Civil Engineering, Faculty of Science and Engineering, Swansea University, Swansea, UK

Correspondence
Pengfei Liu, Institute of Highway Engineering, RWTH Aachen University, Aachen, Germany.
Email: liu@isac.rwth-aachen.de

Abstract

Pavement distress detection is critical for road safety and infrastructure longevity. Although nighttime inspections offer advantages such as reduced traffic and enhanced operational efficiency, challenges like low visibility and noise hinder their effectiveness. This paper presents IllumiShiftNet, a novel model that transforms low-light images into high-quality, daylight-like representations for pavement distress detection. By employing unpaired image translation techniques, aligned nighttime–daytime datasets are generated for supervised training. The model integrates a lightEnhance generator, multiscale feature discriminators, and distress-focused loss function, ensuring accurate reconstruction of critical pavement details. Experimental results show that IllumiShiftNet achieves a state-of-the-art peak signal-to-noise ratio of 28.5 and a structural similarity index measure of 0.78, enabling detection algorithms trained on daytime data to perform effectively on nighttime imagery. The model demonstrates robust performance across varying illuminance levels, adverse weather conditions, and diverse road types while maintaining real-time processing capabilities. These results establish IllumiShiftNet as a practical solution for nighttime pavement monitoring.

1 | INTRODUCTION

The identification of pavement distress is a critical aspect of maintaining the safety and longevity of road infrastructure, which is fundamental to modern transportation systems. Early detection of pavement deterioration is essential for implementing timely maintenance, as unaddressed surface damage can escalate into more severe conditions, leading to higher repair costs, traffic disruptions, and safety risks (Tong et al., 2020). Effective pavement distress detection enables the development of maintenance strategies aimed at mitigating deterioration and extending the service life of roadways (Malekloo et al., 2024). In recent years, significant advancements have been made in automated detection techniques, ranging from traditional image processing algorithms to advanced deep learning frameworks (Rafiei & Adeli, 2017). These methods have enhanced the accuracy and efficiency of pavement condition assessment (Pan et al., 2023).

Most existing approaches for pavement distress detection are designed for daytime operations, utilizing the favorable lighting conditions that facilitate high-resolution imagery and clearer pavement visibility (Zhu et al., 2024). However, daytime inspections often face significant challenges. High traffic volumes during the day increase the likelihood of obstructed camera views and pose safety risks to both inspection personnel and road users (J. Li et al., 2021). Heavy traffic during the day can also obstruct the inspection equipment's view of the pavement, limiting the ability to capture complete and consistent data for analysis. In contrast, nighttime detection offers several operational advantages that address these limitations. Roads generally experience lower traffic volumes at night, resulting in fewer potential conflicts between inspection vehicles and regular traffic, thus creating a safer working environment for inspection crews (Montenegro & Flores-Calero, 2022). The reduced traffic also allows for smoother and higher speed data acquisition, which can significantly enhance the efficiency of the inspection process. Moreover, the absence of harsh sunlight eliminates issues related to glare and shadowing, which are common during daytime operations, enabling more consistent imaging conditions. Beyond these technical benefits, nighttime inspections provide transportation agencies with greater scheduling flexibility, allowing for optimization of limited resources and equipment usage. These factors collectively highlight the potential of nighttime detection as a safer, faster, and more efficient alternative under appropriate conditions.

Despite its advantages, nighttime pavement distress detection presents significant technical challenges, primarily due to the inherent limitations imposed by low ambient light. Insufficient illumination significantly reduces the signal-to-noise ratio (SNR) in captured images, diminishing the quantity of meaningful information available for analysis (Zhou et al., 2025). From the perspective of Shannon's information theory, the information capacity of a communication channel—such as an imaging sensor—is directly proportional to its SNR and bandwidth. Low-light conditions effectively constrain this capacity, reducing the amount of recoverable information and impairing the ability to distinguish subtle pavement distress features. Conventional compensatory approaches, including extended exposure times and increased sensor gain, often introduce secondary problems like motion blur and noise amplification, which can further mask critical pavement details (Gu et al., 2020). The technical challenge extends beyond simple image brightness—noise patterns can mimic or obscure actual pavement texture variations, leading to both false positives and missed detections in distress identification algorithms. While specialized hardware solutions provide partial remedies by enhancing signal acquisition, they typically involve significant cost and operational complexity. On the software side, image processing techniques for denoising and contrast enhancement often struggle with the fundamental trade-off between noise suppression and detail preservation, frequently resulting in over-smoothed images where subtle but important pavement distress indicators are lost (Chen et al., 2023). Furthermore, the practical implementation of pavement monitoring systems increasingly demands real-time processing capabilities across hardware platforms with varying computational power. Effective solutions must maintain operational efficiency across different GPUs and processing units without sacrificing enhancement quality (Y. Liu et al., 2025). The challenge extends beyond merely enhancing low-light images—algorithms must balance complicated image transformation with consistent processing speeds on different computational devices, from high-performance workstations to more constrained systems. This combination of challenges creates a high-entropy imaging environment where uncertainty dominates, substantially limiting the reliability and precision of automated distress detection systems operating under nighttime conditions without specialized adaptation.

To address the challenges of nighttime pavement distress detection, researchers have explored three primary strategies: (1) traditional image preprocessing techniques, (2) training detection models on nighttime datasets, and (3)transforming nighttime images into daylight-like representations using image translation methods. Traditional image preprocessing includes methods such as brightness adjustment, contrast enhancement, and noise reduction (Maryami et al., 2024). These techniques aim to improve the visual clarity of nighttime images and are computationally efficient and easy to implement. However, they primarily enhance the appearance of the images without recovering or increasing the underlying information. Since no new data are introduced, these methods are limited in their ability to address the inherent noise and signal loss caused by low-light conditions (Rafiei et al., 2022). As a result, they often fail to manage significant variations in lighting, leading to over-enhancement or loss of critical pavement details. Training detection models directly on nighttime datasets (Miao et al., 2020) is another approach aimed at addressing the challenges of low-light conditions. This strategy allows the models to learn and adapt to the unique characteristics of nighttime imagery, potentially improving detection accuracy under such conditions. However, the collection and annotation of sufficiently large and diverse nighttime datasets are labor-intensive and costly (Rafiei et al., 2024). Additionally, despite extensive training, these models often struggle to overcome the inherent quality degradation of nighttime images caused by noise, low signal levels, and insufficient contrast. These limitations reduce the scalability and generalizability of the models, making it challenging to apply them effectively across diverse real-world scenarios (Alam et al., 2020).

A promising strategy for nighttime pavement distress detection involves the use of image translation techniques (Y. Zhang & Zhang, 2024) to convert nighttime images into daylight-like representations. This approach normalizes lighting conditions, enabling the use of robust detection models that were originally developed and optimized for daytime imagery. By employing the extensive advancements in daytime pavement detection algorithms, this method has the potential to enhance detection performance without requiring entirely new model architectures tailored to low-light conditions. Despite its potential, this approach faces significant challenges (Pereira et al., 2020). One major obstacle is the scarcity of strictly aligned nighttime–daytime datasets. Advanced image translation models typically require paired datasets where each nighttime image corresponds to a daytime image captured under identical or closely matched conditions, ensuring consistency in perspective and content. However, collecting such datasets is time-intensive and logistically challenging, often compelling researchers to rely on approximate or unpaired datasets, which compromises translation accuracy and generalizability. Another key challenge lies in the inherent difficulty of translating nighttime images into realistic and detailed daylight representations. Nighttime images are characterized by weaker signals, increased noise, and reduced contrast, making it difficult to recover subtle but critical details, such as pavement cracks. Conventional image enhancement algorithms and standard machine learning approaches often fall short in restoring fine features and preserving structural integrity during the translation process. As a result, visual artifacts and degraded image quality can occur, particularly in regions of critical detail, negatively affecting downstream detection tasks. These challenges highlight the need for innovative solutions that not only address dataset limitations but also improve the robustness and fidelity of nighttime-to-daytime image translations.

To address these challenges, this paper introduces IllumiShiftNet, a novel model for transforming nighttime pavement images into realistic daylight representations. The approach utilizes an unpaired image translation model to generate paired training datasets, combining multiscale feature discriminators with a distress-focused loss function to ensure both global coherence and precise reconstruction of pavement distress regions. This design enables detection models trained solely on daytime data to effectively process nighttime imagery, demonstrating robust adaptability across varying illuminance levels, adverse weather conditions, and diverse road types. IllumiShiftNet outperforms state-of-the-art methods with superior image quality metrics while maintaining real-time processing capabilities. The primary contributions of this paper are as follows:

1. Proposing IllumiShiftNet, a novel model that transforms nighttime pavement images into daylight-like representations for accurate distress detection under low-light conditions.
2. Utilizing existing unpaired image translation models to generate strictly aligned paired datasets, creating high-quality training data without extensive manual collection.
3. Introducing a distress-focused loss function that prioritizes critical crack details, improving downstream detection accuracy.
4. Enabling detection models trained on daytime data to achieve high performance on translated nighttime images without requiring nighttime-specific training.
5. Demonstrating robust adaptability across varying illuminance levels and adverse weather conditions, while maintaining consistent performance across diverse pavement types.
6. Achieving better performance compared to other methods with the highest peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM) scores, while meeting real-time requirements with optimized inference speeds on mainstream GPU platforms.

