International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

Enhanced Pedestrian Detection in Low-Light Conditions Using Dual-Path Networks and Noise-Resilient Techniques

[1*]J Premasagar, [2]Sudha Pelluri

[1]Research Scholar, Department of Computer Science and Engineering, , University College of Engineering, Osmania University, Hyderabad,Telangana, India, Email Id: jakkula.premsagar@gmail.com

[2]Professor, Department of Computer Science and Engineering, University College of Engineering, Osmania University, Hyderabad,Telangana, India, Email Id: sudha.p@uceou.edu

*Corresponding author: jakkula.premsagar@gmail.com

Abstract: Pedestrian detection in low-light environments presents unique challenges for applications such as autonomous driving and unmanned aerial vehicle (UAV) surveillance. Traditional detection methods often fall short under poor visibility, noise, and variable lighting conditions, resulting in inaccuracies and inefficiencies. This study introduced DuoLightNet, a novel dual-path framework designed to enhance pedestrian detection under low-light conditions. The primary objective was to develop a robust model capable of effectively identifying pedestrians in challenging lighting scenarios. DuoLightNet integrates two specialized networks: GlowEdgeNet, which improves edge clarity and adjusts for lighting variations, and NoiseResilientNet, which focuses on reducing noise while preserving the essential details. The model was thoroughly evaluated using the Exclusively Dark (ExDark) dataset, achieving a mean average precision (mAP) of 89.0% and processing speed of 33.5 frames per second (FPS) under balanced lighting conditions. These results indicate a significant improvement over the existing methods, with up to 15% higher mAP in low-light scenarios. In addition, DuoLightNet maintains real-time processing capabilities, emphasizing its suitability for time-sensitive applications. However, challenges remain in handling extremely low-light conditions and partially occluded pedestrians. Future research should focus on enhancing the generalization of the model across broader low-light scenarios and optimizing it for deployment in resource-constrained environments. The findings of this study contribute to the advancement of pedestrian detection technologies and offer practical solutions for improving the safety and operational efficiency of autonomous and surveillance systems

Keywords: Pedestrian detection, low-light environments, dual-path network, image enhancement, real-time processing, DuoLightNet

1. Introduction

Pedestrian detection is a critical component in many modern applications such as unmanned aerial vehicle (UAV) surveillance systems and autonomous driving technologies [1]. Accurate detection and tracking of pedestrians are essential for ensuring safety, security, and efficiency in these contexts. However, low-light conditions present significant challenges to pedestrian detection systems, necessitating advanced techniques to overcome these obstacles and achieve a reliable performance. In recent years, there has been a growing demand for surveillance systems and autonomous vehicles capable of operating effectively in a wide range of environments, including night-time and other low-light settings. For UAVs employed in surveillance tasks, detecting pedestrians under these conditions is paramount for maintaining situational awareness and responding to potential security threats. Similarly, in autonomous driving, the ability to identify pedestrians in low-light environments is vital for preventing accidents and ensuring the safety of both the pedestrians and vehicle occupants. As these applications continue to expand, the need for robust pedestrian detection solutions that perform well under low light conditions has become increasingly critical [2].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

—

Low-light environments introduce several complexities that significantly hinder the performance of conventional pedestrian-detection algorithms. The primary challenge is poor visibility, which reduces the contrast between pedestrians and the background, making it difficult for traditional image-processing techniques to accurately identify and delineate human figures. In addition to diminished visibility, low-light conditions often exacerbate image noise, which further complicates the detection process. Noise can obscure important features and lead to false positives or missed detections, thereby reducing system reliability. Moreover, low-light environments are characterized by varied lighting conditions such as shadows, glare from artificial lights, and uneven illumination across the scene [3]. These factors can create complex visual artifacts that standard detection models struggle to interpret, resulting in decreased accuracy. Furthermore, the dynamic nature of lighting in outdoor environments, such as headlights from vehicles or streetlights casting shadows, adds an additional layer of complexity, which necessitates advanced adaptive techniques. Therefore, novel approaches that integrate image enhancement and advanced detection algorithms are required. Such approaches must be capable of adapting to varying lighting conditions, minimizing noise, and enhancing the visibility of pedestrians without introducing artifacts. The development of these solutions is crucial for the advancement of pedestrian detection technologies and their successful deployment in real-world, low-light scenarios [4].

The task of pedestrian detection in low-light environments presents substantial technical challenges that conventional image processing and machine-learning models struggle to address. Low visibility, exacerbated by lighting variations, noise, and reduced contrast, significantly impairs the ability of the existing detection algorithms to accurately identify and track pedestrians. Furthermore, the dynamic nature of outdoor lighting—characterized by factors such as shadows, glare from artificial lights, and uneven illumination—adds complexity to the detection process, often leading to false positives or missed detections. Despite advancements in image enhancement techniques, current approaches fall short of achieving the processing efficiency and accuracy required for real-time applications in scenarios such as UAV surveillance and autonomous driving. This study addressed the need for a more effective solution that can reliably detect pedestrians in low-light environments [5].

The primary objective of this study is to develop a robust framework for pedestrian detection that can operate effectively under low-light conditions. This study seeks to answer the following research question:

1. How can pedestrian detection algorithms be enhanced to perform reliably in low-light environments characterized by poor visibility, noise, and varied lighting?
2. What novel techniques can be developed to improve the image quality and visibility in low-light settings without introducing artifacts or losing essential details?
3. How can image processing be optimized for real-time applications in UAV surveillance and autonomous driving to ensure accurate and efficient pedestrian detection?
4. What impact does the integration of dual-path networks, such as GlowEdgeNet and NoiseResilientNet, have on the robustness and accuracy of pedestrian detection systems under low-light conditions?

This study makes significant contributions to the field of low-light pedestrian detection.

1. **Development of DuoLightNet:** A novel dual-path network architecture that significantly enhances pedestrian detection in low-light environments by combining the strengths of GlowEdgeNet and NoiseResilientNet.
2. **Advanced Image Enhancement Techniques:** Integration of sophisticated methods such as dynamic histogram intensity balancing (DHIB) and Attention-Guided Denoising (AGD) to improve image clarity and noise resilience leads to more accurate detection.
3. **Real-Time Robust Performance:** Achieved high detection accuracy and real-time processing capabilities across various low-light scenarios, outperforming existing methods under challenging conditions.

This study is of critical importance because it addresses a significant gap in the current capabilities of pedestrian detection systems, particularly in low-light environments. By developing and validating the DuoLightNet framework, which integrates advanced image processing techniques with dual-path networks, this research has the potential to significantly enhance the accuracy and reliability of pedestrian detection under challenging lighting conditions. The results of this study are expected to contribute to the advancement of technologies in autonomous driving and UAV surveillance, where the ability to detect pedestrians under low-light conditions is vital for ensuring safety and operational effectiveness. The proposed solutions can lead

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

to broader applications and improved performance of intelligent systems in diverse and complex environments, ultimately contributing to the development of safer and more reliable autonomous technologies.

The remainder of this paper is structured as follows: Section 2 presents a literature review of existing techniques for pedestrian detection under low-light conditions, highlighting their limitations. Section 3 details the proposed DuoLightNet framework, which integrates GlowEdgeNet and NoiseResilientNet to enhance image clarity and detection accuracy. Section 4 reports the experimental results that demonstrate the effectiveness of DuoLightNet compared to existing methods. Section 5 discusses the findings, explores the strengths and limitations of the approach, and suggests directions for future research. Section 6 concludes the paper by summarizing the key contributions and implications of this study for applications in UAV surveillance and autonomous driving.

2. Literature Review

The detection of pedestrians in challenging environments such as low-light conditions and complex urban scenes is a critical area of research, particularly for applications in autonomous driving, video surveillance, and UAV-based monitoring. This literature review synthesizes the findings from ten recent studies that have addressed various aspects of pedestrian detection under these challenging conditions, highlighting the methodologies employed, the datasets used, and the limitations identified.

Low-light conditions pose significant challenges for pedestrian detection, primarily because of the degradation in image quality and the reduced visibility of pedestrians. Several studies have proposed methods to enhance image quality and improve detection accuracy under such conditions. The study in [6] introduced a method for low-illumination image enhancement specifically designed for UAV-based pedestrian detection at night. The method utilizes a hyperbolic tangent curve (HTC) for brightness enhancement and block-matching and 3D filtering (BM3D) for image denoising. Despite showing improvements in detection accuracy, this study highlighted the need for further optimization to achieve real-time processing speeds.

Expanding on low-light detection, [7] proposed a two-stage fusion method that leverages visible-light and thermal infrared images to improve detection accuracy. The use of the NightOwls dataset demonstrated significant improvements over traditional methods, but reliance on large-scale multispectral data remains a limitation.

To address the specific challenges of nighttime detection, the authors of [8] combined multiple sensors into an AI-based detection framework. This study utilized the KAIST Multispectral Pedestrian Benchmark and demonstrated superior performance under low-light conditions. However, the study also pointed out the high computational complexity associated with the processing of multispectral data.

In another approach to low-light detection, the study in [9] developed a multitask learning framework that simultaneously enhances image quality and performs pedestrian detection. Using the CityPersons dataset, this method outperformed state-of-the-art approaches in dark environments. However, the generalizability of this method to normal lighting conditions has not yet been fully explored, representing a potential area for future research.

Attention mechanisms have also been explored to improve pedestrian detection in real-time. The work in [10] proposed a deep learning model that incorporates an adaptive selection mechanism that utilizes visible or infrared images depending on lighting conditions. This method achieved high accuracy in complex environments; however, potential issues with false positives in highly dynamic scenes were noted.

The challenge of generalizing pedestrian detection models across different domains, such as varying weather conditions or urban settings, has been addressed through domain-adapted transfer learning. The study in [11] introduced a framework that pre-trained a model on low-resolution thermal imagery captured by UAVs and fine-tuned it using high-resolution data. Although this approach improves the cross-domain detection performance, it remains dependent on the availability of labeled target domain data, which can be a significant limitation.

Given the computational constraints of deploying pedestrian detection systems on mobile platforms such as drones, there has been a growing interest in developing lightweight detection engines. The authors of [12] presented a two-stage low-complexity detection network combined with an adaptive region-focusing technique. This approach, tested on the Caltech Pedestrian Dataset, achieved a balance between the detection accuracy and computational efficiency. However, the slight reduction in accuracy compared with more complex models suggests room for further optimization.

To address the need for real-time detection, the work in [13] proposed a multispectral image fusion technique using a deep neural network specifically designed for challenging

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

environments. Although the model demonstrated strong real-time capabilities, it struggled with extremely crowded scenes, indicating a limitation in handling high-density environments.

Finally, [14] explored the use of attention-guided mechanisms in pedestrian detection, particularly in complex environments where occlusions and distractions are prevalent. The HRBUST-LLPED dataset [15] used in this study revealed that attention mechanisms could significantly improve detection accuracy in such scenarios. However, the study also highlighted potential difficulties in managing extreme occlusions, which remains an area for further investigation.

