

Research Article

# LUMEN: Low-light Unified Multi-stage Enhancement Network to Improve RetinaFace-Based Face Detection

Rivaldi Ramadani <sup>1</sup>, De Rosal Ignatius Moses Setiadi <sup>2,\*</sup>, Imanuel Harkespan <sup>1</sup>, Kristoko Dwi Hartomo <sup>3</sup>, and Christian Arthur <sup>4</sup>

<sup>1</sup> Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang 50131, Indonesia; e-mail : rivaldiramadani73@gmail.com; harkespan@dsn.dinus.ac.id

<sup>2</sup> Research Group for Quantum Computing and Materials Informatics, Faculty of Computer Science, Universitas Dian Nuswantoro, Semarang 50131, Indonesia; e-mail : moses@dsn.dinus.ac.id

<sup>3</sup> Faculty of Information Technology, Satya Wacana Christian University, Salatiga 50715, Indonesia; e-mail : kristoko@uksw.edu

<sup>4</sup> School of Transdisciplinary, University of Technology Sydney, Ultimo, NSW 2007, Australia; e-mail : christian.arthur@student.uts.edu.au

\* Corresponding Author : De Rosal Ignatius Moses Setiadi

**Abstract:** Face detection in low-light conditions remains challenging due to underexposure, noise, and unstable contrast, which significantly degrade the performance of convolutional-based models. Conventional enhancement techniques, such as Histogram Equalization, Contrast-Limited Adaptive Histogram Equalization (CLAHE), and Low-Light Image Enhancement (LIME), often improve brightness but introduce visual artifacts and lack robustness for face detection. This study proposes Low-light Unified Multi-stage Enhancement Network (LUMEN), a lightweight multi-stage image enhancement pipeline designed as a preprocessing module to improve RetinaFace-based face detection under low-light conditions. LUMEN integrates Multiscale Retinex with Color Restoration, adaptive gamma correction, CLAHE-based local contrast enhancement, controlled image fusion, and Non-Local Means denoising to jointly stabilize illumination, preserve texture, and maintain visual naturalness. The method is evaluated on a low-light subset of the Human Faces Object Detection Dataset using RetinaFace. Detection performance is assessed using detection rate and confidence score, while visual quality is evaluated using no-reference metrics such as the Natural Image Quality Evaluator (NIQE) and the Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE). Experimental results show that LUMEN achieves a face detection rate of 91% with a high confidence score (0.9545), representing an improvement of 33 percentage points over raw low-light images (58%), while maintaining detection performance comparable to CLAHE, which ranks as the second-best method, LUMEN delivers superior perceptual quality, evidenced by the lowest BRISQUE score, as well as more stable visual appearance with reduced noise amplification and fewer contrast artifacts in low-light facial regions. LUMEN also achieves the lowest BRISQUE score, indicating superior texture preservation and perceptual stability. Ablation studies confirm that Retinex and CLAHE are the most critical components for detection robustness, while gamma correction, fusion, and denoising mainly contribute to visual naturalness. These results demonstrate that LUMEN provides an effective and practical preprocessing solution for low-light face detection.

Received: September, 9<sup>th</sup> 2025

Revised: December, 12<sup>th</sup> 2025

Accepted: December, 24<sup>th</sup> 2025

Published: December, 25<sup>th</sup> 2025

Curr. Ver.: December, 25<sup>th</sup> 2025



Copyright: © 2025 by the authors.

Submitted for possible open access publication under the terms

and conditions of the Creative

Commons Attribution (CC BY

SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>)

**Keywords:** Computer vision; Face detection; Low-light image enhancement; LUMEN; RetinaFace; Retinex; Surveillance systems.

## 1. Introduction

Lighting quality is a crucial factor in computer vision systems, particularly in applications such as security monitoring, public surveillance, and biometric identification. Under low-light conditions, cameras often fail to capture sufficient facial texture and structural details,

resulting in unstable face detection performance. Numerous reports indicate that a large proportion of criminal incidents, nighttime activities, and security events are recorded under dark or semi-dark environments, making robust face detection in low-light images highly valuable from a practical perspective [1]–[4]. These illumination limitations position image enhancement as a critical preprocessing step prior to face detection in low-light scenarios and, consequently, as a challenging and relevant research topic.