2 | RELATED WORKS

2.1 | Traditional image preprocessing

Traditional image preprocessing techniques for nighttime pavement distress detection include illumination adjustment, contrast enhancement, and noise reduction methods. These approaches aim to improve the visual clarity of low-light images through direct pixel manipulation.

Illumination adjustment methods correct insufficient brightness in nighttime images, utilizing techniques such as histogram equalization prior (F. Zhang et al., 2023), gamma correction, and multiscale Retinex algorithms (Zotin, 2020) to enhance visibility in darker regions. Contrast enhancement techniques increase the distinction between pavement features and their surroundings through unified frameworks integrating denoising (L. Li et al., 2015), wavelet-based enhancement, and contrast-limited adaptive histogram equalization approaches (Yuan et al., 2023). Noise reduction algorithms mitigate increased noise in low-light imagery while preserving structural details. Key approaches include bilateral filtering (Gavaskar & Chaudhury, 2018), nonlocal means denoising, and hybrid optimization methods (Rajinikanth & Razmjooy, 2023).

Despite computational efficiency, traditional preprocessing methods have inherent limitations. They often apply global adjustments that inadequately address complex illumination–reflectance relationships in nighttime scenes. Without a semantic understanding of pavement features, these methods struggle to selectively enhance critical regions. The fundamental trade-off between noise reduction and detail preservation frequently results in either excessive noise or loss of fine crack information.

2.2 | Unpaired data-based image translation

In scenarios where strictly aligned nighttime–daytime image pairs are unavailable, generative models trained on unpaired datasets have emerged as a practical alternative. These approaches aim to learn mappings between two domains without the need for perfectly matched samples.

Several notable methods have been developed in this area. An unsupervised low-light image enhancement model based on an improved CycleGAN(Cycle Generative Adversarial Network) framework (Tang et al., 2024) utilizes an enhanced U-Net structure, with adaptive instance normalization (AdaIN) to learn the style of normal-light images. Cycle-Retinex (Wu et al., 2023) combines Retinex theory with an inline CycleGAN framework, decomposing the low-light enhancement task into two sub-tasks: illumination enhancement and reflectance restoration. Undark-GAN(Undark Generative Adversarial Network) focuses on enhancing dark regions without overexposing or oversaturating well-lit areas, effectively preserving structural details (Parihar et al., 2021).

Unpaired data-based approaches excel in style transfer tasks such as converting daytime to nighttime images, where information is primarily discarded rather than recovered. However, translating nighttime images into realistic daytime representations is significantly more complex, requiring recovery of critical information that is inherently absent or degraded in low-light conditions. These methods show substantial limitations in preserving fine crack details and subtle pavement texture variations, particularly under weak lighting where minimal structural information is available. The absence of direct supervision between unpaired images restricts the accurate reconstruction of fine-grained features crucial for pavement distress detection, resulting in translations that may appear visually plausible but lack the precise details necessary for reliable analysis.

2.3 | Paired data-based image translation

In contrast to unpaired methods, paired data-based solutions require strictly aligned image pairs, where both images share identical perspectives and visual content, captured under different lighting conditions. Although this approach is costly and logistically challenging, it enables models to learn precise mappings between specific low-light inputs and their well-lit counterparts, resulting in superior performance and strong information recovery capabilities. AugGAN(Cross domain adaptation with GAN-based data augmentation) (Lin et al., 2020), a GAN(Generative Adversarial Network)-based data augmentation method, effectively transforms road-driving images into the desired domain while preserving image objects. A model based on Pix2Pix (C. Liu & Xu, 2022) is trained on paired images to identify the optimal approach for converting nighttime images to daytime images, significantly enhancing the detection performance of CNN(Convolutional Neural Network) models. To reduce domain gaps, the concept of domain adaptation was introduced (Hong et al., 2023), enabling the transfer of target domain samples to the source domain known to the object detector, such as converting nighttime images to daytime images (Schutera et al., 2020). VQ-VAE(Vector Quantized Variational Autoencoder) generates compact and expressive discrete latent representations, making it an ideal choice for preserving critical details and textures in images (Koohestani et al., 2024), particularly excelling in low-light image enhancement. Diffusion modelshave recently garnered attention in low-level vision tasks, applying iterative refinement to capture both global context and fine details in the translated images (Jiang et al., 2023). Although paired data-based methods typically achieve higher translation fidelity, they demand significant effort to collect strictly aligned nighttime and daytime images. Since the effectiveness of these approaches is closely tied to the availability of high-quality paired datasets, developing efficient strategies for acquiring and maintaining such data remains a critical challenge.

3 | METHODOLOGY

The proposed framework (Figure 1a) addresses nighttime pavement distress detection through a three-stage pipeline. CycleGAN-turbo generates a high-quality paired dataset by transforming unpaired images into aligned pairs, overcoming data scarcity issues. IllumiShiftNet then converts low-light nighttime images into realistic daylight-like representations using an encoder–decoder generator with skip connections for preserving structural details, while multiscale discriminators ensure image quality across various resolutions. The trained model transforms nighttime images into daylight-like representations that can be directly processed by existing detection models

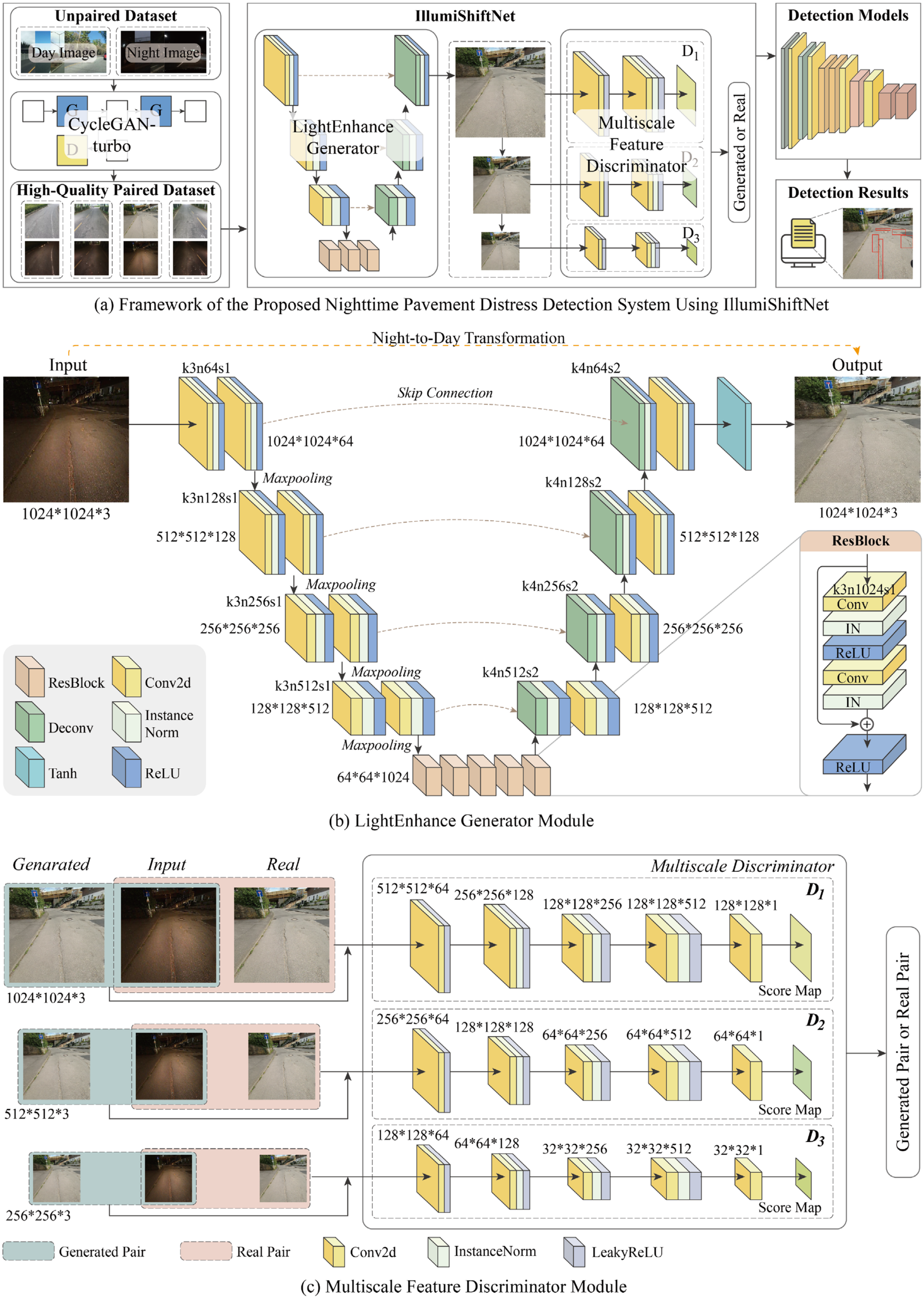


FIGURE 1 Framework and Architectural components of the proposed IllumiShiftNet for nighttime pavement distress detection.

trained on daytime data, eliminating domain gaps without requiring additional nighttime training data. A distress-focused loss function enhances critical pavement feature recovery, ensuring reliable downstream detection performance.