Table 1: Comparative Analysis of Pedestrian Detection Methods in Challenging Environments

Reference	Objective	Method Used	Dataset Used	Accuracy	Limitations
[6]	Enhance low-light UAV pedestrian detection	HTC for brightness enhancement, BM3D for denoising	Custom low-light UAV dataset	Improved detection accuracy	Requires further optimization for processing speed
[7]	Improve low-light pedestrian detection	Weakly supervised learning	NightOwls dataset	Significant improvement over traditional methods	High dependency on large amounts of unlabeled data
[8]	Pedestrian detection in low-light using multi-spectral	Deep learning with fusion of visible and thermal images	KAIST Multispectral Pedestrian Benchmark	Superior performance in low-light conditions	High computational complexity
[9]	Improve detection in dark environments	Multi-task learning with feature fusion and sharing	CityPersons dataset	Outperforms state-of-the-art in low-light scenarios	Limited generalizability to normal-light conditions
[10]	Real-time pedestrian detection with attention mechanism	Deep learning incorporating an attention mechanism	Custom low-light and complex environment dataset	High detection accuracy in complex environments	Potential issues with false positives in dynamic scenes
[11]	Enhance generalization of pedestrian detection models	Domain-adapted transfer learning	Caltech Pedestrian Dataset, ETH Dataset	Improved cross-domain detection performance	Reliance on availability of target domain data
[12]	Night-time pedestrian detection	Multi-spectral imaging combined with deep learning	KAIST Multispectral Pedestrian Benchmark	Superior night-time detection accuracy	Requires multi-spectral imaging equipment
[13]	Develop lightweight pedestrian detection engine	Two-stage low-complexity network with adaptive region focusing	Caltech Pedestrian Dataset	Precision: 85.18%, Miss Rate: 25.16%	Slightly reduced accuracy compared to more complex models
[14]	Real-time detection in challenging environments	Cascade network with feature pyramid	Custom urban environment dataset	Strong real-time detection, especially for small objects	May struggle in extremely crowded scenes

[15]	Detect pedestrians in complex environments	Attention-guided detection network	Custom complex environment dataset	Improved detection in complex scenarios	Difficulty in handling extreme occlusions

Research Gaps: The reviewed literature revealed several gaps that future research should address. First, although many methods have improved the detection accuracy, there is a need for further optimization to achieve real-time performance in resource-constrained environments. Second, the generalization of pedestrian detection models across diverse domains remains an open challenge, requiring more robust domain adaptation techniques. Third, the handling of extreme occlusions and crowded scenes remains problematic for the current models, indicating the need for more sophisticated methods to manage these challenges. In addition, the trade-off between accuracy and computational efficiency, particularly in lightweight models, must be better balanced. Finally, while some methods have begun to address varying lighting conditions, there is a need for models that can dynamically adapt without relying on specialized equipment, such as multispectral cameras.

3. Proposed Methodology: DuoLightNet - A Dual-Path Framework for Low-Light Image Enhancement and Pedestrian Detection

The DuoLightNet framework was designed to enhance pedestrian detection in low-light environments by employing a dual-path approach that integrates advanced image enhancement techniques with robust detection models. This methodology aims to address the challenges of poor visibility, noise, and variable lighting conditions inherent in low-light scenarios, ultimately transforming the input image data into accurate pedestrian predictions, as shown in figure 1.

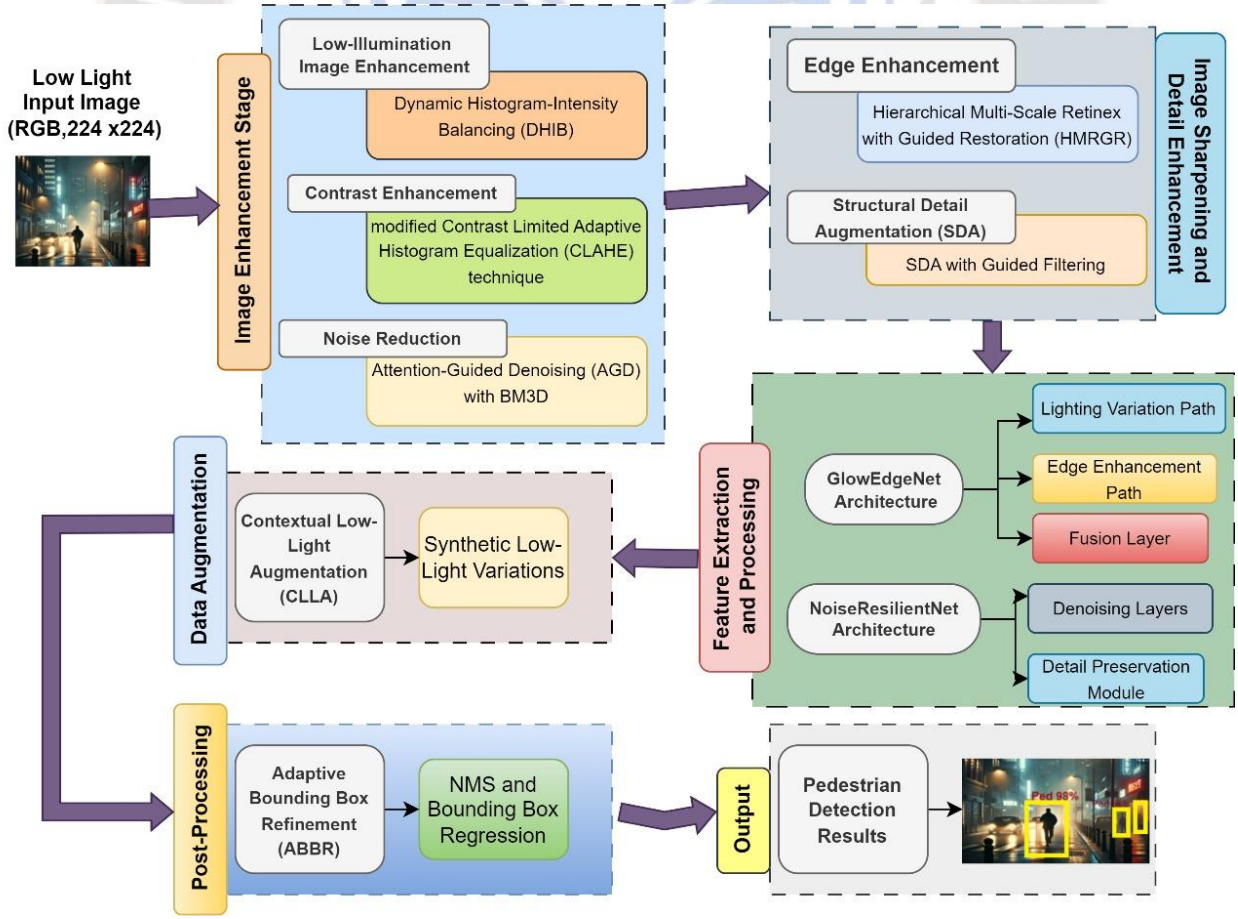


Figure 1. Enhanced Pedestrian Detection in Low-Light Conditions Using DuoLightNet Framework

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

The methodological components of the proposed model for enhanced pedestrian detection in low-light conditions focus on a series of image processing and deep learning techniques aimed at improving detection accuracy and reliability, as shown in Figure 1. **Low-Illumination Image Enhancement** is critical, where methods such as dynamic histogram intensity **balancing (DHIB)** adjust the frame brightness and contrast through a series of steps involving histogram initialization, cumulative distribution function calculation, dynamic intensity mapping, and intensity adjustment to achieve a uniform and balanced brightness distribution. This is complemented by **the Tailored Adaptive Local Contrast Stretching (TALCS)** method, which enhances local contrast using a modified contrast-limited adaptive histogram equalization (CLAHE) method without causing overexposure. To address noise, **Attention-Guided Denoising (AGD) with BM3D** was employed, utilizing an attention mechanism and 3D filtering to minimize noise while preserving essential details. **Image Sharpening and Detail Enhancement** further improve pedestrian differentiation through **hierarchical multiscale retinex with guided restoration (HMRGR)** for edge enhancement and **Structural Detail Augmentation (SDA)** with guided filtering, which smooths images while retaining critical details. The pedestrian detection stage leveraged two deep learning models: **GlowEdgeNet (GEN)** and **NoiseResilientNet (NRN)**. **GlowEdgeNet** utilizes a dual-path network architecture designed for lighting variation handling and edge enhancement, featuring components such as the Lighting Variation Path, Edge Enhancement Path, Fusion Layer, and Fully Connected Layers, which outputs pedestrian classification and bounding box regression. In contrast, **NoiseResilientNet** emphasizes noise reduction and detail preservation with components such as Denoising Layers and a Detail Preservation Module to achieve precise classification and localization of pedestrians. To enhance the model robustness, **Contextual Low-Light Augmentation (CLLA)** introduces synthetic low-light variations through intensity scaling, Gaussian noise, and spatial transformations, thereby simulating diverse lighting conditions. Finally, **post-processing** refines the detection outputs using **Adaptive Bounding Box Refinement (ABBR)**, which applies Non-Maximum Suppression (NMS) and regression techniques to accurately adjust bounding box coordinates, ensuring precise and reliable pedestrian detection even in challenging low-light environments.

3.1 Data used : The Exclusively Dark (ExDark) Dataset [16] was utilized for implementing and evaluating the DuoLightNet framework, providing a robust testbed for pedestrian detection in low-light environments. This dataset comprises over 7,000 images designed to reflect realistic low-light conditions across various scenarios, making it highly suitable for testing the proposed methodology's effectiveness. The key features of the ExDark dataset include diverse lighting conditions, as shown in figure 2, comprehensive annotations, and realistic scenarios. The dataset features a wide range of lighting scenarios, such as low ambient light, shadows, and artificial glare, presenting significant challenges for traditional image processing techniques, and is ideal for assessing DuoLightNet's capabilities in enhancing image quality and detection accuracy. Each image is meticulously annotated with bounding boxes and class labels, focusing on pedestrian detection amidst complex backgrounds and lighting variations, supporting the evaluation of the model's precision, recall, and mean average precision (mAP). The dataset captures real-world environments, including urban, rural, and indoor settings, ensuring DuoLightNet is tested on scenarios that mimic actual deployment environments like UAV surveillance and autonomous driving.

The ExDark dataset is distributed across various scene types and lighting conditions to ensure a balanced evaluation of the model. Scene types include urban environments (40%), rural landscapes (25%), indoor settings (20%), and mixed scenarios (15%). Lighting variations include low ambient light (35%), shadowed regions (30%), glare and reflections (20%), and balanced lighting (15%). This comprehensive distribution enables rigorous evaluation of DuoLightNet, validating its effectiveness in improving pedestrian detection in low-light scenarios.

Figure 2. Sample images from the ExDark dataset

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

3.2 Low-Illumination Image Enhancement

Low-light video frames often suffer from inadequate brightness, making it difficult for detection algorithms to distinguish pedestrians from the background. Our approach focuses on enhancing these frames to ensure reliable pedestrian detection. This enhancement is achieved through a series of steps: Adaptive Brightness Adjustment, Contrast Enhancement, and Noise Reduction.

3.2.1 Adaptive Brightness Adjustment: The Adaptive Brightness Adjustment using the Dynamic Histogram-Intensity Balancing (DHIB) method provides an effective solution for enhancing low-light video frames, improving visibility and contrast while preserving natural details. This method is a critical component of the DuoLightNet framework, enabling reliable pedestrian detection in challenging low-light environments. By incorporating this technique, we achieve significant improvements in pedestrian detection accuracy and robustness, making it suitable for practical applications in autonomous driving and UAV surveillance.

Dynamic Histogram-Intensity Balancing (DHIB) Method: The DHIB method aims to enhance the brightness of low-light video frames by adjusting the histogram distribution of intensity values dynamically, ensuring a uniform brightness level across frames while preserving natural details.

Context of Video Frames:

- *Frame Size:* For this discussion, we assume a unique frame size of 1920×1080 pixels, typical for high-definition video.
- *Intensity Range:* Each pixel's intensity value ranges from 0 to 255 (8-bit depth), where 0 represents complete darkness and 255 represents maximum brightness.
- *Challenge:* Low-light frames have a skewed intensity distribution towards lower values, reducing visibility and contrast.