Research on face and object detection under low-light conditions has progressed significantly in recent years. Various deep learning–based detectors, such as FaceNet, YOLO, GhostFaceNets, and other Convolutional Neural Network (CNN)-based models, have demonstrated high accuracy under normal lighting conditions [5]–[7]. However, when applied to dark images, these models often degrade substantially due to increased noise, uneven intensity distributions, and low visual contrast [8]. This observation indicates that the visual quality of the input images strongly influences the robustness of detection models.

To address these limitations, several studies have proposed low-light face detection approaches through adaptive illumination enhancement [9]–[12], perceptual low-light enhancement [8], and Deep Color Consistent Networks [7]. These methods aim to improve detection performance by either modifying detector architectures or applying learning-based illumination correction. Although effective in certain scenarios, these approaches typically entail high computational complexity, reliance on large-scale training data, and limited flexibility for deployment in real-world operational environments, such as video surveillance systems.

YOLOv5 detectors are also commonly used as benchmarks in this context. YOLOv5 is known for its high detection speed but is relatively sensitive to noise, while GhostFaceNets offer computational efficiency yet exhibit instability under extreme lighting conditions [5], [13]. Meanwhile, deep illumination-aware detection methods that incorporate specialized illumination modules require additional training processes and complex configurations [8], [14]. These findings suggest that although detector architectures can be adapted to handle low-light conditions, optimal performance remains highly dependent on the quality of the input images. This conclusion is consistent with prior studies emphasizing that preprocessing—particularly image enhancement—plays a crucial role in improving detection robustness under low-light conditions [8], [15], [16].

RetinaFace is a modern face detection model designed to handle variations in pose, scale, and illumination through a Feature Pyramid Network (FPN) architecture and context modules [17]. Despite its strong architectural design, most existing evaluations of RetinaFace focus on normal lighting conditions and do not systematically explore integrating image enhancement as a preprocessing step in low-light conditions. Meanwhile, various enhancement techniques—including histogram-based methods such as Histogram Equalization (HE) and CLAHE [11], [18], Retinex-based approaches [6], [19], illumination estimation–based methods such as Low-Light Image Enhancement (LIME) [9], [20], and luminance adaptation via gamma correction [13], have been reported to improve detection performance across multiple computer vision models, including RetinaFace. However, comprehensive analyses examining how illumination enhancement optimizes RetinaFace performance under low-light conditions remain limited, highlighting a significant research gap.

Among existing enhancement techniques, LIME is recognized for its effectiveness in improving low-light image visibility with low computational complexity and without requiring training data [10], [20]. Nevertheless, several studies report that LIME, when used alone, is prone to producing excessive illumination artifacts and lacks robustness in maintaining color consistency and fine local details, particularly in facial regions with subtle textures [10]. These findings indicate that although LIME has strong potential as an enhancement component, it must be combined with other techniques to achieve more stable, reliable image quality in face detection applications.

Furthermore, improvements in image brightness do not always correlate with improved perceptual visual quality. Certain enhancement methods may increase detection accuracy while simultaneously introducing artifacts, over-contrast, or color distortion that degrade human visual interpretability. In surveillance and forensic contexts, image quality must also support manual analysis by human operators. Human-based perceptual evaluation is inherently subjective, inconsistent, and costly [15]. Therefore, this study adopts objective no-reference image quality metrics, namely NIQE and BRISQUE, which are grounded in principles of the Human Visual System (HVS), to provide quantitative and standardized visual quality assessment [2].

Based on these considerations, this research proposes a customized multi-technique enhancement approach designed to effectively improve low-light image quality while simultaneously supporting RetinaFace-based face detection and preserving visual interpretability for manual inspection. The proposed approach integrates multiple enhancement techniques into a staged pipeline that adaptively balances illumination correction, contrast enhancement, and preservation of facial structural information, as suggested by prior studies on low-light enhancement using illumination correction and adaptive luminance adjustment [9], [21].