3.1 | Generation of strictly aligned night–day pairs

To address the scarcity of paired nighttime and daytime data, a method is proposed to transform unpaired datasets into strictly aligned pairs. The CycleGAN-turbo (Parmar et al., 2024) model is trained on publicly available unpaired nighttime and daytime datasets (Yu et al., 2020), learning bidirectional mappings between the two domains. Once trained, the model is used to convert daytime images into their nighttime counterparts, ensuring identical perspectives and visual content. This process generates high-quality paired datasets with consistent alignment, capturing both lighting conditions under similar camera angles and environmental settings, as illustrated in the left section of Figure 1a. A significant advantage of this approach is that bounding box annotations for pavement distress are created on the original daytime images, where visibility is optimal, and then directly transferred to the corresponding generated nighttime images. This annotation transfer method ensures consistent labeling across the paired dataset without requiring separate manual annotation of low-visibility nighttime images, which would be both time-consuming and prone to errors.

The CycleGAN-turbo model's bidirectional nature enables it to learn both daytime-to-nighttime and nighttime-to-daytime transformations simultaneously from unpaired data. This framework specifically utilizes the daytime-to-nighttime conversion capability to create synthetic nighttime images, providing aligned pairs that serve as a robust foundation for supervised training. These paired datasets enable the downstream image translation network to benefit from stronger supervision and improved training accuracy compared to models trained on unpaired data alone.

3.2 | Architecture of illumishiftnet

IllumiShiftNet is designed to transform low-light nighttime images $I_{\text{night}} \sim p_{\text{night}}(I)$ into realistic daytime-like images $\hat{I}_{\text{day}} \sim p_{\text{day}}(I)$ with enhanced illumination and preserved details. The framework comprises two key components: the LightEnhance generator and the multiscale feature discriminator, as shown in Figure 1a.

3.2.1 | LightEnhance generator module

The LightEnhance generator employs an encoder–decoder architecture with skip connections (Siddique et al., 2021), designed to transform low-light nighttime images I_{night} into realistic daytime-like representations $\hat{I}_{\text{day}}$, as illustrated in Figure 1b. The encoder progressively extracts hierarchical features while reducing spatial dimensions, capturing both low-level textures and high-level semantic properties.

The encoder downsamples the input image through convolutional layers with instance normalization and ReLU(Rectified Linear Unit) activation. The ResBlock employs two consecutive 3×3 convolutional layers, which provide an optimal balance between receptive field coverage and computational efficiency while maintaining fine-grained feature representation essential for preserving crack details (Simonyan & Zisserman, 2014). Instance normalization (Huang & Belongie, 2017) is specifically chosen over batch normalization to preserve instance-specific statistics, preventing information overload and ensuring stable training across varying lighting conditions. This normalization strategy is particularly important for nighttime images, where illumination varies dramatically across different samples. At each encoder level, feature maps at progressively reduced resolutions capture multiscale information, from fine textures to broader structural patterns.

The decoder reconstructs high-resolution outputs through transposed convolutional layers, with skip connections from corresponding encoder layers. These skip connections directly link encoder and decoder layers of matching dimensions, allowing the network to bypass the bottleneck for certain information flows. This mechanism is crucial for preserving high-frequency details that might otherwise be lost during downsampling, such as fine crack patterns and texture variations. The skip connection design addresses the challenge of nighttime noise processing, as the main encoder–decoder path handles global illumination enhancement while these direct pathways help distinguish between random noise patterns and actual pavement texture. This architecture allows the network to simultaneously enhance global brightness while retaining local structural details, effectively suppressing noise without sacrificing critical crack information.

The generator's objective is to produce images $\hat{I}_{\text{day}}$ that are indistinguishable from real daytime images I_{day}

when evaluated by the discriminator. The generator loss is defined as

$$\mathcal{L}_G = \mathbb{E}_{I_{\text{night}} \sim p_{\text{night}}(I)} \left[(D(I_{\text{night}}, G(I_{\text{night}})) - 1)^2 \right], \quad (1)$$

where $G(I_{\text{night}}) = \hat{I}_{\text{day}}$ is the generator output, D represents the discriminator, and $p_{\text{night}}(I)$ is the probability distribution of nighttime images.

3.2.2 | Multiscale feature discriminator module

The multiscale feature discriminator (Rao et al., 2023) module consists of three parallel discriminators D_1, D_2, and D_3, which operate at different resolutions: full, half, and quarter scales, respectively, as shown in Figure 1c. The primary motivation for this multiscale design stems from the understanding that pavement distress features exist at various spatial scales—from microscopic cracks to macroscopic potholes and broader degradation patterns. A single-scale discriminator would focus on either fine details or global structures, but not both simultaneously.

These discriminators provide complementary feedback to the generator by evaluating different aspects of the translated images. The full-scale discriminator D_1 assesses fine-grained details such as thin cracks and surface texture, the half-scale discriminator D_2 evaluates medium-scale features like crack patterns and local continuity, while the quarter-scale discriminator D_3 focuses on global consistency and larger structural elements. This complementary evaluation ensures that the generator receives comprehensive feedback across all relevant scales for pavement analysis.

Each discriminator evaluates the realism and consistency of the generator's output $\hat{I}_{\text{day}}$ relative to the real daytime image I_{day}, conditioned on the input nighttime image I_{night}. The losses from these discriminators are combined with equal weights, ensuring a balanced contribution from each scale without biasing the model toward any particular resolution level. This approach provides comprehensive evaluation across multiple scales simultaneously, improving the overall quality of the translated images. The discriminator loss is formulated as

$$\begin{aligned} \mathcal{L}_D =& \mathbb{E}_{(I_{\text{night}}, I_{\text{day}}) \sim p_{\text{data}}} \left[(D(I_{\text{night}}, I_{\text{day}}))^2 \right] \\ &+ \mathbb{E}_{I_{\text{night}} \sim p_{\text{night}}} \left[(1 - D(I_{\text{night}}, G(I_{\text{night}})))^2 \right], \end{aligned} \quad (2)$$

where p_{data} represents the joint distribution of real nighttime and daytime image pairs, p_{night} is the distribution of nighttime images, $G(I_{\text{night}}) = \hat{I}_{\text{day}}$ is the generator output, and D is the discriminator.

3.2.3 | Distress-focused loss function

The distress-focused loss function is designed to enhance the accurate reconstruction of pavement distress areas, which are critical for downstream detection tasks. This specialized loss component prioritizes regions identified by bounding boxes $b \in \mathcal{B}$, where $\mathcal{B}$ represents the set of bounding boxes highlighting pavement distress areas.

The distress-focused loss applies a mean squared error penalty specifically to these regions:

$$\mathcal{L}_{\text{distress}} = \lambda_b \sum_{b \in \mathcal{B}} \|\hat{I}[b] - I_{\text{real}}[b]\|_2^2, \quad (3)$$

where $\hat{I}$ and I_{real} denote the generated and real daytime images, respectively, and λ_b is a weighting factor that emphasizes the importance of these regions.

This targeted approach recognizes that not all image regions carry equal importance for pavement distress detection. By explicitly emphasizing distress regions, the loss function guides the generator to allocate more attention to accurately translating these critical areas. This approach aligns with principles from attention mechanisms, where models learn to focus computational resources on the most informative regions of the input.

The weight parameter λ_b is set to 1 in the implementation. This value balances the emphasis on the distressed regions with the general image translation objective. A higher weight would risk over-focusing on bounding box regions at the expense of overall image quality, while a lower weight would provide insufficient guidance for distress-specific enhancement.

Bounding boxes are chosen over pixel-wise segmentation masks for several reasons. Pavement distress detection requires not only the accurate reconstruction of the crack or distress itself but also the surrounding contextual pixels that provide essential information about the pavement surface condition. Bounding boxes effectively capture both the distress and its immediate context, enabling the model to learn the relationship between the crack pattern and its surrounding texture. This approach aligns with how detection models typically operate, focusing on regions rather than isolated pixels.

Although the current implementation primarily targets crack detection, the loss function's design is inherently extensible to other types of pavement distress. The same bounding box approach can be applied to potholes, raveling, rutting, and other surface defects without modification to the underlying methodology. This flexibility ensures that the framework can adapt to diverse pavement monitoring requirements across different infrastructure scenarios.

The overall loss function combines three components: the generator loss $\mathcal{L}_G$, the discriminator loss $\mathcal{L}_D$, and the distress-focused loss $\mathcal{L}_{\text{distress}}$. It is defined as

$$\mathcal{L}_{\text{total}} = \mathcal{L}_G + \mathcal{L}_D + \mathcal{L}_{\text{distress}}, \tag{4}$$

This combined approach ensures that the model simultaneously optimizes for image realism, structural fidelity, and accurate reconstruction of critical pavement distress regions.

4 | EXPERIMENT

4.1 | Experiment setting

The dataset preparation for this study involved a multi-stage process: (1) Initially, the CycleGAN-turbo (Parmar et al., 2024) model was trained on the BDD100k dataset (Yu et al., 2020) to learn bidirectional mappings between nighttime and daytime domains. (2) This pretrained model was then applied to transform 2330 daytime pavement images, collected from Aachen, Germany, and Nanjing, China, supplemented with samples from the RDD2022 dataset (Arya et al., 2024). This transformation process generated 2330 synthetic nighttime images that maintained precise spatial alignment with their daytime counterparts, resulting in 4660 total images (2330 paired samples) for IllumiShiftNet training. (3) To evaluate real-world performance, an additional dataset of 529 authentic nighttime images was collected under varied illuminance conditions: 194 images at >15 lux, 183 images at 11–15 lux, and 152 images at <1 lux. These images served as inputs to the trained IllumiShiftNet model, with the translated outputs subsequently used to assess downstream detection performance. Throughout all experiments, datasets were consistently divided into training and testing sets following an approximate 5:1 ratio to ensure reliable evaluation.