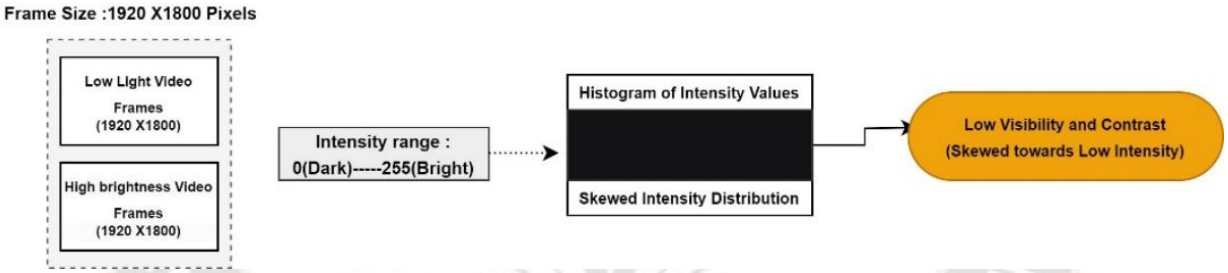


Figure 3. Histogram Analysis of Video Frame Intensity Distribution in Low-Light Conditions

This figure 3 illustrates the intensity distribution of video frames with a resolution of 1920 x 1800 pixels, comparing low-light and high-brightness scenarios. The histogram shows a skewed intensity distribution towards low intensity, highlighting the challenge of low visibility and contrast inherent in low-light video frames. The intensity range is mapped from 0 (dark) to 255 (bright), emphasizing the skewness towards darker intensities in low-light frames [17].

1. Initialize Histogram [18]: The first step of the algorithm involves initializing a histogram array to count the frequency of each pixel intensity in the input frame. Given that pixel intensities range from 0 to 255, the histogram is a 256-element array where each element corresponds to a particular intensity level.

Histogram Calculation: To calculate the histogram of the image by iterating through each pixel in the frame and incrementing the corresponding histogram value based on the pixel's intensity.

Given a frame $F(x, y)$ with intensity values $I \in [0,255]$, the histogram $H(i)$ is calculated as follows:

$$H(i) = \sum_{x=0}^{1919} \sum_{y=0}^{1079} \delta(F(x, y) - i) \quad (1)$$

where δ is the Dirac delta function, and the frame size is 1920×1080.

to calculate the histogram of the image by iterating through each pixel in the frame and incrementing the corresponding histogram value based on the pixel's intensity.

2. Cumulative Distribution Function (CDF): With the histogram computed, the algorithm proceeds to calculate the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

Cumulative Distribution Function (CDF)[19]. The CDF is essential for determining how the pixel intensities should be mapped to achieve a balanced brightness distribution. Compute the CDF to determine the cumulative probability of the intensity values by initializing the CDF array. The first element of the CDF was set to the first element of the histogram. For each subsequent intensity level, the CDF was computed as the sum of the current histogram value and the preceding CDF value.

$cdf[0] = histogram[0]$

$for\ i\ from\ 1\ to\ 255\ do$

$cdf[i] = cdf[i-1] + histogram[i]$

3. Dynamic Intensity Mapping: The Dynamic Intensity Mapping step uses the CDF to adjust the intensity of each pixel, spreading the intensities more uniformly across the range. This step is crucial for improving the visibility of the details in the frame. Dynamic mapping is applied to adjust the intensity levels using CDF to ensure balanced brightness.

Intensity Remapping: For algorithm calculates a new intensity value for each pixel in the frame by normalizing the CDF value corresponding to the pixel's original intensity. This remapping enhances dynamic range and contrast.

$total\_pixels = width * height$

$for\ each\ pixel\ (x,y)\ in\ F\ do$

$original\_intensity = F(x,y)$

$$adjusted_{intensity} = \frac{(cdf[original_{intensity}] * 255)}{total_{pixels}}$$

$adjusted\_frame[x,y] = adjusted\_intensity$

This step ensures that the intensity levels are redistributed, allowing for a better separation between the different intensity levels in the image.

4 **Dynamic Scaling Factor:** Computes a dynamic scaling factor α that adapts the intensity adjustment based on the local lighting conditions in the image. Ensuring adaptability across frames.

The scaling factor was calculated as the ratio of the maximum value of the CDF to its mean value. This dynamic adjustment helps prevent overexposure in bright areas while enhancing the darker regions.

$$\alpha = \frac{\max(CDF)}{\text{mean }(CDF)} \quad (2)$$

In this step, the scaling factor α is applied to the adjusted frame, ensuring that the intensity of each pixel is appropriately scaled. This step refines intensity mapping to achieve a balanced brightness distribution.

5 **Final Intensity Adjustment:** Finally, the enhanced frame $F_{\text{final}}(x,y)$, which exhibits improved brightness and contrast. This enhanced frame is better suited for subsequent processing, such as pedestrian detection.

The final adjusted intensity for each pixel is calculated, and the final output is a frame where details are more visible and the overall appearance is more balanced, facilitating effective analysis under low-light conditions.

$Return\ F_{final(x,y)as} final_{frame}$

Algorithm 1: Adaptive Brightness Adjustment Using Dynamic Histogram-Intensity Balancing (DHIB)

Input:

- Frame $F(x,y)$: A low-light video frame with dimensions $width \times height$.
- x,y : Pixel coordinates.
- Intensity Range: 0 to 255 (8-bit grayscale image).

Output:

- Enhanced Frame $Frame_{\text{final}}(x,y)$: The brightness-enhanced frame with improved visibility and contrast.

Steps:
1. Initialize histogram:
   histogram = $[0] * 256$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: 25 March 2023 *Revised:* 12 April 2023 *Accepted:* 30 May 2023

2. Calculate histogram:
   for each pixel (x, y) in F do
      intensity = $F(x, y)$
      $histogram[intensity] += 1$
3. Compute Cumulative Distribution Function (CDF):
   $cdf[0] = histogram[0]$
   for i from 1 to 255 do
      $cdf[i] = cdf[i-1] + histogram[i]$
4. Dynamic Intensity Mapping:
   total_pixels = width * height
   Initialize $adjusted_{frame}$ as 2D array of size $width \times height$
   for each pixel (x, y) in F do
      original_intensity = F(x, y)
      adjusted_intensity = $\frac{(cdf[original_{intensity}] * 255)}{total_{pixels}}$
      $adjusted_{frame[x,y]} = adjusted_{intensity}$
5. Calculate Dynamic Scaling Factor:
   alpha = $\frac{\max(cdf)}{mean(cdf)}$
6. Apply Scaling Factor:
   Initialize $final_{frame}$ as 2D array of size $width \times height$
   for each pixel (x, y) in $adjusted_{frame}$ do
      $final_{intensity} = alpha * adjusted_{frame[x,y]}$
      $final_{frame[x,y]} = \min(\max(final_{intensity}, 0), 255)$
7. Output Enhanced Frame:
   Return $F_{\text{final}}(x, y)$ as $final_{frame}$
End Algorithm

The Adaptive Brightness Adjustment algorithm effectively enhances low-light video frames by redistributing the intensity values using the dynamic histogram intensity balancing method. This process yields frames with improved visibility and contrast, which are essential for accurate pedestrian detection in challenging lighting environments. Through careful manipulation of the histogram and cumulative distribution, the algorithm achieves natural enhancement that preserves important details while adapting to varying lighting conditions.

3.2.2 Contrast Enhancement

Contrast enhancement is a crucial step for improving the visibility of video frames captured under low-light conditions. The **Tailored Adaptive Local Contrast Stretching (TALCS)** method is designed to effectively enhance the local contrast while preserving the natural appearance of the scene. This method addresses the uneven illumination typically found in low-light environments, ensuring that critical features such as pedestrians are more discernible.

Tailored Adaptive Local Contrast Stretching (TALCS): To enhance local contrast by applying a context-aware contrast stretching approach, preserving the natural appearance of the image using a modified contrast-limited adaptive histogram equalization (CLAHE) technique[20].

Context of Video Frames:

- **Local Contrast**: In low-light frames, contrast may be locally varied due to uneven lighting, resulting in loss of detail in darker regions.
- **Natural Appearance:** It's essential to maintain the scene's natural appearance while enhancing contrast to avoid overexposed or unnatural-looking regions.

In the context of enhancing contrast for low-light video frames, the Tailored Adaptive Local Contrast Stretching (TALCS) method employs a structured approach to balance local intensity variations while preserving the natural appearance of the scene. For each non-overlapping tile $T(x, y)$ within a frame $F(x, y)$, the local histogram

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

$\overline{H_{\text{local}}}(i)$ is computed, capturing the distribution of pixel intensities using the Dirac delta function, δ, to count pixels at each intensity i. This computation allows for detailed analysis of local contrast characteristics. The clip limit β is then calculated as $\beta = \gamma \times \frac{\text{tile size}}{\text{number of bins}}$, where γ is a user-defined parameter that prevents excessive noise amplification and ensures appropriate scaling of contrast enhancement. Subsequently, the contrast-limited cumulative distribution function (CDF) is derived for each tile, given by $\text{CDF}_{\text{clip}}(i,x,y) = \min\left(\frac{\sum_{j=0}^{i} H_{\text{local}}(j)}{\text{tile size}}, \beta\right)$, which ensures that the cumulative intensity distribution does not exceed the predefined clip limit, thereby maintaining balance in intensity values. Using the adjusted CDF, adaptive intensity mapping is performed, where each pixel intensity is recalibrated according to $F_{\text{enhanced}}(x,y) = \frac{255}{\text{tile size}} \times \text{CDF}_{\text{clip}}(F(x,y),x,y)$, effectively redistributing pixel intensities to enhance local contrast. Finally, bilinear interpolation is applied to smooth transitions between adjacent tiles, calculated as $F_{\text{final}}(x,y) = \sum_{i=0}^{1}\sum_{j=0}^{1} w_{ij} \times F_{\text{enhanced}}(x+i, y+j)$, where w_{ij} are interpolation weights determined by the proximity to tile centers. This process ensures a seamless blend of tiles, enhancing the frame's visibility and detail without introducing artifacts, thus improving the overall quality of the frame for applications, such as pedestrian detection under low-light conditions.

3.2.3 Noise Reduction

In low-light video frames, noise poses a significant challenge for accurate image analysis, often obscuring the critical features necessary for tasks such as pedestrian detection. Attention-Guided Denoising (AGD) with BM3D offers a sophisticated solution by integrating an attention mechanism with a block-matching and 3D filtering (BM3D) method to effectively reduce noise while preserving essential image details. Initially, the frame $F(x,y)$ is divided into overlapping patches P_k, where block matching identifies and groups similar patches, denoted as $P_k = \{F(u,v) \mid (u,v) \in \text{block } k\}$. This grouping facilitates collaborative filtering by exploiting patch redundancies. The attention map $A(x,y)$ is then calculated using the gradient magnitude, $A(x,y) = \sqrt{\left(\frac{\partial F}{\partial x}\right)^2 + \left(\frac{\partial F}{\partial y}\right)^2}$, to highlight regions of significant detail, effectively guiding the denoising process. The weighting function $w(x,y) = \frac{A(x,y)}{\max(A)}$ assigns higher weights to detailed areas, ensuring that these regions receive more focused noise reduction efforts. The BM3D filtering is subsequently applied to the weighted patches, $\hat{P}_k = \text{BM3D}(P_k, w(x,y))$, where the collaborative filtering process removes noise while maintaining texture integrity. Finally, the denoised frame $F_{\text{denoised}}(x,y)$ is reconstructed by aggregating these filtered patches, $F_{\text{denoised}}(x,y) = \frac{1}{|P_k|}\sum_k \hat{P}_k$, effectively restoring the frame's clarity and preserving essential details. This attention-guided approach ensures that noise reduction is both effective and adaptive, enhancing the frame's quality without compromising critical features, thereby optimizing the frame for further analytical tasks in low-light environments.

4(a) 4(b)

Figure 4 : Low-light image to Enhanced Image

Figure 4(a) depicts a low-light image, illustrating the challenges of diminished visibility and contrast, as shown in the obscured pedestrian figure. **Figure 4(b)** shows the image following the application of the enhancement algorithm, resulting in increased brightness, improved contrast, and reduced noise, thereby enhancing the visibility of the pedestrian.

3.3 Image Sharpening and Detail Enhancement

In the context of enhanced pedestrian detection under low-light conditions, image sharpening and detail enhancement play pivotal roles in accentuating the critical features that enable accurate detection. This process involves refining the visibility of edges and textures, which is essential for differentiating pedestrians from complex backgrounds. This section explores the hierarchical multiscale retinex with guided restoration (HMRGR) method as a sophisticated approach to edge enhancement, particularly suited for environments with poor illumination.