Conceptually, the proposed pipeline combines global–local illumination correction, local contrast enhancement, and adaptive luminance adjustment to produce a more stable and informative intensity distribution. Such multi-stage approaches have been widely reported to overcome the limitations of single enhancement methods, which typically focus on only one aspect—such as illumination or contrast—without considering overall visual stability [11], [18]. Additionally, luminance adaptation via gamma correction has been shown to effectively recover details in dark regions while preventing over-enhancement in bright areas [13], [22], [23].

The proposed pipeline is introduced as the Low-light Unified Multi-stage Enhancement Network (LUMEN). This designation emphasizes that the method is not merely an ad hoc combination of enhancement techniques, but a fully integrated system with a consistent processing flow, rather than a learning-based neural architecture. By providing images with stabilized illumination, enhanced local contrast, and adaptively adjusted luminance, LUMEN is designed to improve the reliability of RetinaFace-based face detection under low-light conditions, where visual degradation often leads to the loss of critical texture and edge structures required for detection. The main contributions of this study are summarized as follows:

- A unified multi-stage low-light image enhancement pipeline (LUMEN) is proposed as a detector-agnostic preprocessing strategy to improve the robustness of RetinaFace-based face detection under low-light conditions, without modifying the detector architecture.
- A systematic evaluation framework is introduced to jointly analyze perceptual image quality and face detection performance, enabling recall-oriented assessment under low-light scenarios.
- A structured ablation study is conducted to quantitatively and visually examine the functional contributions of each enhancement module, revealing how illumination normalization, contrast enhancement, and noise suppression collectively support stable low-light face detection.

The rest of this paper is organized as follows: Section 2 reviews related works on low-light image enhancement and face detection. Section 3 presents the proposed LUMEN pipeline and its implementation details. Section 4 reports the experimental setup, quantitative and qualitative results, and ablation studies. Finally, Section 5 concludes the paper and outlines limitations and directions for future work.

## 2. Preliminaries

### 2.1. Multiscale Retinex with Color Restoration (MSRCR)

Retinex is a fundamental theory in image processing introduced by Land to explain how the human visual system separates reflectance and illumination information of an object [19]. Within the Retinex framework, an image is assumed to be formed as the product of reflectance and illumination components. Under low-light conditions, the illumination component is severely degraded, making it difficult to observe the reflectance details of objects, including facial structures [6], [15], [21]. The most basic Retinex model is Single Scale Retinex (SSR), which is formulated through a logarithmic comparison between the original image intensity and an illumination estimate obtained using a Gaussian surround function, as expressed in Equation (1) [6].

$$R_{SSR}(x, y) = \log I(x, y) - \log [I(x, y) * F(x, y)] \quad (1)$$

Where  $I(x, y)$  denotes the pixel intensity at location  $(x, y)$ , the symbol  $*$  represents the convolution operation, and  $F(x, y)$  is a Gaussian function used to approximate the local illumination.

Although SSR is effective at correcting global illumination, it tends to be unstable under complex, spatially varying illumination conditions. To overcome this limitation, Jobson et al. proposed

Multiscale Retinex (MSR), which combines multiple Gaussian scales to simultaneously capture both global and local illumination characteristics [6], [22]. The MSR formulation is given in Equation (2):

$$R_{MSR}(x, y) = \sum_{k=1}^K w_k R_{SSR_k}(x, y) \quad (2)$$

Where  $K$  denotes the number of Gaussian scales,  $w_k$  is the weight assigned to each scale with  $\sum w_k$ , and  $R_{SSR_k}$  represents the SSR output at the  $k$ -th scale. To mitigate the color distortion commonly introduced by MSR, the method is combined with a Color Restoration Function (CRF), forming Multiscale Retinex with Color Restoration (MSRCR). The CRF is defined in Equation (3) [15], [16], [19]:

$$CRF(x, y) = \beta \left[ \log(\alpha I(x, y)) - \log \left( \sum_c I_c(x, y) \right) \right] \quad (3)$$

Where  $\alpha$  and  $\beta$  are gain parameters, and  $I_c(x, y)$  denotes the intensity of the  $c$ -th color channel at pixel location  $(x, y)$ .