Images were acquired using a Hikvision MV-CE060-10UM industrial camera with a WL1224-4MP 12 mm lens. Experiments were conducted on a system with Ubuntu 22.04, dual Intel 5318Y processors, and four RTX 4090 GPUs. The software setup utilized Python 3.8 and PyTorch 2.1. The hyperparameter settings for IllumiShiftNet training are summarized in Table 1.

To ensure training stability and convergence, several advanced techniques were implemented. Instance normalization (Huang & Belongie, 2017) was applied to reduce the instability in generator outputs, minimizing the risk of sudden loss fluctuations and mode collapse by normalizing features at the instance level. Spectral normalization (Miyato et al., 2018) was employed in discriminator networks to constrain the Lipschitz constant of network layers, preventing the discriminator from becoming overly confident and maintaining gradient stability throughout training. Additionally, a replay buffer (Eysenbach et al., 2019) mechanism was integrated to reduce oscillations in the adversarial training process by storing and randomly sampling from previously generated images, which helps mitigate mode collapse and stabilizes the generator's learning progression.

TABLE 1 Hyperparameter settings for IllumiShiftNet

Parameter	Value
Number of epochs	200
Batch size	8
Learning rate	2×10^{-4}
Optimizer type	Adam
Adam β_1	0.5
Input image size	1024×1024

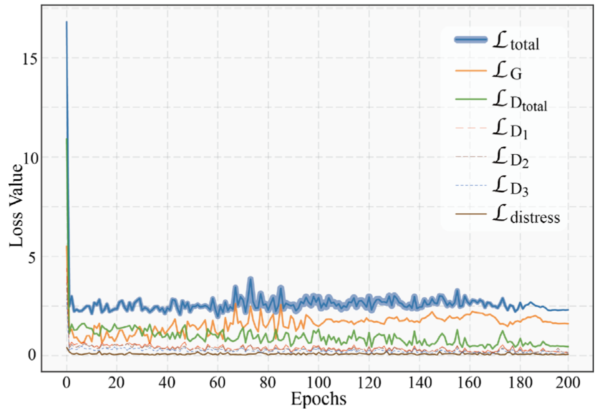


FIGURE 2 Convergence patterns of IllumiShiftNet training loss components.

The training process demonstrated consistent convergence, as illustrated in Figure 2, which shows the progression of various loss components over 200 epochs. After an initial phase of rapid loss reduction, all components gradually stabilize, with the total loss and individual discriminator losses reaching equilibrium values. This convergence pattern validates the effectiveness of the training strategy and stability mechanisms implemented in the IllumiShiftNet framework.

A demonstration of the IllumiShiftNet model is available at http://yuaniscute.pink:2225/, where both the generated paired dataset and source code can be accessed.

4.2 | Evaluating indicator

To evaluate IllumiShiftNet's performance, two sets of metrics are employed: image quality metrics and detection accuracy metrics.

For image quality assessment, SSIM and PSNR (Hore & Ziou, 2010) measure the fidelity between translated images and ground truth. SSIM quantifies the structural similarity between the translated image $\hat{I}$ and the ground truth image I_{real}:

$$\text{SSIM}(\hat{I}, I_{\text{real}}) = \frac{(2\mu_{\hat{I}}\mu_{I_{\text{real}}} + C_1)(2\sigma_{\hat{I}I_{\text{real}}} + C_2)}{(\mu_{\hat{I}}^2 + \mu_{I_{\text{real}}}^2 + C_1)(\sigma_{\hat{I}}^2 + \sigma_{I_{\text{real}}}^2 + C_2)}. \quad (5)$$

Higher SSIM values indicate better structural preservation. PSNR measures pixel-level fidelity and is defined as

$$\text{PSNR} = 10 \cdot \log_{10}\left(\frac{\text{MAX}^2}{\text{MSE}}\right), \quad (6)$$

where MAX represents the maximum pixel value and MSE is the mean squared error between images. A higher PSNR value signifies greater similarity to the ground truth image.

For detection accuracy, mean average precision (mAP) at 50% intersection over union threshold and F1-score evaluate the model's performance in identifying pavement distress.

5 | DISCUSSION

5.1 | Experiment results

The experimental results demonstrate the effectiveness of the proposed approach, encompassing both the paired dataset generation using CycleGAN-turbo and the night-to-day image translation performance of IllumiShiftNet.

Figure 3a presents the paired dataset generated by CycleGAN-turbo (Parmar et al., 2024), representing selections from the complete set of 2330 image pairs. The top row shows the original daytime images, while the bottom row depicts the corresponding nighttime images generated by the model. These generated pairs exhibit consistent alignment in spatial content and structural integrity, addressing the critical challenge of paired data scarcity. The quality of the synthetic nighttime images is quantitatively validated by a Fréchet inception distance score of 26.4, indicating strong perceptual similarity to real nighttime images while maintaining the essential structural information from the source daytime images. This score demonstrates that the generated data effectively mimics the distribution of real nighttime imagery while preserving the underlying pavement features necessary for training.

Figure 3b demonstrates the results of IllumiShiftNet applied to real-world nighttime images. The top row contains the original nighttime images, characterized by low visibility, high noise levels, and reduced contrast. The bottom row displays the corresponding outputs from IllumiShiftNet, which successfully translates the low-light images into daytime-like representations. Importantly, the translated images exhibit a significant increase in entropy values, as shown in the figure. This entropy increase reflects the model's ability to recover information technically present in nighttime images but masked by low SNR. According to information theory, higher entropy indicates greater information content (Zhou et al., 2023), suggesting IllumiShiftNet performs selective nonlinear transformations that amplify weak signals while suppressing noise. This functions as sophisticated denoising that preserves underlying signal structure while removing stochastic noise components (Goyal et al., 2020). The recovered details include fine crack patterns and pavement textures, demonstrating the model's effectiveness in enhancing both global illumination and critical local features.

The combined results validate the effectiveness of the proposed framework. CycleGAN-turbo effectively generates high-quality paired datasets, while IllumiShiftNet demonstrates superior night-to-day translation performance. By ensuring the preservation of critical crack features and recovering lost information, the framework enhances the overall suitability of the translated images for downstream pavement distress detection tasks.

5.2 | Ablation study of model components

To evaluate the individual contribution of each component within the proposed IllumiShiftNet framework, an ablation study was conducted using the set of 2330 image pairs. Three comparative models were designed based on the removal of key architectural elements:

- *Baseline*: This model includes the basic generator and discriminator without the multiscale feature discriminator or distress-focused loss.
- *M1*: Adds the distress-focused loss function to the baseline model to prioritize pavement distress regions during training.
- *M2*: Integrates multiscale feature discriminators with the baseline model to enhance multiresolution consistency but excludes the distress-focused loss.

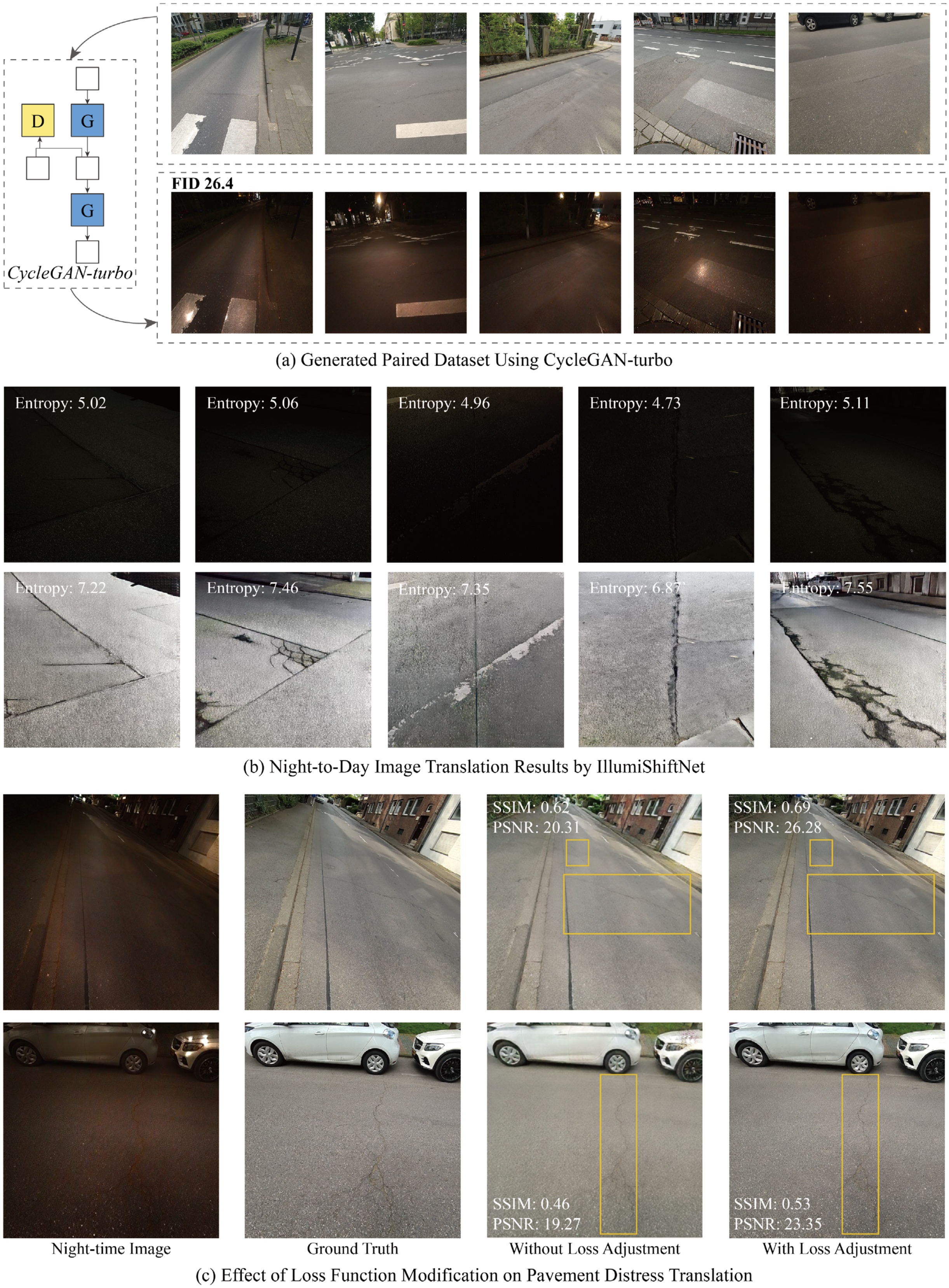


FIGURE 3 Dataset generation, translation results, and ablation study of IllumiShiftNet.