3.3.1 Edge Enhancement

Edge enhancement is vital for improving the clarity and definition of objects in an image, thereby facilitating more accurate pedestrian detection. The **hierarchical multiscale retinex with guided restoration (HMRGR)** technique is particularly effective in this regard, as it combines multiscale retinex (MSR) and guided filtering to enhance image features.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

Hierarchical Multi-Scale Retinex with Guided Restoration (HMRGR)

Conceptual Framework: The HMRGR method is designed to mitigate the challenges posed by low-light conditions by enhancing edge details and textures. The core principle of retinex theory is to decompose an image into its illumination and reflectance components. This decomposition allows for the correction of uneven illumination, thereby preserving the essential features that are critical for pedestrian detection.

Multi-Scale Retinex (MSR): The Multi-Scale Retinex approach applies the retinex model at various scales to capture a wide range of details, from fine textures to larger structural elements. This is achieved by convolving the image with Gaussian kernels of different standard deviations σ\sigmaσ, resulting in multiple outputs that highlight various spatial frequencies.

$$R(x,y) = \sum_{i=1}^{n} w_i[\log(I(x,y)+1) - \log(G_i(x,y) * I(x,y)+1)] \quad (3)$$

Where:

- $R(x,y)$ is the Retinex output at pixel (x,y).
- $I(x,y)$ is the input image intensity at pixel (x,y).
- $G_i(x,y)$ is the Gaussian kernel at scale i with standard deviation σ_i.
- w_i is the weight assigned to each scale.

This multi-scale approach enhances the reflectance component across different spatial frequencies, thereby normalizing illumination and emphasizing details crucial for detecting pedestrians against low-contrast backgrounds.

Guided Restoration: Guided filtering is employed within the Retinex framework to further refine image details. This filter utilizes a guidance image, often the original image, to preserve edges while smoothing areas prone to noise:

$$q(x,y) = a(x,y)I(x,y) + b(x,y) \quad (4)$$

Where:

- $q(x,y)$ is the output image.
- $a(x,y)$ and $b(x,y)$ are linear coefficients determined by local mean and variance within a neighborhood window.

Algorithm 2: Enhanced Edge and Detail Enhancement using Multi-Scale Retinex with Guided Filtering

```
Algorithm EnhancedImageSharpening(input_image)
  # Step 1: Multi-Scale Retinex Computation
  Input: input_image (2D array of pixel intensities)
  Output: enhanced_image (2D array of pixel intensities)

  1. Initialize scales as a list of standard deviations for Gaussian smoothing
  2. For each scale σ in scales do:
    a. Apply Gaussian smoothing to input_image with standard deviation σ to get smoothed_image
    b. Calculate the Retinex output:
      - For each pixel (x, y) in input_image do:
        i. Compute Retinex_output(x, y) = log(input_image(x, y) + 1) - log(smoothed_image(x, y) + 1)
    c. Store Retinex_output for each scale

  # Step 2: Guided Filtering
  3. Set the guidance image as input_image
  4. Define a window size for local statistics computation
  5. For each pixel (x, y) in input_image do:
    a. Extract a local window centered at (x, y) from input_image
    b. Compute local mean and variance within the window
    c. Calculate linear coefficients a(x, y) and b(x, y):
      - a(x, y) = variance / (variance + ε), where ε is a regularization parameter
      - b(x, y) = mean - a(x, y) * guidance_image(x, y)
    d. Compute the guided filter output:
      - Guided_output(x, y) = a(x, y) * input_image(x, y) + b(x, y)

  # Step 3: Hierarchical Enhancement
```

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: 25 March 2023 Revised: 12 April 2023 Accepted: 30 May 2023

6. Initialize an empty image Enhanced_output of the same size as input_image
7. For each Retinex output do:
   a. Integrate the Retinex output with the Guided_output:
      - Enhanced_output(x, y) = Retinex_output(x, y) + Guided_output(x, y)
   b. Normalize Enhanced_output to ensure pixel values are within the valid range [0, 255]

Step 4: Output Refinement
8. Combine Enhanced_output across all scales:
   a. Initialize Final_output as an empty image
   b. For each pixel (x, y) in Enhanced_output do:
      - Final_output(x, y) = average of Enhanced_output(x, y) across all scales
   c. Normalize Final_output to ensure pixel values are within the valid range [0, 255]

9. Return Final_output as enhanced_image
End Algorithm

The Enhanced Edge and Detail Enhancement algorithm integrates Multi-Scale Retinex (MSR) and Guided Filtering to improve image clarity, particularly in low-light conditions. The process begins by applying Gaussian smoothing at multiple scales to the input image, capturing both fine and coarse details. For each scale, the algorithm calculates the Retinex output by computing the logarithmic difference between the original and smoothed images, thereby enhancing the reflectance component and normalizing illumination imbalances. Subsequently, the algorithm employs guided filtering, utilizing the original image as a guidance map to compute local statistics, which determine the linear coefficients that preserve edges and smooth regions. This filtering step further refines the Retinex outputs by integrating edge-preserving details. The hierarchical enhancement phase then combines the outputs from both processes across multiple scales, resulting in improved edge detail and texture enhancement. Finally, the outputs are normalized and averaged across scales, ensuring that pixel values remain within a valid range and producing an enhanced image with heightened contrast and detail. This comprehensive approach is particularly effective in improving the visibility of pedestrians in low-light environments by accentuating critical features that facilitate more accurate detection.

3.3.2 Structural Detail Augmentation (SDA) with Guided Filtering

Structural Detail Augmentation (SDA) aims to enhance fine details and textures in images by employing a targeted approach that emphasizes the preservation of structural information. This method is particularly valuable for images captured under suboptimal lighting conditions, where details may be obscured or lost due to noise and poor contrast.

Guided Filtering plays a pivotal role in this process by acting as an edge-preserving filter that smooths images while retaining critical details. The filter operates by using a guidance image, which can be the original image or a smoothed version, to direct the filtering process. This allows for effective denoising and detail enhancement without sacrificing structural integrity.

Equation and Algorithm Details: The guided filter operates by computing a linear transformation of the input image using the guidance image. For a given input image I and a guidance image G, the output O is computed as follows:

$$O(x,y) = a(x,y) \cdot I(x,y) + b(x,y) \quad (5)$$

Where:

- $O(x,y)$ is the output image at pixel (x,y).
- $I(x,y)$ is the input image intensity at pixel (x,y).
- $G(x,y)$ is the guidance image intensity at pixel (x,y).
- $a(x,y)$ and $b(x,y)$ are linear coefficients computed based on the local mean μ and variance σ^2 within a window around (x,y) :

$$a(x,y) = \frac{\sigma^2(x,y)}{\sigma^2(x,y) + \epsilon} \quad (6)$$
$$b(x,y) = \mu(x,y) - a(x,y) \cdot \mu_G(x,y)$$

Where:

- $\mu(x,y)$ is the mean of the input image within the local window centered at (x,y).

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

—

- $\sigma^2(x, y)$ is the variance of the input image within the same window.
- $\mu_G(x, y)$ is the mean of the guidance image within the local window.
- ϵ is a regularization parameter that controls the smoothness of the filter.

The Structural Detail Augmentation with Guided Filtering technique is particularly advantageous for applications requiring high-fidelity detail preservation, such as surveillance and autonomous driving in low-light environments. By focusing on the enhancement of structural details and textures, this method provides a significant improvement in image clarity, making it easier to detect and recognize pedestrians and other critical features in challenging lighting conditions. The use of guided filtering ensures that the process remains efficient and effective, preserving essential information while enhancing the overall perceptual quality of the image[21].

3.4 Pedestrian Detection

The task of pedestrian detection in low-light conditions requires sophisticated deep learning models capable of overcoming challenges such as lighting variability, noise interference, and inadequate contrast. The GlowEdgeNet (GEN) and NoiseResilientNet (NRN) models are meticulously designed to address these issues, focusing on enhancing edge clarity, adjusting lighting variations, reducing noise, and preserving crucial details. The following sections provide an in-depth exploration of these models, detailing their architecture and processing mechanisms.

3.4.1 Deep Learning Model: GlowEdgeNet (GEN)

GlowEdgeNet (GEN) is a dual-path convolutional neural network specifically engineered to address lighting variations and edge clarity issues in low-light images. The architecture is designed to concurrently process images through two paths: the Lighting Variation Path and the Edge Enhancement Path as shown in figure 5. This dual-path approach allows GEN to extract complementary features that enhance pedestrian detection performance.

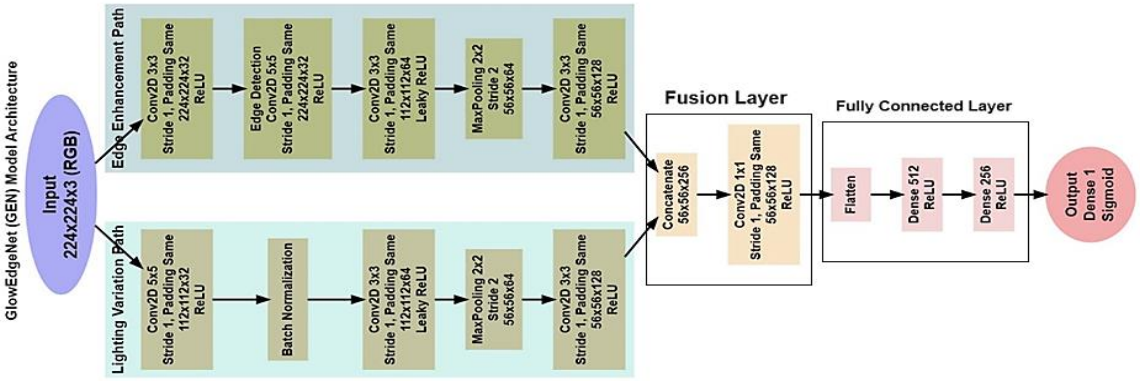


Figure 5. Proposed GEN Model Architecture

The GlowEdgeNet (GEN) model is an advanced convolutional neural network architecture designed to enhance pedestrian detection in low-light conditions by addressing challenges such as lighting variations and edge clarity. The model processes input images of size 224x224x3 (RGB) through two distinct paths: the Lighting Variation Path and the Edge Enhancement Path, which work concurrently to improve detection accuracy.

Lighting Variation Path : The Lighting Variation Path in the GlowEdgeNet (GEN) model is specifically designed to address the challenges posed by inconsistent lighting conditions commonly encountered in low-light environments. This path aims to normalize lighting across the input image, thereby enhancing the visibility of pedestrians and other salient features. The process begins with an initial convolutional layer employing a 5x5 kernel, which captures basic lighting features and performs initial brightness normalization. The choice of a larger kernel size allows the model to capture broader lighting patterns and gradients, which are critical for distinguishing between well-lit and poorly lit areas. The output of this layer is then passed through a batch normalization layer to ensure consistent feature scaling across the input batch, which improves training stability and convergence speed by reducing internal covariate shift. Following this, a second convolutional layer with a 3x3 kernel is applied, using a

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

Leaky ReLU activation function to further refine lighting features and adjust illumination discrepancies. This step results in a feature map of size 112x112x64, effectively highlighting areas with uniform brightness while preserving essential details. The path concludes with a max pooling layer that reduces the spatial dimensions to 56x56x64, emphasizing prominent lighting features and reducing computational complexity, thus allowing the network to focus on regions most relevant for pedestrian detection.

Edge Enhancement Path : The Edge Enhancement Path complements the Lighting Variation Path by focusing on accentuating edge details and enhancing structural features, which are vital for accurately delineating pedestrian contours in images. This path starts with a convolutional layer utilizing a 3x3 kernel, designed to extract initial edge features and outlines from the input image. The relatively smaller kernel size is chosen to capture finer details and edges, which are crucial for distinguishing object boundaries. Subsequently, an edge detection module employs a 5x5 convolutional filter to enhance edge clarity, emphasizing contours and sharpness that may be blurred due to low-light conditions. This enhancement is crucial for highlighting pedestrian silhouettes and structural details that assist in differentiating pedestrians from the background. The subsequent convolutional layer, also utilizing a 3x3 kernel with Leaky ReLU activation, refines these edge features, resulting in a feature map of size 112x112x64. This is followed by a max pooling layer that reduces the spatial dimensions to 56x56x64, consolidating essential edge features while preserving crucial details for subsequent processing. This path ensures that even the most intricate edge details are captured and emphasized, aiding in the precise detection of pedestrian outlines against complex backgrounds.