Several studies have reported that Retinex-based methods and their variants provide significant illumination stability for deep learning-based face detection systems [2], [8], [16]. Given that RetinaFace heavily relies on consistent edge cues and gradient distributions [17], the preprocessing pipeline in this study begins with MSRCR. The underlying hypothesis is that Retinex-based illumination recovery yields a more stable reflectance representation, thereby enhancing the RetinaFace feature extractor's sensitivity under low-light conditions.

## 2.2. Contrast Limited Adaptive Histogram Equalization (CLAHE)

Contrast Limited Adaptive Histogram Equalization (CLAHE) is a local contrast enhancement method that improves on Adaptive Histogram Equalization (AHE) [11], [18], [24]. While AHE can enhance micro-level details, it often leads to excessive noise amplification in dark regions. CLAHE addresses this issue by limiting the histogram through a clip limit parameter, ensuring that contrast enhancement remains controlled. In CLAHE, the image is divided into several local grids, and the histogram of each grid is clipped using a predefined threshold before intensity redistribution. The intensity transformation in CLAHE can be expressed by Equation (4) [11]:

$$I_{CLAHE}(x, y) = CDF_{\gamma}(I(x, y)) \cdot (L - 1) \quad (4)$$

Where  $CDF_{\gamma}$  denotes the cumulative distribution function with clip limit  $\gamma$ , and  $L$  represents the number of intensity levels.

Numerous studies have shown that CLAHE is effective at highlighting facial structures such as the contours of the nose, lips, and eye regions [15], [25], [26]. In the proposed pipeline, CLAHE is applied to the luminance (Y) channel to avoid color distortion. The underlying hypothesis is that local contrast enhancement from CLAHE will strengthen facial spatial structures, thereby facilitating RetinaFace's recognition of critical facial features.

## 2.3. Adaptive Gamma Correction

Adaptive Gamma correction is a nonlinear transformation aimed at adjusting image luminance based on global intensity characteristics [13]. In low-light conditions, adaptive gamma correction is used to enhance dark regions without over-enhancement. The adaptive gamma value can be derived from the image histogram distribution, as formulated in Equation (5) [13], [16]:

$$\gamma = 1 - CDF(\mu) \quad (5)$$

Where  $\mu$  denotes the mean intensity value of the image. The gamma transformation is then given by Equation (6):

$$I'(x, y) = 255 \left( \frac{I(x, y)}{255} \right)^{\gamma} \quad (6)$$

Where  $I(x, y)$  and  $I'(x, y)$  denote the pixel intensities before and after gamma transformation, respectively.

The literature indicates that gamma correction is effective in stabilizing luminance prior to contrast enhancement stages [2], [8], [27]. In this pipeline, gamma correction is applied to normalize the global intensity distribution before CLAHE, with the hypothesis that this initial normalization helps prevent oversaturation and preserve subtle facial details crucial for RetinaFace detection.

## 2.4. Non-Local Means (NL-Means) Denoising

Non-Local Means (NL-Means) is a patch-similarity-based denoising technique that exploits global image information. Unlike Gaussian smoothing, which considers only local neighborhoods, NL-means compares patches across the entire image to preserve important textures, making it well suited for low-light images that often suffer from high sensor noise [15]. The basic NL-means model is expressed in Equation (7) :

$$NL[v](x) = \sum_{y \in \Omega} w(x, y)v(y) \quad (7)$$

Where  $v(y)$  denotes the pixel intensity at location  $y$ ,  $\Omega$  represents the search domain, and  $w(x, y)$  is the similarity weight defined in Equation (8):

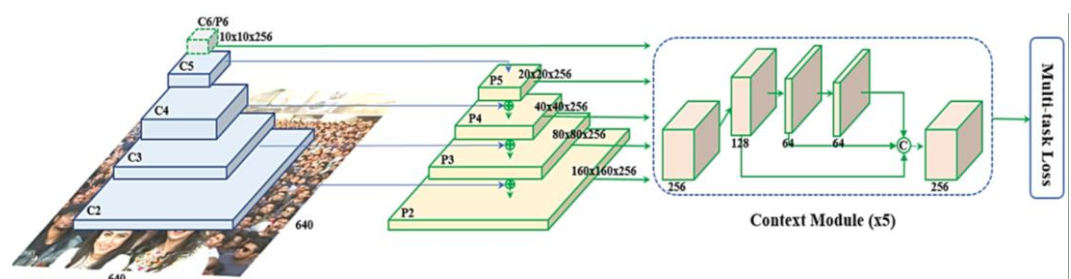
$$w(x, y) = \frac{1}{Z(x)} \exp \left( - \frac{\|v(N_x) - v(N_y)\|^2}{h^2} \right) \quad (8)$$