TABLE 2 Ablation study results.

Model	Multiscale feature	Distress-focused loss	PSNR(↑)	SSIM(↑)
Baseline	✗	✗	22.64	0.61
M1	✗	✓	25.37	0.69
M2	✓	✗	26.91	0.71
IllumiShiftNet	✓	✓	**28.50**	**0.78**

Abbreviations: PSNR: peak signal-to-noise ratio; SSIM: structural similarity index measure.

The quantitative results in Table 2 highlight the contributions of the multiscale discriminator and the distress-focused loss. Compared to the baseline model, adding the distress-focused loss alone (M1) improves PSNR by 2.73, showing its effectiveness in enhancing structural accuracy in distress regions. This suggests that the proposed loss helps the model focus on reconstructing critical pavement features.

Using only the multiscale discriminator (M2) yields even better results, with PSNR increasing by 4.27 and SSIM by 0.10 over the baseline. This demonstrates that the multiscale discriminator helps capture both fine textures and the global structure.

Combining both components in the full IllumiShiftNet model brings the largest gains—PSNR improves by 5.86 and SSIM by 0.17 compared to the baseline. These results confirm that the two modules are complementary: the multiscale discriminator improves global consistency, while the distress-focused loss enhances the reconstruction of fine distress details.

Figure 3c provides visual evidence supporting these findings, particularly for M1. Without the distress-focused loss, translated images fail to restore fine pavement cracks accurately. In contrast, when the distress-focused loss is applied, visible crack patterns are more distinct and consistent with the ground truth, emphasizing its critical role in recovering key distress features under low-light conditions.

5.3 | Comparison with other models and real-time performance analysis

The comparison between IllumiShiftNet and other image translation models, including Pix2pix (Isola et al., 2017), CycleGAN-turbo (Parmar et al., 2024), and Brownian bridge diffusion models (BBDM) (B. Li et al., 2023), is summarized in Table 3 and Figure 4a. The evaluation metrics include PSNR, SSIM, and inference speed, highlighting the performance of each model in terms of translation quality and computational efficiency.

TABLE 3 Comparison of image translation models.

Model	PSNR	SSIM	Speed (ms)
Ours (IllumiShiftNet)	**28.5**	**0.78**	146.7
Pix2pix	24.2	0.63	**82.9**
CycleGAN-turbo	15.6	0.56	1160
BBDM	27.8	0.74	1897

Abbreviations: BBDM, Brownian bridge diffusion models; PSNR: peak signal-to-noise ratio; SSIM: structural similarity index measure.

As shown in Figure 4a, Pix2pix broadly restores the overall structure of nighttime images but struggles with preserving finer details. Visual artifacts and noise frequently appear in high-texture regions, such as crack edges and textured pavement surfaces. Although it offers the fastest inference speed among the compared models, its limited detail fidelity and presence of noise reduce its suitability for high-precision tasks like pavement distress detection.

CycleGAN-turbo introduces noticeable color distortions in the translated outputs, especially over pavement regions, which often appear unnaturally saturated or tinted. These artifacts compromise visual realism and may confuse downstream detection models. Moreover, its relatively slow inference time limits its applicability for real-time scenarios.

The diffusion-based BBDM model provides improvements in texture preservation and lighting consistency compared to GAN-based approaches. However, it suffers from extremely high computational complexity, resulting in the slowest inference speed among all models. This makes it impractical for real-time detection applications.

In comparison, IllumiShiftNet significantly outperforms the competing models, especially the unpaired translation method CycleGAN-turbo. It achieves an improvement of 12.9 in PSNR and 0.22 in SSIM, highlighting its superior ability to generate high-fidelity, structurally accurate translations. As shown in Figure 4a, IllumiShiftNet produces visually clean outputs, free from the noise, color distortions, and artifacts commonly observed in the results of CycleGAN-turbo and other baselines.

Furthermore, as shown in Figure 5, IllumiShiftNet achieves real-time performance across mainstream NVIDIA GPUs when optimized with lower-precision formats. On all tested platforms–including RTX 4090, 3090, and 4070—both FP16 and INT8 quantization reduce inference times well below the 33-ms threshold required for 30 FPS real-time processing. The real-time capability enables deployment as a preprocessing step for continuous video stream processing, where each frame undergoes night-to-day translation before detection analysis. These results demonstrate the model's practicality for deployment in nighttime pavement inspection scenarios.

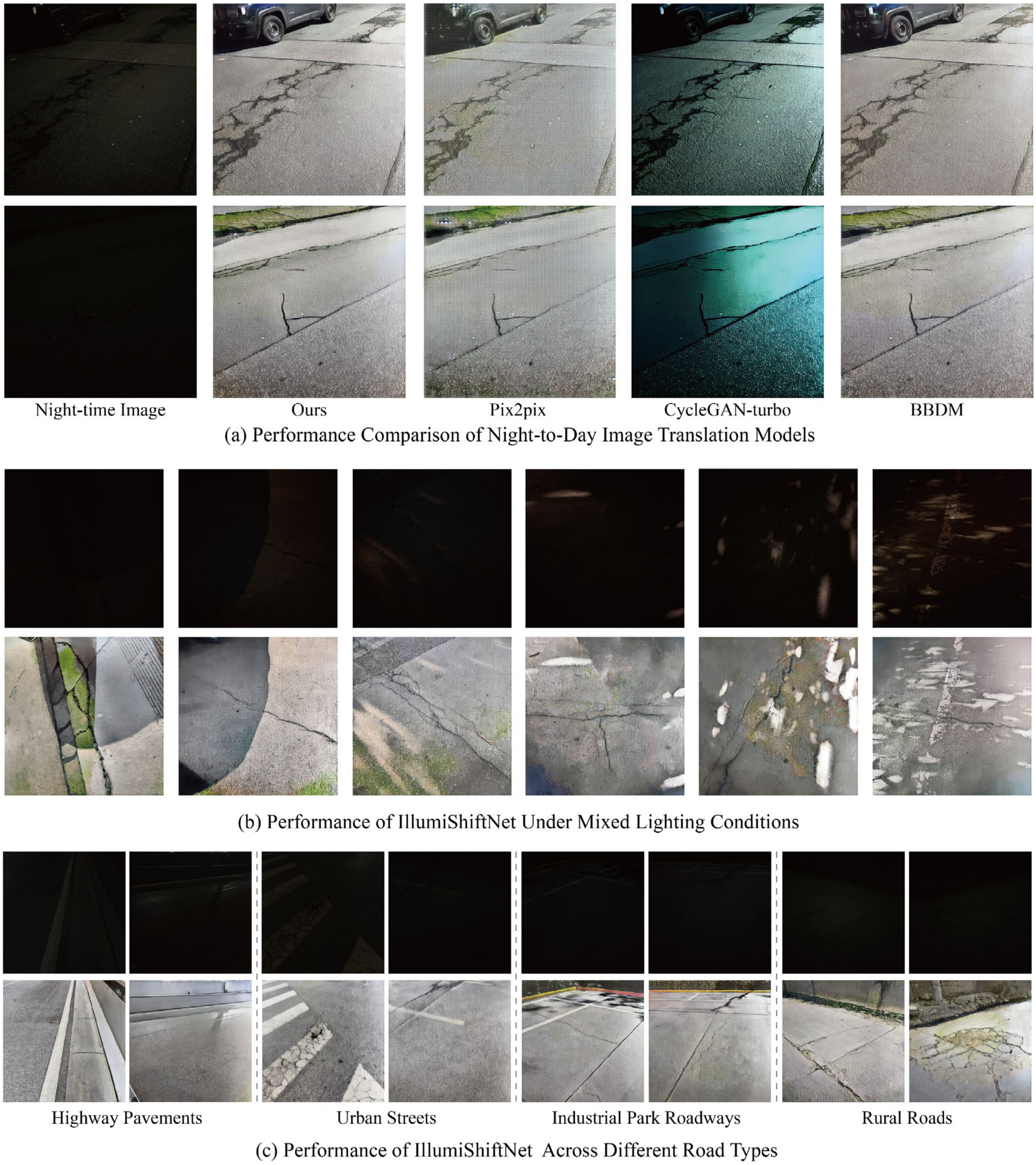


FIGURE 4 Comprehensive performance analysis of IllumiShiftNet: Model comparisons and generalization evaluation.