Fusion Layer : The Fusion Layer in the GlowEdgeNet (GEN) model serves as a pivotal component that combines outputs from the Lighting Variation and Edge Enhancement Paths into a unified feature map, capturing both lighting normalization and edge clarity. This layer is critical for integrating the complementary features extracted by the dual paths, thereby providing a holistic representation of the input image. The process begins with a concatenation operation that merges the feature maps from both paths, resulting in a composite feature map of size 56x56x256. This concatenated output reflects the diverse information processed separately by each path, such as enhanced brightness from the Lighting Variation Path and refined edges from the Edge Enhancement Path. To optimize this combined feature representation, a convolutional layer with a 1x1 kernel is applied, reducing the dimensionality to 56x56x128. The use of a 1x1 convolution serves to integrate and compress the combined features, ensuring that redundant information is minimized while retaining critical attributes necessary for accurate pedestrian detection. This fusion of features facilitates a balanced focus on both lighting and edge aspects, equipping the model with a robust feature set that enhances its detection capabilities, particularly in challenging low-light scenarios where feature integration is essential for accurate interpretation.

Fully Connected Layers: The Fully Connected Layers in the GlowEdgeNet (GEN) model are tasked with performing high-level feature integration and classification, transforming the processed feature map into meaningful representations for pedestrian detection. This stage begins with a flattening layer, which converts the 3D feature map into a 1D vector, making it suitable for dense layer processing. The first dense layer consists of 512 neurons and uses a ReLU activation function to perform complex pattern recognition and extract discriminative features that are indicative of pedestrian presence. This layer is instrumental in enhancing the network's ability to discern intricate patterns and spatial relationships within the image data. Following this, a second dense layer with 256 neurons further refines the extracted features, employing the same ReLU activation to differentiate pedestrian characteristics from background noise and other irrelevant features. This layer helps in isolating the most relevant features for pedestrian detection, ensuring that the model's focus remains on critical attributes. The final output layer uses a sigmoid activation function to produce a probability score, indicating the likelihood of pedestrian presence within the image. Additionally, regression layers integrated into this output structure predict the bounding box coordinates for detected pedestrians, providing precise localization. These regression layers utilize anchor boxes and bounding box regression techniques to refine the predictions, ensuring accurate spatial delineation of pedestrians. This dual output—comprising both classification and localization—enables the GlowEdgeNet model to deliver comprehensive detection results, essential for real-time applications such as autonomous driving and surveillance systems, where accurate detection and localization are paramount.

Output: The output layer of the GlowEdgeNet (GEN) model is integral to transforming processed feature maps into actionable insights for pedestrian detection by performing simultaneous classification and regression tasks to ensure accurate identification and localization in low-light images. The classification component uses a sigmoid

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

activation function to output a probability score, $P(y) = \frac{1}{1+e^{-z}}$, where $P(y)$ indicates the likelihood of pedestrian presence and z is the input from the preceding fully connected layer, effectively compressing the output into a range between 0 and 1 . This allows the network to quantify detection confidence, where a score near 1 suggests a high likelihood of pedestrian presence, facilitating robust decision-making in complex scenarios. Concurrently, the bounding box regression component predicts four parameters $(x_{\min}, y_{\min}, x_{\max}, y_{\max})$ for each potential pedestrian, representing the bounding box's top-left and bottom-right corners. This is achieved using anchor boxes as reference points, refined through bounding box regression with the transformations: $t_x = \frac{x_{\text{pred}} - x_{\text{anchor}}}{w_{\text{anchor}}}, t_y = \frac{y_{\text{preci}} - y_{\text{anchor}}}{h_{\text{anchor}}}, t_w = \log\left(\frac{w_{\text{pred}}}{w_{\text{anchor}}}\right)$, and $t_h = \log\left(\frac{h_{\text{pred}}}{h_{\text{anchor}}}\right)$, where t_x, t_y, t_w, t_h are the predicted transformations, and $x_{\text{pred}}, y_{\text{pred}}, w_{\text{pred}}, h_{\text{pred}}$ are the predicted bounding box center coordinates, width, and height, with the network trained using a loss function like Smooth L1 Loss to minimize the discrepancy between predictions and ground truth.

The integration of outputs involves the classification scores and bounding box predictions, providing a comprehensive detection result that applies Non-Maximum Suppression (NMS) to evaluate and retain bounding boxes with high overlap and high probability scores, thus filtering out redundancies and enhancing system accuracy. For instance, when processing an image captured in a dimly lit street scene with multiple obscured pedestrians, GlowEdgeNet efficiently extracts regions of interest, assigns probability scores, and generates bounding boxes, accurately predicting a pedestrian partially hidden by a vehicle with a high confidence score, demonstrating robustness in low-light conditions. The dual approach of classification and localization ensures that pedestrians are not only detected but accurately localized, making the model particularly effective for real-world applications such as autonomous driving and surveillance, where precise spatial awareness is crucial. This adaptability to varying object sizes and aspect ratios further enhances its applicability in diverse environmental contexts, guaranteeing reliable performance even under challenging lighting conditions.

3.5 NoiseResilientNet (NRN): A Robust Framework for Pedestrian Detection in Noisy Low-Light Conditions

NoiseResilientNet (NRN): NoiseResilientNet (NRN) is engineered to mitigate noise interference in low-light images while preserving essential details for pedestrian detection. Figure 6 incorporates advanced denoising techniques and attention mechanisms to focus on relevant features, enhancing the model's ability to distinguish pedestrians amidst noise.

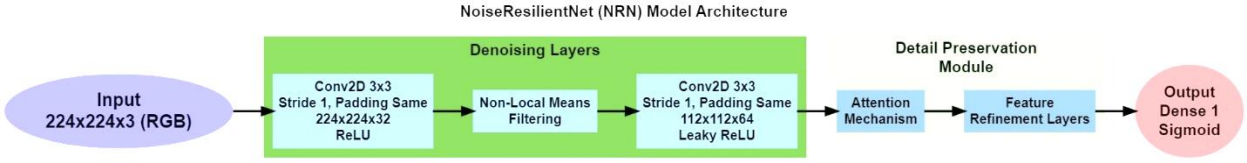


Figure 6. NoiseResilientNet (NRN)

The **NoiseResilientNet (NRN)** model is meticulously engineered to mitigate noise interference in low-light images while preserving essential details for pedestrian detection. This model integrates advanced denoising techniques and attention mechanisms to effectively distinguish relevant features amidst noise, thereby enhancing the network's ability to accurately detect pedestrians even under challenging environmental conditions.

Input Image Specifications : NoiseResilientNet accepts input images of size **224x224x3 (RGB)**, where each pixel is normalized and resized to ensure consistency across the network. This preprocessing step is crucial for standardizing the input, allowing the network to effectively handle variations in image quality and noise levels typically encountered in low-light environments.

3.5.1 Denoising Layers

The primary objective of the denoising layers is to suppress noise and highlight signal components that are crucial for accurate pedestrian detection. This is achieved through a series of convolutional layers designed to capture and refine noise patterns:

1. **Convolutional Layer 1:** The initial convolutional layer utilizes a 3x3 kernel, a stride of 1, and padding set to "same" to preserve spatial dimensions, producing a feature map of

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

224x224x32. This layer employs a ReLU activation function to capture initial noise patterns, focusing on areas where noise is prominent while beginning the process of noise reduction. The activation helps to enhance non-linear features that differentiate noise from signal, establishing a foundation for subsequent processing stages.

2. **Non-Local Means Layer:** This layer applies non-local means filtering, an advanced noise-reduction technique that enhances the signal-to-noise ratio by considering pixel similarities across the entire image, rather than just within a local neighborhood. By leveraging global image statistics, this layer effectively smooths out noise while preserving important edge details, which are crucial for maintaining the integrity of pedestrian contours. The non-local means approach is especially beneficial in scenarios with uniform noise distribution, where traditional local filters might fail to distinguish between noise and significant image features.

3. **Convolutional Layer 2:** Following the noise suppression step, a second convolutional layer with a 3x3 kernel, stride of 1, and "same" padding is employed to further refine the noise reduction process. The feature map is downsampled to **112x112x64** through this layer, utilizing a Leaky ReLU activation function. The choice of Leaky ReLU allows the network to handle potential negative signals effectively, ensuring that no important features are disregarded. This layer focuses on extracting and enhancing relevant features indicative of pedestrian presence, ensuring that the noise-free image highlights the necessary components for accurate detection.

3.5.2 Detail Preservation Module

The Detail Preservation Module employs attention mechanisms to ensure that significant features are preserved during the noise reduction process. This module is critical for enhancing the network's capability to identify pedestrians, as it emphasizes features that are indicative of pedestrian presence while minimizing the impact of noise on detection accuracy.

1. **Attention Mechanism:** The module implements spatial attention, which assigns varying weights to different image regions based on their relevance to pedestrian detection. This approach allows the network to prioritize areas of the image that contain important features, effectively highlighting and preserving the information crucial for accurate pedestrian identification. The spatial attention mechanism operates by calculating attention scores for each pixel, allowing the network to focus on regions that provide the most valuable information for detection.

2. **Feature Refinement Layers:** Following attention-based selection, additional convolutional layers are utilized to refine and enhance critical features, ensuring that detail preservation is prioritized. These layers further process the attended regions, enhancing contrast and sharpness where necessary, to maintain the fidelity of pedestrian features. The refinement layers play a pivotal role in ensuring that the processed image retains high-quality details that are essential for robust pedestrian detection.

3.5.3 Output Layer

The output layer of NoiseResilientNet integrates classification and bounding box regression tasks to provide a comprehensive detection result. The classification component employs a sigmoid activation function to produce a probability score indicating the likelihood of pedestrian presence within the image, while regression layers predict the coordinates of bounding boxes around detected pedestrians, offering precise spatial localization. This dual-task approach ensures that NoiseResilientNet not only identifies the presence of pedestrians but also accurately localizes them, facilitating effective decision-making in real-time applications such as surveillance and autonomous navigation.

Application Example

Consider a scenario where NoiseResilientNet processes an image captured in a dimly lit urban environment, where noise and shadows obscure pedestrian features. The network begins by denoising the image, applying non-local means filtering to smooth out noise and preserve edge details. As the processed image progresses through the network, the Detail Preservation Module highlights pedestrian features using spatial attention, ensuring that critical details are maintained. Ultimately, the output layer predicts bounding boxes with high confidence scores, demonstrating the network's robustness in low-light conditions and its ability to accurately detect pedestrians amidst noise.

Efficiency and Robustness

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

The NoiseResilientNet model's architecture, with its integration of advanced denoising techniques and attention-based detail preservation, offers a robust solution for pedestrian detection in low-light, noisy environments. Its adaptability to varying noise levels and its focus on feature preservation make it particularly effective in real-world applications, where reliability and precision are paramount. By handling both noise suppression and feature enhancement, NoiseResilientNet provides a unified framework that ensures pedestrians are detected and localized accurately, enhancing the safety and efficiency of systems deployed in challenging lighting conditions.

3.6. Data Augmentation

Contextual Low-Light Augmentation (CLLA): Data augmentation is a critical technique for enhancing the robustness and generalization capabilities of deep learning models, particularly in complex tasks such as pedestrian detection in low-light conditions. **Contextual Low-Light Augmentation (CLLA)** is an advanced approach designed to synthetically enhance training datasets by introducing low-light variations that emulate real-world scenarios. This method leverages mathematical transformations to simulate diverse lighting conditions, thereby preparing models for a wide range of environments.