Where  $N_x$  and  $N_y$  denote patches centered at pixels  $x$  and  $y$ , respectively,  $h$  is the smoothing parameter, and  $Z(x)$  is a normalization factor.

Several recent studies have emphasized that denoising plays an important role in preserving facial textures prior to the detection stage [7]. Therefore, NL-means is placed at the final stage of the proposed pipeline to reduce noise that may be amplified during Retinex processing, gamma correction, and CLAHE. The underlying hypothesis is that NL-means can effectively suppress noise while preserving critical facial details for RetinaFace detection.

## 2.5. RetinaFace Detector

RetinaFace is a single-stage face detector built upon a Feature Pyramid Network (FPN) architecture, designed to simultaneously localize face bounding boxes and facial landmarks [17]. As shown in Figure 1, the model employs multi-scale feature pyramids (P2–P6), dense anchors, and independent context modules, enabling robust detection across a wide range of face scales and poses. Owing to this design, RetinaFace has been widely recognized as one of the most accurate and stable face detection frameworks in modern face detection research [1], [14], [25].



**Figure 1.** RetinaFace single-stage face detection architecture based on FPN with multi-scale features and context modules [17].

Several studies have reported that RetinaFace performance decreases when applied to low-light images [2], [8], [14]. The degradation is commonly associated with reduced gradient information, increased sensor noise, and loss of fine facial texture details under insufficient illumination. These factors negatively affect convolutional feature extraction, as CNN-based detectors rely heavily on edge cues and intensity distributions for reliable face localization. Previous investigations therefore, describe RetinaFace as a strong detection backbone whose performance under low-light conditions is closely related to the visual quality of the input images.

## 2.6. Related Works

Early research on low-light image enhancement was largely influenced by histogram-based methods, with Histogram Equalization (HE) being a classical approach for improving global contrast [28]. Although effective under certain conditions, numerous studies have reported that HE often leads to over-enhancement, noise amplification, and loss of important structural information when intensity distributions are highly uneven [15]. To address these limitations, Pizer et al. developed Adaptive Histogram Equalization (AHE) and subsequently Contrast Limited Adaptive Histogram Equalization (CLAHE), which introduces a clip limit mechanism to suppress excessive noise amplification [11], [18]. The effectiveness of CLAHE was further reinforced by Singh et al. [26], who reported that this method produces more natural local details, particularly in medical imaging. However, several surveys in [2], [15] emphasize that histogram-based approaches remain insufficiently adaptive to extreme illumination variations.

These limitations motivated the development of Retinex-based approaches, which mimic the Human Visual System (HVS) by separating illumination and reflectance components. Land [19] introduced Retinex theory as a foundation of visual perception, which was later extended by Jobson et al. through Multiscale Retinex (MSR) and MSRCR to balance global and local details while preserving color consistency [6]. Although effective, the literature indicates that Retinex-based methods may introduce color distortion and noise amplification in dark regions. Li et al. [21] addressed this issue by incorporating adaptive gamma correction, while the survey by Liu et al. [2] confirmed that Retinex remains a core foundation of many modern low-light enhancement techniques.

Another widely adopted approach is Low-Light Image Enhancement (LIME), introduced by Guo et al. [9], [20]. LIME estimates an illumination map to adaptively normalize lighting without requiring training data. Its primary advantages lie in low computational complexity and suitability for real-time applications. Subsequent studies, such as [29], reported that LIME effectively enhances visibility in dark regions. However, the literature also notes that LIME tends to produce color inconsistency and illumination artifacts when used as a standalone method, particularly in images with extreme illumination variation [2], [10].