5.4 | Generalization and practical performance across varying illuminance and weather conditions

To evaluate the practical applicability of IllumiShiftNet, additional field tests were conducted across varying illumination levels and challenging environmental conditions. Three representative groups of images were captured at different illuminance levels in Aachen, Germany, with selected examples shown in the results. Illuminance values were precisely measured at different times using a calibrated Lux meter, ranging from early evening (274.38 lux) to complete darkness (0.03 lux).

Figure 6a demonstrates the progressive evaluation of IllumiShiftNet's performance across this illuminance spectrum. The top and third rows present original images captured under increasingly challenging lighting conditions. As illuminance decreases, particularly below 20 lux

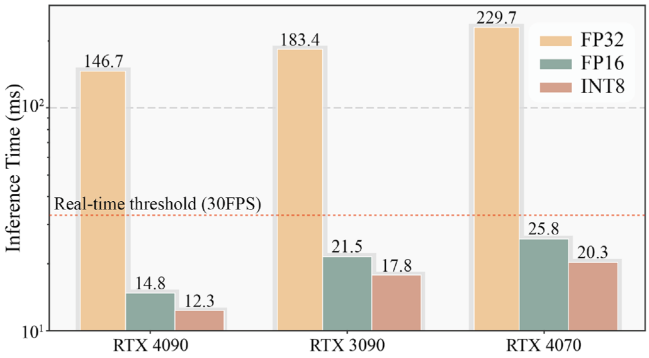


FIGURE 5 Real-time inference capability of IllumiShiftNet on multiple GPU platforms.

(after 6:30 pm), a significant degradation in source image quality becomes evident. Critical pavement features, including cracks and lane markings, become increasingly obscured, posing substantial challenges for conventional detection systems.

The second and fourth rows display the corresponding translated outputs generated by IllumiShiftNet. At higher illuminance levels (274.38 and 85.21 lux), the model consistently produces high-fidelity daylight-like representations with excellent preservation of pavement texture and distress details. This performance remains robust through moderate low-light conditions (16.87 and 1.54 lux), with translations retaining structural coherence and critical feature clarity.

At the extreme illuminance threshold of 0.03 lux (midnight), a notable decline in reconstruction quality becomes apparent. This performance degradation can be attributed to fundamental information theory constraints rather than algorithmic limitations. At such minimal illuminance levels, the original nighttime images operate at exceptionally low SNRs, where photon count becomes the primary limiting factor. According to the Shannon–Hartley theorem, when SNR approaches the noise floor, as in near-total darkness, the recoverable information becomes inherently limited regardless of the sophistication of the recovery algorithm.

Despite these physical constraints, IllumiShiftNet demonstrates remarkable robustness, still producing structurally coherent outputs that retain sufficient information for practical distress detection applications. This demonstrates that even under extreme low-light conditions, the proposed framework maintains operational utility, though with acknowledged precision trade-offs in fine crack detail reconstruction.

To further validate generalization capabilities beyond illuminance variations, additional experiments were conducted under rainy nighttime conditions, as illustrated in Figure 6b. Rainwater introduces multiple imaging challenges including specular reflections, dynamic light scattering, and reduced visibility due to atmospheric interference. Despite these compounding factors, IllumiShiftNet successfully generated high-quality translations that preserved critical pavement features and maintained structural integrity across the tested illuminance range. This robust performance under combined challenges of low light and precipitation demonstrates the model's practical resilience in diverse real-world conditions.

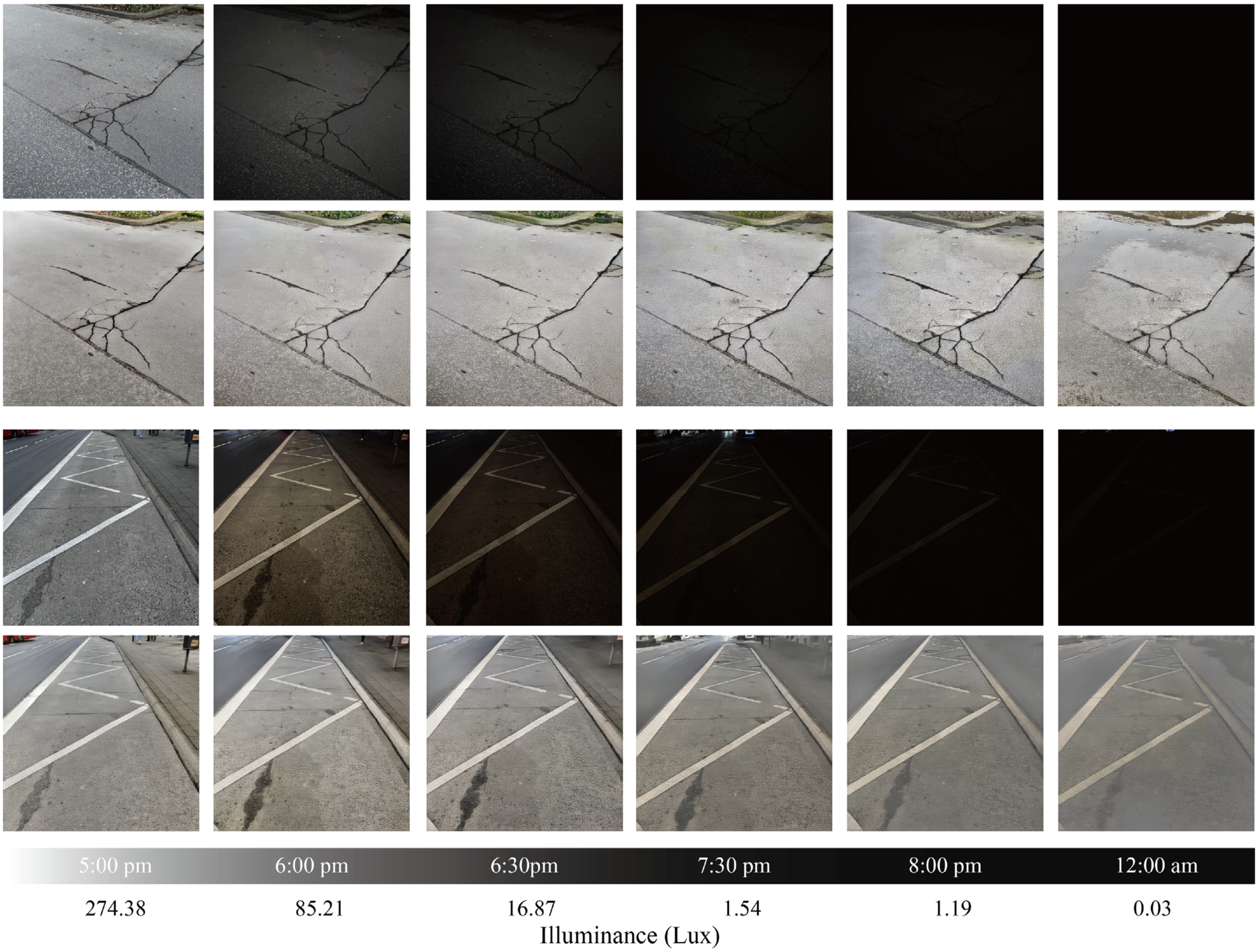
An important observation from these experiments is that standard camera sensors capture substantially more information in low-light conditions than is immediately apparent to human observers. IllumiShiftNet effectively utilizes this latent information, functioning as an advanced signal recovery system that separates meaningful structural data from noise components. This capability reduces reliance on specialized lighting equipment across most operational scenarios, potentially lowering system complexity and deployment costs. The comprehensive evaluation across both illuminance gradients and adverse weather conditions establishes IllumiShiftNet as a versatile solution for nighttime pavement monitoring, demonstrating consistent performance across the operational spectrum typically encountered in real-world inspection scenarios despite fundamental physical limitations at extreme low-light boundaries.

5.5 | Performance of detection models

Crack detection models were evaluated across multiple operational scenarios to validate IllumiShiftNet's effectiveness in enhancing pavement distress detection under low-light conditions and to demonstrate its robustness for downstream applications. Three detection models—YOLOv11 (Khanam & Hussain, 2024), Co-DETR (Zong et al., 2023), and ViTDet (Y. Li et al., 2022)—were assessed using mAP@50 and F1-score, as shown in Table 4. The detection models were trained using the 2330 paired dataset (day and night) and further evaluated on the additional 529 nighttime images with varying illuminance levels.

In Mission 1, models trained on daytime data were tested directly on nighttime images. Performance was poor across all models due to low visibility and high noise levels. In Mission 2, models trained and tested on nighttime images performed better than in Mission 1, but collecting and labeling large-scale nighttime data remains costly and challenging.

Mission 3, serving as the baseline scenario, trained and tested the models on daytime images. As expected, this scenario yielded the best performance metrics, demon-



(a) Performance Across Varying Illuminance Levels for Night-to-Day Image Translation

(b) Generalization Performance Under Rainy Nighttime Conditions

FIGURE 6 IllumiShiftNet performance across illuminance levels and rainy night conditions.

TABLE 4 Detection performance across different missions.