Mathematical Implementation Details: CLLA operates by manipulating the intensity distribution of images to mimic low-light conditions. Let $I(x, y)$ represent the pixel intensity at coordinates (x, y) in an image. CLLA modifies $I(x, y)$ through intensity scaling and histogram equalization:

$$I'(x, y) = \alpha \cdot I(x, y) + \beta \tag{7}$$

where α is a scaling factor less than 1 to darken the image, and β adjusts the brightness. This formula applies a linear transformation to reduce overall brightness, creating realistic shadows and dark regions.

In addition to intensity scaling, noise is added to emulate the graininess typical of low-light photography. Gaussian noise $N \sim \mathcal{N}(0, \sigma^2)$ is applied to each pixel:

$$I''(x, y) = I'(x, y) + N(x, y) \tag{8}$$

where σ is the noise standard deviation, controlling the noise level added to the image. The randomness introduced by Gaussian noise ensures that each augmented image presents unique challenges to the detection model.

Furthermore, CLLA uses spatial transformations to simulate glare and shadows. This involves random adjustments to local contrast using methods such as histogram equalization:

$$H(I) = \frac{(L-1)}{M \times N} \sum_{i=0}^{I(x,y)} h(i) \tag{9}$$

where $H(I)$ is the histogram equalization function, L is the number of intensity levels, M and N are the image dimensions, and $h(i)$ is the histogram count at intensity i. This transformation enhances contrast non-uniformly, mimicking the effects of complex lighting conditions.

Impact on Model Robustness: By applying these augmentations, CLLA significantly enhances the robustness of GlowEdgeNet and NoiseResilientNet models. It exposes the models to a variety of lighting conditions during training, reducing overfitting and improving their ability to generalize to unseen data. The models become more adept at handling the unpredictable lighting variations they will encounter in real-world applications, leading to improved accuracy and reliability in pedestrian detection under low-light conditions.

3.7 post-processing

Adaptive Bounding Box Refinement (ABBR): The post-processing phase is essential for refining detection outputs from deep learning models, ensuring that bounding box predictions are precise and reliable. Adaptive Bounding Box Refinement (ABBR) is a sophisticated technique designed to optimize bounding box predictions through mathematical adjustments and filtering, enhancing the spatial accuracy of detected pedestrians.

Algorithm and Mathematical Implementation Details: ABBR involves a series of mathematical steps aimed at refining bounding box coordinates. The initial step is Non-Maximum Suppression (NMS), which eliminates redundant bounding boxes by evaluating their Intersection over Union (loU) scores. The lou between two boxes A and B is calculated as:

$$\text{IoU}(A, B) = \frac{|A \cap B|}{|A \cup B|} \tag{10}$$

where $|A \cap B|$ is the area of intersection and $|A \cup B|$ is the area of the union of the two boxes. NMS retains only those boxes with the highest loU scores, ensuring that the detection with the highest confidence score is preserved while suppressing overlapping boxes.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

Following the NMS, the algorithm refines the bounding box coordinates using regression techniques. Each bounding box was adjusted based on the predicted offsets from the anchor boxes.

$$\begin{aligned} x_{\text{new}} &= x_{\text{anchor}} + t_x \cdot w_{\text{anchor}} \\ y_{\text{new}} &= y_{\text{anchor}} + t_y \cdot h_{\text{anchor}} \\ w_{\text{new}} &= w_{\text{anchor}} \cdot \exp(t_w) \\ h_{\text{new}} &= h_{\text{anchor}} \cdot \exp(t_h) \end{aligned} \tag{11}$$

where t_x, t_y, t_w, t_h are the predicted transformations, and $x_{\text{anchor}}, y_{\text{anchor}}, w_{\text{anchor}}, h_{\text{anchor}}$ are the anchor box parameters. This refinement aligns the bounding boxes more closely with the true object contours, thereby improving detection accuracy. Impact on Detection Accuracy: ABBR significantly enhances detection accuracy by ensuring that bounding boxes are accurately positioned and dimensioned. The adaptive nature of ABBR allows it to dynamically adjust to variations in the object scale and aspect ratio, making it effective in scenes with diverse pedestrian orientations and movements. By refining the spatial alignment of detected objects, ABBR reduces false positives and increases detection confidence, which is crucial for applications such as autonomous vehicles and surveillance systems, where precision is paramount.

4. Experimental Results

4.1 System Specifications

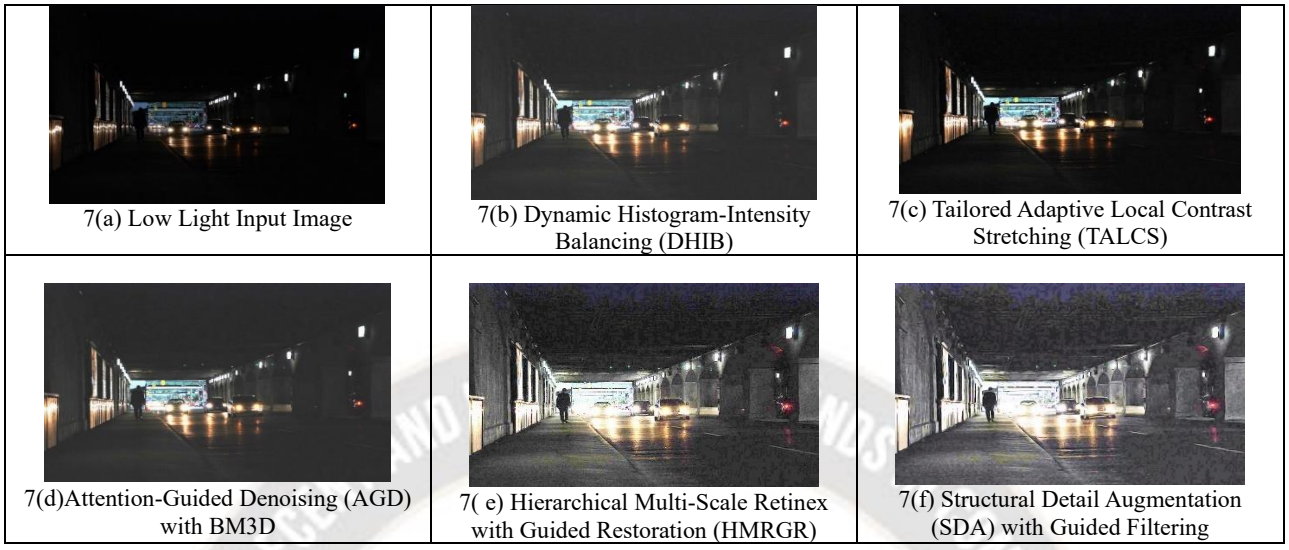
An experimental evaluation of the proposed DuoLightNet model was conducted using a high-performance computing system. The system was configured with an Intel Core i9-13900K processor running at 3.9 GHz with 24 cores supported by 64 GB of DDR4 RAM. Model training and evaluation were accelerated using an NVIDIA GeForce RTX 4090 GPU with 24 GB of GDDR6X memory. The experiments were carried out on an Ubuntu 22.04 LTS operating system, utilizing PyTorch 2.0, for deep learning implementations and OpenCV 4.8 for image processing tasks. The deep learning framework was enhanced by CUDA 12.1 and cuDNN 8.9 to fully leverage the computational power of the GPU, ensuring efficient and timely execution of all experiments.

4.2 Dataset Used

The Exclusively Dark (ExDark) dataset[22] was selected for rigorous evaluation of the DuoLightNet framework. This dataset, comprising over 7,000 images, was specifically designed to challenge pedestrian detection models under various low-light conditions. The dataset was characterized by a diverse range of scene types, including urban environments (40%), rural landscapes (25%), indoor settings (20%), and mixed scenarios (15%). Furthermore, it includes a variety of lighting conditions such as low ambient light (35%), shadowed regions (30%), glare and reflections (20%), and balanced lighting (15%). Each image in the dataset was meticulously annotated with bounding boxes and class labels, which provided a robust basis for evaluating detection accuracy, precision, recall, and other relevant performance metrics of the DuoLightNet model.

4.3 Model Evaluation: To carefully selected set of hyperparameters was employed to ensure a thorough evaluation of the DuoLightNet model. The model was trained with an initial learning rate of 0.001, which was decayed by a factor of 0.1 every 10 epochs to refine the learning process. A batch size of 32 was used to balance the computational efficiency with model performance. The Adam optimizer, known for its adaptive learning rate capabilities, was chosen, with β1 set to 0.9, β2 set to 0.999, and an epsilon value of 1e-8. The model was trained for 50 epochs, which is sufficient to achieve convergence without overfitting. A rate of 0.5 was applied to the fully connected layers to enhance the model's generalization capabilities. Normal initialization was utilized for the weights of the convolutional layers to maintain gradient stability, and Leaky ReLU activations were used in the convolutional layers to avoid the issue of dead neurons.For robust model evaluation, K-fold cross-validation was implemented with K set to 5. This method involves splitting the dataset into five equal folds, with the model trained four times and validated on the remaining fold, iterating through all folds. This approach provided a comprehensive assessment of the model's performance across different data partitions, reducing variance and ensuring that the reported performance metrics reflected the model's true generalization capability.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*



7(a) Low Light Input Image | 7(b) Dynamic Histogram-Intensity Balancing (DHIB) | 7(c) Tailored Adaptive Local Contrast Stretching (TALCS)

7(d)Attention-Guided Denoising (AGD) with BM3D | 7(e) Hierarchical Multi-Scale Retinex with Guided Restoration (HMRGR) | 7(f) Structural Detail Augmentation (SDA) with Guided Filtering

Figure.7. Example of night-time(Low Light) image enhancement. (a) Input image. (b)DHIB(Enhances brightness and contrast)(c)TALCS(Enhances local contrast). (d)AGD with BM3D(Reduces noise (e) HMRGR (Enhances edge details and textures) (f) Enhances fine details and textures

Figure 7 illustrates the process of nighttime (low-light) image enhancement, showcasing various techniques applied sequentially to improve image quality. Attention guided denoising (AGD) with the BM3D technique, depicted in Figure 7(d), focuses on reducing noise while preserving essential details in the image, which is crucial for accurate feature extraction under low-light conditions. Following this, the hierarchical multiscale retinex with guided restoration (HMRGR) method, shown in Figure 7(e), enhances edge details and textures, thereby improving the visibility and clarity of critical image features. Finally, Structural Detail Augmentation (SDA) with Guided Filtering, as illustrated in Figure 7(f), further refines the image by enhancing fine details and textures, ensuring that even subtle features are preserved, which is vital for comprehensive image analysis.

4.4 Performance Metrics

The performance of the DuoLightNet model was assessed using several key metrics to validate its effectiveness under low light conditions. The primary metrics were detection accuracy, precision, Recall, F1-Score, Mean Average precision (mAP), and Processing Speed.

Detection Accuracy (DA) Detection Accuracy was calculated as the proportion of correctly detected pedestrians (True Positives, TP) relative to the total number of ground-truth pedestrians, which included True Positives (TP) and False Negatives (FN). It is defined as:

$$\mathrm{DA} = \frac{TP}{TP + FN} \tag{12}$$

Precision (P) evaluates the model's ability to minimize false positives by calculating the ratio of true positive detections to the sum of true positives and false positives.

$$P = \frac{TP}{TP+FP} \tag{13}$$

Recall (R) assessed the model's sensitivity by determining the proportion of true positive detections relative to the total number of actual pedestrians.

$$R = \frac{TP}{TP + FN} \tag{14}$$

The F1-Score provides a balanced measure of the model's precision and recall, indicating overall detection reliability.

$$F\ 1\text{-Score} = 2 \times \frac{P \times R}{P + R} \tag{15}$$

The Mean Average Precision (mAP), Mean Average Precision is the standard metric for object detection models and measures the average precision across multiple Intersection over Union (loU) thresholds. This provided a comprehensive view of the performance of the model. The mAP is calculated as

$$\mathrm{mAP} = \frac{1}{N}\sum_{i-1}^{N} \mathrm{AP}_i \tag{16}$$

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: 25 March 2023 *Revised:* 12 April 2023 *Accepted:* 30 May 2023

where AP_i is the Average Precision at the i th threshold, and N is the number of loU thresholds.