In addition, luminance-based approaches such as adaptive gamma correction have been widely used due to their computational efficiency. Rahman [13] proposed a histogram-based adaptive gamma correction suitable for real-time applications. Nevertheless, several studies have shown that this method is less effective when applied independently, especially for images with severely uneven contrast and illumination [2], [21]. As demand for improved visual quality increased, several studies proposed image–fusion–based approaches. Mishra and Palkar [30] demonstrated that wavelet fusion and pyramid fusion can significantly enhance textural details, albeit at high computational cost. Singh et al. [26] combined CLAHE with wavelet fusion to improve structural detail; however, such fusion-based methods are less suitable for time-constrained systems such as video surveillance due to their computational overhead.

Advances in modern computer vision have also driven the emergence of deep learning–based low-light image enhancement methods. Semi-supervised approaches by Yang et al. [8] and the Deep Color Consistent Network proposed by Zhang et al. [14] significantly improve perceptual quality and color consistency. In addition, Zhang et al. [7] introduced residual learning for denoising low-intensity images. Despite achieving high-quality results, most deep learning–based methods require large-scale datasets, long training times, and substantial computational resources, making them less ideal for real-time applications with hardware constraints [2], [8], [14].

In the context of face detection, the literature consistently highlights illumination quality as a key factor affecting detector performance. Studies such as GhostFaceNets [5] and GhostNet [31] improve model efficiency but still experience significant performance degradation under low-light conditions. The survey by Zhao et al. [1] emphasizes that poor illumination leads to the loss of critical facial features essential for CNN-based feature extraction. Fu et al. [32] demonstrated that applying image enhancement prior to detection can significantly improve accuracy, while studies in [8] and [14] further confirmed the importance of preprocessing for enhancing the robustness of deep learning–based detectors.

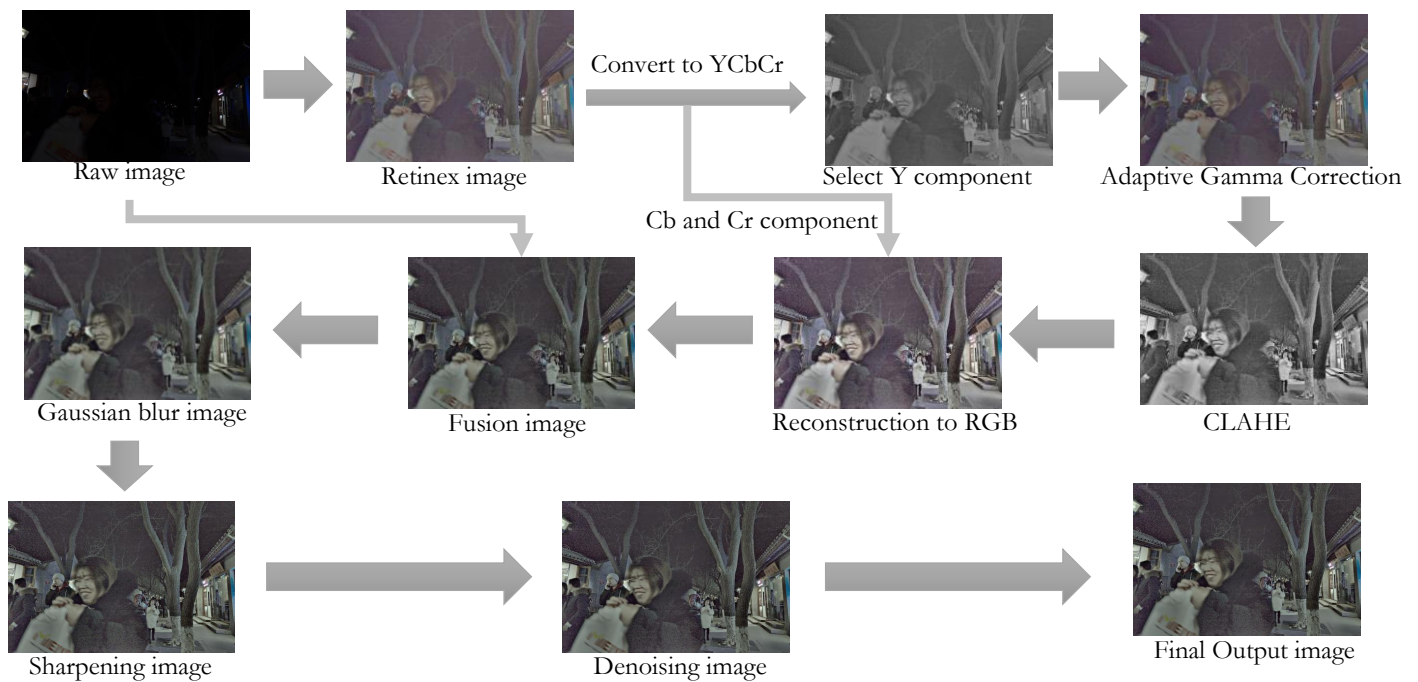
Although RetinaFace [17] is currently one of the most powerful face detectors, the literature indicates that integrating multi-technique enhancements—such as Retinex,