Mission	Mission 1 Day-trained, Night-tested		Mission 2 Night-trained, Night-tested		Mission 3 Day-trained, Day-tested		Mission 4 Day-trained, Night-to-day-tested (IllumiShiftNet)	
Model	**mAP@50**	**F1 score**	**mAP@50**	**F1 score**	**mAP@50**	**F1 score**	**mAP@50**	**F1 score**
YOLOv11	0.48	0.43	0.71	0.68	0.94	0.91	0.91	0.87
Co-DETR	0.53	0.50	0.67	0.65	0.92	0.88	0.87	0.81
ViTDet	0.40	0.34	0.64	0.63	0.91	0.86	0.88	0.83

Abbreviation: mAp: mean average precision.

TABLE 5 Detection performance of Mission 4 across different illuminance levels.

Model	Low illuminance (<1 lux)		Medium illuminance (1–15 lux)		High illuminance (>15 lux)	
	mAP@50	**F1 score**	**mAP@50**	**F1 score**	**mAP@50**	**F1 score**
YOLOv11	0.82	0.76	0.89	0.83	0.93	0.89
Co-DETR	0.75	0.69	0.84	0.77	0.90	0.85
ViTDet	0.77	0.71	0.85	0.78	0.89	0.84

Abbreviation: mAp: mean average precision.

strating the effectiveness of models when operating within the same domain.

In Mission 4, models were trained on daytime data but tested on nighttime images translated to the daytime domain using IllumiShiftNet. Performance was nearly identical to the baseline, demonstrating that image translation effectively bridges the domain gap. YOLOv11 achieved a mAP@50 of 0.91, showing that detection models can operate reliably on translated images without retraining on nighttime data.

The comparison between Mission 1 and Mission 4 underscores the value of image translation models like IllumiShiftNet. By using such a framework, Mission 4 achieves performance levels comparable to daytime detection, offering a practical and scalable solution for high-accuracy nighttime pavement crack detection without requiring extensive nighttime-specific datasets.

To further evaluate the robustness of Mission 4 under varying lighting conditions, detection performance was analyzed across three illuminance levels, as shown in Table 5. Based on field measurements in Figure 6a, conditions were categorized as high illuminance (>15 lux, early evening with ambient lighting), medium illuminance (1–15 lux, typical nighttime with minimal lighting), and low illuminance (<1 lux, near-complete darkness).

At high illuminance levels, all models maintained near-optimal performance, closely approaching daytime benchmarks. Under medium illuminance conditions, detection accuracy declined slightly, with YOLOv11's mAP decreasing by approximately 4%, though performance remained robust. These results indicate that moderate ambient lighting suffices for effective translation and detection.

Under extremely low-light conditions (<1 lux), all models showed more substantial performance drops, with mAP decreasing by 11–15% compared to higher illuminance scenarios. This limitation reflects the severe loss of visual information in original images, where critical pavement features become nearly absent due to extremely low SNRs. While IllumiShiftNet recovers significant detail under such challenging conditions, the available information in near-total darkness imposes a fundamental ceiling on recoverable quality. Combining the proposed method with auxiliary lighting, such as low-power LED fill lights, could further mitigate this limitation in real-world deployments.

5.6 | Generalization performance under mixed lighting and across diverse road types

To evaluate IllumiShiftNet's adaptability under real-world conditions, experiments were conducted across scenarios with partial street lighting interference and diverse road types. Such lighting often leads to nonuniform illumination, posing challenges for nighttime pavement inspection.

Figure 4b demonstrates the model's performance under such challenging conditions. IllumiShiftNet successfully processes images with heterogeneous lighting, maintaining overall structural coherence and pavement feature preservation. The model effectively handles transition zones between illuminated and dark regions, demonstrating adaptability to spatially varying illumination patterns. However, in regions with strong contrast caused by direct light sources, the enhanced images may exhibit slight overexposure. As illustrated in the last two image pairs of Figure 4b, high-intensity illumination can lead to localized brightness saturation, which affects the preservation of fine pavement details. This limitation is primarily attributed to the training data, which did not explicitly account for brightness balancing under such extreme lighting conditions.

Testing was also conducted across four distinct road categories: highway pavements, urban streets, industrial park roadways, and rural roads, as illustrated in Figure 4c. Images were acquired using a Nikon D850 DSLR camera with a 24–70 mm *f*/2.8 lens (aperture: *f*/4.0, shutter speed: 1/60s, ISO: 400), representing a different imaging system from the industrial camera used in earlier experiments. IllumiShiftNet produced consistent high-quality translations across all road categories. The model successfully preserved longitudinal crack details and lane markings on highways, captured complex crack networks in urban settings, maintained wide-crack patterns in industrial environments, and achieved reliable translations on rural roads despite material inconsistency.

6 | CONCLUSIONS

This paper presents IllumiShiftNet, a novel model that transforms nighttime pavement images into daylight-like representations for enhanced distress detection. The model effectively preserves critical pavement details while addressing the challenges of low visibility and noise in nighttime imagery. The quantitative evaluation demonstrates superior image quality with a PSNR of 28.5 and SSIM of 0.78, significantly outperforming state-of-the-art alternatives.

IllumiShiftNet utilizes an unpaired image translation model to generate paired training datasets while incorporating a distress-focused loss function that ensures accurate reconstruction of critical crack details. These innovations effectively balance global illumination enhancement with the preservation of fine structural details essential for distress detection.

Comparative analysis demonstrated that IllumiShiftNet achieves superior performance in both qualitative and quantitative metrics, eliminating common issues such as color distortions, noise artifacts, and detail loss. Quantization techniques enable real-time inference across GPU platforms, with latency consistently below the 33-ms threshold for 30 FPS operation.

Experimental results revealed that IllumiShiftNet maintains effective performance from moderate low-light down to challenging near-darkness conditions. At the extreme threshold of 0.03 lux, while reconstruction quality declined, the model still produced structurally coherent outputs sufficient for practical distress detection. The system demonstrated similar robustness under rainy conditions, validating its utility under diverse environmental challenges.

Detection performance analysis established that IllumiShiftNet effectively bridges the domain gap between nighttime and daytime imagery. By enabling detection models trained on daytime datasets to perform effectively on translated nighttime images, the model eliminates the need for extensive nighttime-specific training data, achieving performance comparable to native daytime detection.

Generalization testing across diverse road types, mixed lighting conditions, and different imaging equipment confirmed the model's adaptability and consistent performance across varying pavement and distress patterns for real-world deployment.

Despite these achievements, limitations exist in extreme low-light conditions (below 1 lux) where performance degradation occurs due to fundamental information constraints, and safety considerations for inspection vehicles operating in darkness remain a practical concern. Future research will explore several directions to address current limitations and enhance system capabilities. Enhanced loss function architectures and algorithmic improvements targeting ultra-low-light image enhancement could improve performance at extreme low-light boundaries and potentially eliminate the need for additional lighting equipment. Extending the framework to support semantic segmentation models would enable more precise distress delineation, while addressing varying camera distances and angles will ensure reliable performance in diverse real-world deployment scenarios. Expanding the framework to other adverse weather conditions would further enhance practical applicability. These advancements would collectively advance the autonomy and effectiveness of nighttime pavement monitoring systems.

ACKNOWLEDGMENTS

The language and grammar of this manuscript were partially refined using a large language model (Claude). This paper is funded by the German Research Foundation (DFG) under LI 3613/5-2 with Project ID 414936990.

Open access funding enabled and organized by Projekt DEAL.

REFERENCES

Alam, K. M. R., Siddique, N., & Adeli, H. (2020). A dynamic ensemble learning algorithm for neural networks. *Neural Computing and Applications*, *32*(12), 8675–8690.

Arya, D., Maeda, H., Ghosh, S. K., Toshniwal, D., & Sekimoto, Y. (2024). Rdd2022: A multi-national image dataset for automatic road damage detection. *Geoscience Data Journal*, *11*(4), 846–862.

Chen, Z., Pawar, K., Ekanayake, M., Pain, C., Zhong, S., & Egan, G. F. (2023). Deep learning for image enhancement and correction in magnetic resonance imaging—State-of-the-art and challenges. *Journal of Digital Imaging*, *36*(1), 204–230.

Eysenbach, B., Salakhutdinov, R. R., & Levine, S. (2019). Search on the replay buffer: Bridging planning and reinforcement learning. *Advances in Neural Information Processing Systems*, *32*, 15246–15257.

Gavaskar, R. G., & Chaudhury, K. N. (2018). Fast adaptive bilateral filtering. *IEEE Transactions on Image Processing*, *28*(2), 779–790.

Goyal, B., Dogra, A., Agrawal, S., Sohi, B. S., & Sharma, A. (2020). Image denoising review: From classical to state-of-the-art approaches. *Information Fusion*, *55*, 220–244.

Gu, C., Lu, X., He, Y., & Zhang, C. (2020). Blur removal via blurred-noisy image pair. *IEEE Transactions on Image Processing*, *30*, 345–359.

Hong, D., Zhang, B., Li, H., Li, Y., Yao, J., Li, C., Werner, M., Chanussot, J., Zipf, A., & Zhu, X. X. (2023). Cross-city matters: A multimodal remote sensing benchmark dataset for cross-city semantic segmentation using high-resolution domain adaptation networks. *Remote Sensing of Environment*, *299*, 113856.

Hore, A., & Ziou, D. (2010). Image quality metrics: PSNR vs. SSIM. In *2010 20th international conference on pattern recognition* (pp. 2366–2369). IEEE.