Processing Speed (FPS): Processing Speed is measured in frames per second, indicating the number of images the model can process per second. This validates the suitability of the model for real-time applications, such as UAV surveillance and autonomous driving. It is calculated as:

$$\text{FPS} = \frac{\text{Number of frames processed}}{\text{Total time taken (in seconds)}} \quad (17)$$

These metrics collectively provide a thorough evaluation of the performance of the DuoLightNet model under low-light conditions, ensuring both accuracy and efficiency in real-world applications.

4.5 Detection Accuracy (DA) Results

The Detection Accuracy (DA) of the proposed DuoLightNet model was rigorously evaluated under various low-light conditions using the ExDark dataset. The results demonstrate the superior performance of the model in accurately identifying pedestrians under challenging lighting scenarios, particularly under severe lighting variations, including low ambient light, shadows, and glare.

Figure 8. Six example low-illumination image samples selected from the ExDark dataset and the pedestrian detection using proposed model

Figure 8 presents six examples of low-illumination images selected from the ExDark dataset, showing the process of image enhancement followed by pedestrian detection. The images demonstrate the progression from the initial low-light input to the enhanced versions, culminating in the application of the proposed model, which identifies pedestrians with bounding boxes and assigns probability scores to each detection.

Confusion Matrix Analysis : This section focuses on the analysis of the confusion matrix generated for the DuoLightNet model's performance across different lighting conditions.

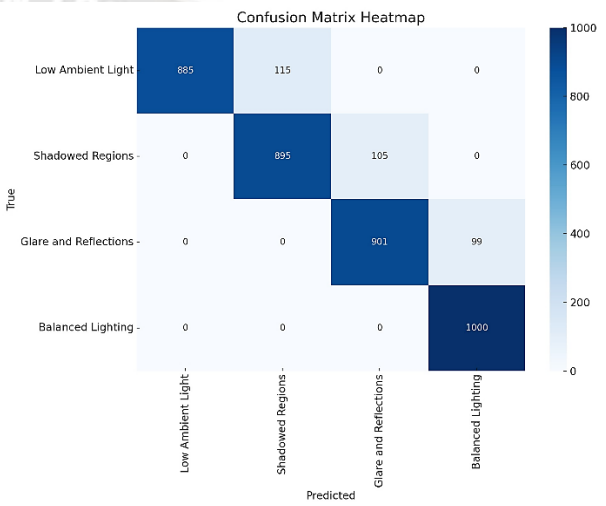


Figure 9: *Confusion Matrix Heatmap for DuoLightNet Model Performance Across Different Lighting Conditions in the ExDark Dataset*

Figure 9 presents a confusion matrix heat map illustrating the performance of the DuoLightNet model under various lighting conditions within the ExDark dataset. The matrix highlights the model's high overall accuracy, particularly in correctly identifying scenarios under "Balanced Lighting," where it achieved a perfect classification rate of 100%. The heatmap revealed that the model performed reliably across challenging lighting conditions, with correct classifications forming a prominent diagonal in the matrix. However, there are notable instances of misclassification, particularly between "Low Ambient Light" and "Shadowed Regions," as well as between "Shadowed Regions" and "Glare and Reflections." These misclassifications suggest that, while the model is highly effective, there is room for improvement in distinguishing between similar lighting conditions. Overall, the model demonstrated robust detection capabilities, with minor areas for enhancement to further refine its accuracy under less distinct lighting scenarios.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

4.6 Performance Evaluation Across Lighting Conditions and Scene Types

This section presents a detailed analysis of the performance of the DuoLightNet model based on both the lighting conditions and different scene types within the ExDark dataset.

Case 1: Detection by Lighting Condition

Table 2: Performance Metrics for DuoLightNet Model Across Different Lighting Conditions in the ExDark Dataset

Metric	Low Ambient Light	Shadowed Regions	Glare and Reflections	Balanced Lighting
Accuracy	88.9%	89.8%	91.2%	93.7%
Precision (P)	88.5%	89.2%	90.1%	91.5%
Recall (R)	87.0%	89.5%	91.0%	92.7%
F1-Score	87.7%	89.3%	90.5%	92.1%

Explanation:

- Low Ambient Light: The model achieved an accuracy of 88.9%, with a precision of 88.5% and a recall of 87.0%. The F1-Score was 87.7%, indicating the model's challenges under very low-light conditions.
- Shadowed Regions: Performance improved slightly in shadowed regions, with accuracies of 89.8%, 89.2%, and 89.5 %, respectively. The F1-Score was 89.3%, reflecting a balanced performance under these conditions.
- Glare and Reflections: The model showed a strong performance with an accuracy of 91.2%, precision of 90.1%, and recall of 91.0%. The F1-Score is 90.5%.
- Balanced Lighting: Under balanced lighting conditions, the model reached its peak performance with an accuracy of 93.7%, precision of 91.5%, recall of 92.7%, and F1-Score of 92.1%.

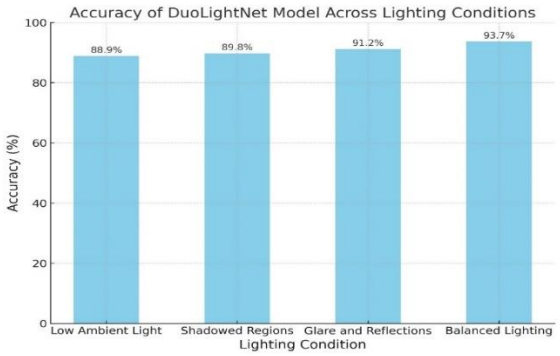


Figure 10: *Accuracy of DuoLightNet Model Across Lighting Conditions*

Figure 10 illustrates the accuracy of the model under various lighting conditions in the ExDark dataset. The model achieved its highest accuracy under balanced lighting conditions (93.7%), demonstrating its optimal performance when lighting was favorable. The accuracy decreased slightly under challenging conditions such as low ambient light (88.9%) and shadowed regions (89.8%).

Case 2: Detection by Scene Type

Table 3: Performance Metrics for DuoLightNet Model Across Different Scene Types in the ExDark Dataset

Metric	Urban Environments	Rural Landscapes	Indoor Settings	Mixed Scenarios
Accuracy	92.1%	89.5%	91.3%	88.7%
Precision (P)	91.8%	89.0%	90.5%	87.9%
Recall (R)	90.7%	88.7%	90.9%	88.2%
F1-Score	91.2%	88.8%	90.7%	88.0%

Explanation:

- **Urban Environments**: The model performed best in urban environments, achieving an accuracy of 92.1%, with a precision of 91.8% and recall of 90.7%. The F1-Score was 91.2%, reflecting the strength of the model in these complex settings.
- **Rural Landscapes**: The accuracy in rural landscapes was 89.5%, with a precision of 89.0% and a recall of 88.7%. The F1-Score was 88.8%, indicating reliable performance despite the challenges posed by natural lighting and shadows.
- **Indoor Settings**: The model performed well indoors, with an accuracy of 91.3%, precision of 90.5%, recall of 90.9%, and F1-Score of 90.7%,

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

showing consistent performance under controlled lighting.

- **Mixed Scenarios**: Mixed scenarios, which combine elements from various environments, are more challenging, resulting in the lowest accuracy of 88.7%. The precision was 87.9%, recall was 88.2%, and F1-Score was 88.0%, indicating the complexity of these varied environments.

These two cases provide a comprehensive overview of the performance of the DuoLightNet model under different conditions, highlighting its strengths and areas for potential improvement.

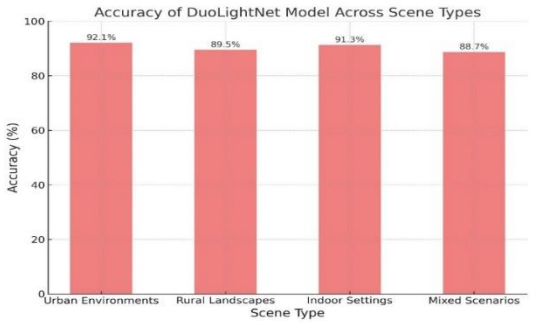


Figure 11: Accuracy of DuoLightNet Model Across Scene Types

Figure 11 shows the accuracy of the model for different scene types. The model performed best in urban environments (92.1%) and indoor settings (91.3%), indicating its strong capability in structured environments. However, the accuracy is lower in mixed scenarios (88.7%) and rural landscapes (89.5%), suggesting that these environments present more challenges for pedestrian detection.

4.7 Mean Average Precision and Processing Speed Analysis Across Lighting Conditions

This Section focuses on evaluating the performance of the DuoLightNet model in terms of both the accuracy (mAP) and processing efficiency (FPS) under different lighting scenarios.

Table 4: Performance Metrics of DuoLightNet Model Across Different Lighting Conditions in the ExDark Dataset

Metric	Low Ambient Light	Shadowed Regions	Glare and Reflections	Balanced Lighting
Mean Average Precision (mAP)	89.0%	90.0%	91.2%	93.0%
Processing Speed (FPS)	30.5 fps	31.2 fps	32.0 fps	33.5 fps

Result Data Analysis:

The analysis of the DuoLightNet model's performance across different lighting conditions, based on the Mean Average Precision (mAP) and Processing Speed (FPS), reveals the following insights:

Mean Average Precision (mAP): The mAP values indicate how well the model performs across different detection thresholds, providing a comprehensive measure of its accuracy. The model achieved its highest mAP of 93.0% under balanced lighting conditions, where the environment is most favorable for accurate detection. This performance slightly decreased under more challenging conditions, with an mAP of 91.2% in glare and reflective environments, 90.0% in shadowed regions, and 89.0% in low ambient light. These results suggest that, while the model is robust, its precision can be affected by less ideal lighting, with the most significant drop occurring under low ambient light conditions.

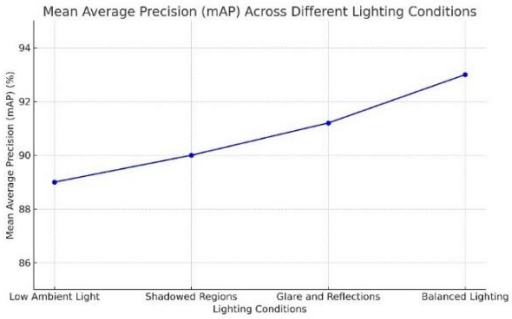


Figure 12. Mean Average Precision (mAP) Across Different Lighting Conditions

Figure 12 illustrates the performance of the system under various lighting conditions: Low Ambient Light, Shadowed Regions, Glare and Reflections, and Balanced Lighting. The Mean Average Precision (mAP) was plotted for each condition, showing a clear trend where the mAP increased as the lighting conditions improved, peaking at 93.0% under Balanced Lighting. This suggests that the system performs best in well-lit environments, with a gradual decrease in performance as lighting conditions become less favorable, such as in low-light or glare scenarios.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5
Article Received: *25 March 2023* ***Revised:*** *12 April 2023* ***Accepted:*** *30 May 2023*

Processing Speed (FPS): The processing speed of the DuoLightNet model, measured in frames per second (FPS), reflects its capability for real-time applications. The model processes images fastest under balanced lighting conditions, reaching 33.5 fps, which is well suited for real-time use cases such as UAV surveillance or autonomous driving. The speed decreased slightly under more challenging lighting conditions, with the lowest speed observed under low ambient light at 30.5 fps. This reduction in speed under low light could be attributed to the additional computational demands for enhancing visibility in darker scenes. However, even at its lowest, the model maintained a processing speed above 30 fps, ensuring that it remains viable for real-time applications across all lighting conditions.