CLAHE, and adaptive gamma correction—into a unified preprocessing pipeline for RetinaFace has rarely been explored. Most existing studies focus on improving detector architectures rather than enhancing input image quality [1], [17], [31]. Moreover, objective visual quality assessment based on Human Visual System–oriented metrics such as NIQE and BRISQUE has become increasingly important for large-scale applications without reference images [2], [15].

Based on the above review, it can be concluded that no single method can comprehensively address all challenges associated with low-light images. Traditional and deep learning–based approaches each possess distinct strengths and limitations. Furthermore, research integrating multiple enhancement techniques into a unified pipeline as preprocessing for modern face detectors such as RetinaFace remains limited. This research gap motivates the development of LUMEN, an integrated enhancement approach designed to achieve more stable illumination, preserved textural details, and improved face detection accuracy under low-light conditions.

### 3. Proposed Method

This study proposes the Low-light Unified Multi-stage Enhancement Network (LUMEN) as an integrated preprocessing pipeline to enhance low-light image quality prior to



**Figure 2.** Detailed workflow of the LUMEN enhancement pipeline, showing sequential stages of illumination normalization, luminance enhancement, contrast adaptation, fusion, and noise suppression for low-light face images.

RetinaFace-based face detection. LUMEN is designed as a multi-stage approach that combines global–local illumination correction, adaptive luminance adjustment, local contrast enhancement, and noise and sharpness control in a sequential and coordinated manner. Unlike single enhancement methods that generally address only one specific visual aspect, LUMEN targets overall image quality stability. Thus, this pipeline not only supports more robust feature extraction in CNN-based detectors but also maintains image visual legibility for manual analysis. An overview of the LUMEN method flow is shown in Figure 2.

#### 3.1. Original Low-Light Image

The initial stage of LUMEN takes the raw low-light image as the baseline input. According to the literature [15], [33]–[35], low-light images are generally characterized by insufficient illumination, poor signal-to-noise ratio, and loss of structural details, particularly in facial regions. In LUMEN, the original image is not entirely discarded; instead, it is partially preserved for use in the fusion stage (Subsection 3.8). This strategy aims to maintain visual naturalness

and prevent overly artificial appearances that may arise from aggressive multi-stage enhancement.

### 3.2. Retinex (Adaptive MSRCR)

The theoretical foundation of Retinex has been discussed in Section 2.1; therefore, this subsection focuses on implementation details and parameterization. LUMEN employs an MSRCR approach using three Gaussian scales, namely  $\sigma \in \{15, 80, 250\}$ . The small scale  $\sigma = 15$  is used to emphasize fine-grained details, the medium scale  $\sigma = 80$  balances texture and contrast, while the large scale  $\sigma = 250$  is intended for global illumination correction. This multi-scale combination enables simultaneous recovery of local details and overall illumination normalization, as recommended in several previous studies [31], [36], [37].