Huang, X., & Belongie, S. (2017). Arbitrary style transfer in real-time with adaptive instance normalization. In *Proceedings of the IEEE international conference on computer vision* (pp. 1501–1510). IEEE.

Isola, P., Zhu, J.-Y., Zhou, T., & Efros, A. A. (2017). Image-to-image translation with conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1125–1134). IEEE.

Jiang, H., Luo, A., Fan, H., Han, S., & Liu, S. (2023). Low-light image enhancement with wavelet-based diffusion models. *ACM Transactions on Graphics (TOG)*, *42*(6), 1–14.

Khanam, R., & Hussain, M. (2024). *Yolov11: An overview of the key architectural enhancements*. arXiv. https://doi.org/10.48550/arXiv.2410.17725

Koohestani, F., Babak, Z. N. S., Karimi, N., & Samavi, S. (2024). Preserving details in darkness: A VQ-VAE based approach with SSIM loss for low-light image enhancement. In *2024 IEEE world AI IoT congress (AIIoT)* (pp. 342–348). IEEE.

Li, B., Xue, K., Liu, B., & Lai, Y.-K. (2023). BBDM: Image-to-image translation with Brownian bridge diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern Recognition*, (pp. 1952–1961). IEEE.

Li, J., Xu, Z., Fu, L., Zhou, X., & Yu, H. (2021). Domain adaptation from daytime to nighttime: A situation-sensitive vehicle detection and traffic flow parameter estimation framework. *Transportation Research Part C: Emerging Technologies*, *124*, 102946.

Li, L., Wang, R., Wang, W., & Gao, W. (2015). A low-light image enhancement method for both denoising and contrast enlarging. In *2015 IEEE international conference on image processing (ICIP)* (pp. 3730–3734). IEEE.

Li, Y., Mao, H., Girshick, R., & He, K. (2022). Exploring plain vision transformer backbones for object detection. In *European conference on computer vision* (pp. 280–296). Springer.

Lin, C.-T., Huang, S.-W., Wu, Y.-Y., & Lai, S.-H. (2020). GAN-based day-to-night image style transfer for nighttime vehicle detection. *IEEE Transactions on Intelligent Transportation Systems*, *22*(2), 951–963.

Liu, C., & Xu, B. (2022). A night pavement crack detection method based on image-to-image translation. *Computer-Aided Civil and Infrastructure Engineering*, *37*(13), 1737–1753.

Liu, Y., Liu, F., Huang, Y., Hu, J., Zhang, W., & Hou, Y. (2025). The real-time pavement distress detection system based on edge-cloud collaborative computing. *IEEE Transactions on Intelligent Transportation Systems*, Advance online publication. https://doi.org/10.1109/TITS.2025.3544240

Malekloo, A., Liu, X. C., & Sacharny, D. (2024). AI-enabled airport runway pavement distress detection using Dashcam imagery. *Computer-Aided Civil and Infrastructure Engineering*, *39*(16), 2481–2499.

Maryami, R., Arcondoulis, E. J., Guo, J., & Liu, Y. (2024). Experimental investigation of active local blowing on the aerodynamic noise reduction of a circular cylinder. *Journal of Sound and Vibration*, *578*, 118360.

Miao, Y., Liu, F., Hou, T., Liu, L., & Liu, Y. (2020). A nighttime vehicle detection method based on Yolo v3. In *2020 Chinese automation congress (CAC)* (pp. 6617–6621). IEEE.

Miyato, T., Kataoka, T., Koyama, M., & Yoshida, Y. (2018). *Spectral normalization for generative adversarial networks*. arXiv. https://doi.org/10.48550/arXiv.1802.05957

Montenegro, B., & Flores-Calero, M. (2022). Pedestrian detection at daytime and nighttime conditions based on Yolo-v5. *Ingenius*, *27*, 85–95.

Pan, N., Liu, H., Wu, D., Liu, C., & Du, Y. (2023). Spatiotemporal matching method for tracking pavement distress using high-frequency detection data. *Computer-Aided Civil and Infrastructure Engineering*, *38*(16), 2257–2278.

Parihar, A. S., Anand, P., Sharma, A., & Grover, A. (2021). UndarkGAN: Low-light image enhancement with cycle-consistent adversarial networks. In *2021 Fifth international conference on computer, communication and signal processing (ICCCSP)* (pp. 1–7). IEEE.

Parmar, G., Park, T., Narasimhan, S., & Zhu, J.-Y. (2024). *One-step image translation with text-to-image models*. arXiv. https://doi.org/10.48550/arXiv.2403.12036

Pereira, D. R., Piteri, M. A., Souza, A. N., Papa, J. P., & Adeli, H. (2020). FEMA: A finite element machine for fast learning. *Neural Computing and Applications*, *32*, 6393–6404.

Rafiei, M. H., & Adeli, H. (2017). A new neural dynamic classification algorithm. *IEEE Transactions on Neural Networks and Learning Systems*, *28*(12), 3074–3083.

Rafiei, M. H., Gauthier, L. V., Adeli, H., & Takabi, D. (2022). Self-supervised learning for electroencephalography. *IEEE Transactions on Neural Networks and Learning Systems*, *35*(2), 1457–1471.

Rafiei, M. H., Gauthier, L. V., Adeli, H., & Takabi, D. (2024). Self-supervised learning for near-wild cognitive workload estimation. *Journal of Medical Systems*, *48*(1), 107.

Rajinikanth, V., & Razmjooy, N. (2023). Design of a system for melanoma diagnosis using image processing and hybrid optimization techniques. In *Metaheuristics and optimization in computer and electrical engineering: Volume 2: Hybrid and improved algorithms* (pp. 241–279). Springer.

Rao, J., Ke, A., Liu, G., & Ming, Y. (2023). Ms-GAN: Multi-scale GAN with parallel class activation maps for image reconstruction. *The Visual Computer*, *39*(5), 2111–2126.

Schutera, M., Hussein, M., Abhau, J., Mikut, R., & Reischl, M. (2020). Night-to-day: Online image-to-image translation for object detection within autonomous driving by night. *IEEE Transactions on Intelligent Vehicles*, *6*(3), 480–489.

Siddique, N., Paheding, S., Elkin, C. P., & Devabhaktuni, V. (2021). U-net and its variants for medical image segmentation: A review of theory and applications. *IEEE Access*, *9*, 82031–82057.

Simonyan, K., & Zisserman, A. (2014). *Very deep convolutional networks for large-scale image recognition*. arXiv. https://doi.org/10.48550/arXiv.1409.1556

Tang, G., Ni, J., Chen, Y., Cao, W., & Yang, S. X. (2023). An improved CycleGAN-based model for low-light image enhancement. *IEEE Sensors Journal*, *24*(14), 21879–21892.

Tong, Z., Yuan, D., Gao, J., & Wang, Z. (2020). Pavement defect detection with fully convolutional network and an uncertainty framework. *Computer-Aided Civil and Infrastructure Engineering*, *35*(8), 832–849.

Wu, K., Huang, J., Ma, Y., Fan, F., & Ma, J. (2023). Cycle-Retinex: Unpaired low-light image enhancement via Retinex-inline cycleGAN. *IEEE Transactions on Multimedia*, *26*, 1213–1228.

Yu, F., Chen, H., Wang, X., Xian, W., Chen, Y., Liu, F., Madhavan, V., & Darrell, T. (2020). Bdd100k: A diverse driving dataset for heterogeneous multitask learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 2636–2645). IEEE.

Yuan, Z., Zeng, J., Wei, Z., Jin, L., Zhao, S., Liu, X., Zhang, Y., & Zhou, G. (2023). Clahe-based low-light image enhancement for robust object detection in overhead power transmission system. *IEEE Transactions on Power Delivery*, *38*(3), 2240–2243.

Zhang, F., Shao, Y., Sun, Y., Gao, C., & Sang, N. (2023). Self-supervised low-light image enhancement via histogram equalization prior. In *Chinese conference on pattern recognition and computer vision (PRCV)* (pp. 63–75). Springer.

Zhang, Y., & Zhang, L. (2024). A generative adversarial network approach for removing motion blur in the automatic detection of pavement cracks. *Computer-Aided Civil and Infrastructure Engineering*, *39*(22), 3412–3434.

Zhou, X., Zheng, X., Shu, T., Liang, W., Kevin, I., Wang, K., Qi, L., Shimizu, S., & Jin, Q. (2023). Information theoretic learning-enhanced dual-generative adversarial networks with causal representation for robust OOD generalization. *IEEE Transactions on Neural Networks and Learning Systems*, *6*(2), 2066–2079.

Zhu, G., Liu, J., Fan, Z., Yuan, D., Ma, P., Wang, M., Sheng, W., & Wang, K. C. (2024). A lightweight encoder–decoder network for automatic pavement crack detection. *Computer-Aided Civil and Infrastructure Engineering*, *39*(12), 1743–1765.

Zong, Z., Song, G., & Liu, Y. (2023). DETRs with collaborative hybrid assignments training. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 6748–6758). IEEE.

Zotin, A. G. (2020). Fast algorithm of image enhancement based on multi-scale Retinex. *International Journal of Reasoning-Based Intelligent Systems*, *12*(2), 106–116.

How to cite this article: Hu, Y., Zhang, H., Hou, Y., & Liu, P. (2025). Advanced low-light image transformation for accurate nighttime pavement distress detection. *Computer-Aided Civil and Infrastructure Engineering*, 1–18. https://doi.org/10.1111/mice.70001