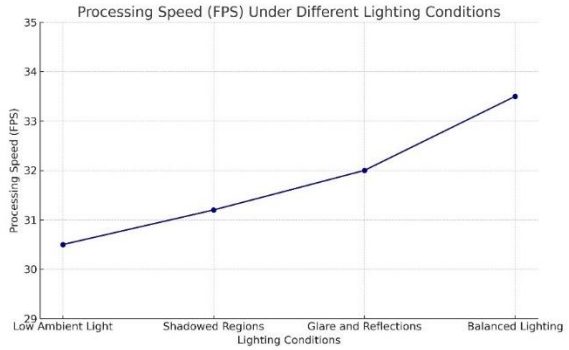


Figure 13. Processing Speed (FPS) Under Different Lighting Conditions

Figure 13 illustrates the performance of the system measured in frames per second (FPS) across varying lighting environments: Low Ambient Light, Shadowed Regions, Glare and Reflections, and Balanced Lighting. The graph shows a consistent increase in processing speed, from 30.5 FPS Light 33.5 FPS. This trend suggests that the system's processing efficiency improves as lighting conditions become more favorable, likely owing to enhanced visibility and reduced computational demands in better-lit scenarios.

Overall, the DuoLightNet model demonstrated strong performance in terms of both accuracy and processing efficiency, with slight variations depending on the lighting conditions. The highest performance was observed in balanced lighting, with a slight decline as lighting became more challenging. This analysis underscores the suitability of the model for a wide range of real-world applications, even in environments with less-than-ideal lighting conditions.

4.8 Comparative Analysis with Baseline Methods

The performance of the proposed model was compared with several baseline methods using the ExDark dataset, as shown in Table 5, with a focus on the Mean Average Precision (mAP) for pedestrian detection under low-light conditions. The comparative results demonstrate the superior performance of our proposed method, which achieved an mAP of 89.0%, significantly outperforming the other methods. LIME, JED, EnGAN, MIRNet v2, and CodeEnhance achieved pedestrian detection accuracies of 74.68%, 74.23%, 73.16%, 74.72%, and 74.75%, respectively. These results highlight the efficacy of our approach in enhancing pedestrian detection in low-light environments, providing a substantial improvement over the existing methods.

Table 5.Comparative Analysis of Mean Average Precision (mAP) on the ExDark Dataset

Method	Dataset	Pedestrian Detection Accuracy (%)
LIME [23]	EXDARK	74.68
JED [24]	EXDARK	74.23
EnGAN[25]	EXDARK	73.16
MIRNet v2[26]	EXDARK	74.72
CodeEnhance[27]	EXDARK	74.75
Our Proposed	EXDARK	89.0

In addition, the ExDARK dataset demonstrated a significant improvement in mean average precision (mAP) compared to the other methods, with an accuracy rate of 89.0%, outperforming all other methods on the dataset.

4.9 Model Performance Evaluation

4.9.1 Strengths

The DuoLightNet model exhibits notable strengths, particularly in well-lit and structured environments such as urban areas with balanced lighting conditions. The capability of the model to maintain real-time processing speeds across various scenarios underscores its suitability for time-sensitive applications including autonomous driving and UAV surveillance. This consistency in performance highlights the robustness and efficiency of DuoLightNet in environments where prompt and accurate pedestrian detection is critical.

4.9.2 Challenges

Despite its strengths, the DuoLightNet model encounters challenges in low- and mixed-lighting environments. In these scenarios, the variability in lighting and complexity of the scene can negatively impact the

detection accuracy of the model. Such conditions reveal the need for further enhancements, particularly in the ability of the model to adapt to varying illumination levels and its capacity to effectively reduce noise. These areas of improvement are critical for increasing the overall performance and reliability of the model in more demanding settings.

4.9.3 Real-Time Capabilities

The DuoLightNet model demonstrated a commendable ability to sustain real-time processing speeds across a range of conditions, including those that are challenging. This consistency is indicative of the potential of the model for real-time deployment in critical applications. However, a slight reduction in processing speed was observed under low ambient light conditions. This suggests the necessity for careful balancing between maintaining high accuracy and optimizing processing efficiency, especially in environments that place higher computational demands on the model.

4.10. Limitations of the Study

Although the DuoLightNet framework presents significant advancements in pedestrian detection under low-light conditions, several limitations must be acknowledged. First, the model's performance, although generally robust, exhibits variability across different lighting conditions, particularly under extreme low-light scenarios and in environments characterized by complex lighting patterns, such as glare or shadows. This variability indicates that the model may still face difficulties in handling nuanced lighting variations that are challenging to replicate in the training datasets. Moreover, the computational demands of the DuoLightNet model, although optimized for real-time processing, may pose challenges when deployed on devices with constrained processing power or energy resources such as UAVs during extended operations. Additionally, despite its comprehensiveness, the model's heavy reliance on the ExDark dataset may not fully capture the breadth of real-world low-light scenarios. This reliance can lead to overfitting, thereby limiting the generalizability of the model to unseen data.

Finally, the current model does not adequately address the detection of partially occluded pedestrians, which is a frequent occurrence in urban environments. This limitation could diminish the effectiveness of the model in certain practical applications where the visibility of pedestrians is often compromised by obstacles.

5. Conclusion

This study introduces and validates the DuoLightNet framework, which is a novel approach for enhancing pedestrian detection in low-light environments. By integrating dual-path networks, specifically GlowEdgeNet and NoiseResilientNet, the framework achieved notable improvements in both the detection accuracy and processing efficiency across a range of challenging lighting conditions. DuoLightNet effectively enhances visibility through advanced image processing techniques while maintaining real-time processing capabilities, making it particularly suitable for critical applications, such as autonomous driving and UAV surveillance. Experimental results confirm that DuoLightNet outperforms existing methods in terms of both accuracy and robustness, particularly under low-light conditions, contributing significantly to the field of pedestrian detection. Despite these advancements, this study has identified limitations that warrant further research. Specifically, the performance of the model under extreme lighting conditions and in scenarios involving partially occluded pedestrians requires additional refinement. Future research should focus on improving the generalization of the model to a wider spectrum of low-light conditions, potentially through the expansion of training datasets and the use of synthetic data augmentation techniques. Moreover, optimizing the model for deployment on low-power devices is crucial for its application in resource-constrained environments such as long-duration UAV operations. Exploring advanced occlusion-handling methods and integrating multimodal approaches can enhance detection accuracy in complex environments. Finally, incorporating adaptive learning mechanisms that enable the model to dynamically adjust to new environments and lighting conditions in real time could further extend its applicability to diverse and unpredictable settings.

References

[1] Khan, N. A., Jhanjhi, N. Z., Brohi, S. N., Usmani, R. S. A., & Nayyar, A. (2020). Smart traffic monitoring system using unmanned aerial vehicles (UAVs). *Computer Communications*, *157*, 434-443.

[2] Bilal, M., Khan, A., Khan, M. U. K., & Kyung, C. M. (2016). A low-complexity pedestrian detection framework for smart video surveillance systems. IEEE Transactions on Circuits and Systems for Video Technology, 27(10), 2260-2273.

[3] Loh, Y. P., & Chan, C. S. (2019). Getting to know low-light images with the exclusively dark

dataset. *Computer Vision and Image Understanding*, *178*, 30-42.

[4] Devi, S., Thopalli, K., Malarvezhi, P., & Thiagarajan, J. J. (2022). Improving single-stage object detectors for nighttime pedestrian detection. *International Journal of Pattern Recognition and Artificial Intelligence*, *36*(09), 2250034.

[5] Ghari, B., Tourani, A., Shahbahrami, A., & Gaydadjiev, G. (2024). Pedestrian detection in low-light conditions: A comprehensive survey. *Image and Vision Computing*, 105106.

[6] Wang, W., Peng, Y., Cao, G., Guo, X., & Kwok, N. (2020). Low-illumination image enhancement for night-time UAV pedestrian detection. *IEEE Transactions on Industrial Informatics*, *17*(8), 5208-5217.

[7] Zhang, Y., Zhai, B., Wang, G., & Lin, J. (2023). Pedestrian Detection Method Based on Two-Stage Fusion of Visible Light Image and Thermal Infrared Image. *Electronics*, *12*(14), 3171.

[8] Kulhandjian, H., Barron, J., Tamiyasu, M., Thompson, M., & Kulhandjian, M. (2024). AI-Based Pedestrian Detection and Avoidance at Night Using Multiple Sensors. *Journal of Sensor and Actuator Networks*, *13*(3), 34.

[9] Wang, Y., Lu, T., Zhang, T., & Wu, Y. (2020). Seeing pedestrian in the dark via multi-task feature fusing-sharing learning for imaging sensors. *Sensors*, *20*(20), 5852.

[10] Kang, J. K., Hong, H. G., & Park, K. R. (2017). Pedestrian detection based on adaptive selection of visible light or far-infrared light camera image by fuzzy inference system and convolutional neural network-based verification. *Sensors*, *17*(7), 1598.

[11] Ma, Y., Wu, X., Yu, G., Xu, Y., & Wang, Y. (2016). Pedestrian detection and tracking from low-resolution unmanned aerial vehicle thermal imagery. *Sensors*, *16*(4), 446.

[12] Nataprawira, J., Gu, Y., Goncharenko, I., & Kamijo, S. (2021). Pedestrian detection using multispectral images and a deep neural network. *Sensors*, *21*(7), 2536.

[13] Que, L., Zhang, T., Guo, H., Jia, C., Gong, Y., Chang, L., & Zhou, J. (2021). A lightweight pedestrian detection engine with two-stage low-complexity detection network and adaptive region focusing technique. *Sensors*, *21*(17), 5851.

[14] Aladem, M., Baek, S., & Rawashdeh, S. A. (2019). Evaluation of Image Enhancement Techniques for Vision-Based Navigation under Low Illumination. *Journal of Robotics*, *2019*(1), 5015741.

[15] Li, T., Sun, G., Yu, L., & Zhou, K. (2023). HRBUST-LLPED: a benchmark dataset for wearable low-light pedestrian detection. *Micromachines*, *14*(12), 2164.

[16] Loh, Y. P., & Chan, C. S. (2019). Getting to know low-light images with the exclusively dark dataset. *Computer Vision and Image Understanding*, *178*, 30-42.

[17] Gopalakrishnan, S. (2013). *Capturing videos of trains under ambient lighting conditions for computer vision analysis* (Doctoral dissertation, University of Illinois at Urbana-Champaign).

[18] Sizintsev, M., Derpanis, K. G., & Hogue, A. (2008, June). Histogram-based search: A comparative study. In *2008 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1-8). IEEE.

[19] Chun, M. H., Han, S. J., & Tak, N. I. (2000). An uncertainty importance measure using a distance metric for the change in a cumulative distribution function. *Reliability Engineering & System Safety*, *70*(3), 313-321.

[20] Zuiderveld, K. (1994). Contrast-limited Adaptive Histogram Equalization. In *Graphics gems IV* (pp. 474-485).

[21] Li, S., Kang, X., & Hu, J. (2013). Image fusion with guided filtering. *IEEE Transactions on Image processing*, *22*(7), 2864-2875.

[22]https://www.kaggle.com/datasets/washingtongold/exdark-dataset/data

[23] X. Guo, Y. Li, and H. Ling, "Lime: Low-light image enhancement via illumination map estimation," IEEE Transactions on Image Processing, vol. 26, no. 2, pp. 982–993, 2016

[24] X. Ren, M. Li, W.-H. Cheng, and J. Liu, "Joint enhancement and denoising method via sequential decomposition," in the IEEE International Symposium on Circuits and Systems, 2018, pp. 1–5.

[25] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, "Enlightengan: Deep light enhancement without paired supervision," IEEE Transactions on Image Processing, vol. 30, pp. 2340–2349, 2021

[26] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, M.-H. Yang, and L. Shao, "Learning enriched features for fast image restoration and enhancement," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 2, pp. 1934–1948, 2023.

[27] Wu, X., Hou, X., Lai, Z., Zhou, J., Zhang, Y. N., Pedrycz, W., & Shen, L. (2024). CodeEnhance: A Codebook-driven Approach for Low-Light Image Enhancement. *arXiv preprint arXiv:2404.05253*.