### 3.3. Color Space Conversion to YCrCb

The Retinex-processed image is converted from the RGB color space to the YCrCb color space to separate luminance and chrominance components. The linear transformation used is given in Equation (9):

$$\begin{aligned} Y &= 0.299R + 0.587G + 0.114B, \\ Cr &= 0.713(R - Y), \\ Cb &= 0.564(B - Y). \end{aligned} \quad (9)$$

Selecting the YCrCb color space enables enhancement operations to focus on the luminance channel (Y), which is most affected in low-light conditions, while preserving the color components (Cr and Cb). This approach effectively prevents color distortion that often occurs when luminance and contrast enhancement are applied directly in the RGB space [38], [39]. Furthermore, luminance–chrominance separation in YCrCb enables more stable and controlled application of adaptive gamma correction and CLAHE. This strategy has been recommended in low-light image enhancement research to preserve color naturalness and visual consistency [21], [40], [41].

### 3.4. Adaptive Gamma Correction

At this stage, adaptive gamma correction (AGC) is applied to the luminance channel Y to normalize the global intensity distribution of low-light images before local contrast enhancement. Gamma correction is a non-linear transformation that adjusts pixel intensities according to the relationship in Equation (6). In LUMEN, the gamma value is adaptively selected from a predefined discrete set  $\gamma \in \{0.5, 0.7, 2.0\}$  according to the overall brightness characteristics of the input image. This design is motivated by the hypothesis that different low-light conditions require different luminance responses. Lower gamma values ( $\gamma = 0.5$ ) are used to aggressively lift severely underexposed regions, moderate values ( $\gamma = 0.7$ ) balance brightness enhancement and visual naturalness in mildly low-light images, while higher values ( $\gamma = 2.0$ ) compress overly bright or unevenly illuminated regions to prevent saturation. These values are selected based on established gamma correction behavior reported in prior studies [42], [43] and empirical stability testing, rather than continuous optimization. The gamma correction stage is applied before CLAHE to normalize the global intensity distribution, enabling subsequent local contrast enhancement to operate more effectively. Although adaptive gamma correction does not directly maximize detection accuracy, it contributes to perceptual consistency and helps prevent excessive contrast amplification in later stages.

### 3.5. Contrast Limited Adaptive Histogram Equalization (CLAHE)

CLAHE is applied to the luminance channel Y to enhance local facial details. The parameters used in LUMEN are clipLimit = 2.5 and tileGridSize =  $8 \times 8$ , following prior studies [2], [11], [26]. This configuration has been shown to effectively preserve critical facial structures—such as the contours of the eyes, nose, and lips—while avoiding excessive noise amplification commonly observed in conventional histogram-based methods.

### 3.6. Reconstruction to RGB Space

After luminance enhancement is completed, the image is reconstructed back into the RGB color space by merging the enhanced luminance channel ( $Y''$ ) with the original chrominance channels ( $Cr, Cb$ ), as expressed in Equation (10):

$$I_{RGB} = \text{merge}(Y'', Cr, Cb) \quad (10)$$

This step ensures that luminance enhancement does not compromise color accuracy, keeping the final image visually natural and color-accurate.

### 3.7. Gaussian Smoothing

Gaussian smoothing is applied to reduce high-frequency noise that is often amplified by the Retinex, gamma correction, and CLAHE stages. The Gaussian filter is defined as:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (11)$$

In LUMEN,  $\sigma = 2$ , is used, which according to [7], [15] provides an optimal balance between noise smoothing and detail preservation. This stage aims to stabilize the image before sharpening, preventing noise from being unintentionally amplified.

### 3.8. Fusion of Enhanced and Original Images

The fusion stage is introduced to control visual naturalness and prevent overly contrast-enhanced appearances. LUMEN combines the enhanced image and the original image using the following formulation:

$$I_f = \alpha I_{enh} + (1 - \alpha) I_{orig} \quad (12)$$

With  $\alpha = 0.7$ , this approach preserves enhancement-induced details while maintaining the original image's natural color and texture characteristics, as recommended in image fusion studies [26], [30].

### 3.9. Sharpening

Sharpening is performed using the unsharp masking method to enhance facial edge structures, as defined by Equation (13). Sharpening is applied after the fusion stage so that edge enhancement operates on an image that is already stabilized and minimally affected by noise [44], [45]. This step improves gradient definition, which is crucial for feature extraction in RetinaFace's convolutional layers.

$$I_{sharp} = I_f + k(I_f - G_\sigma * I_f) \quad (13)$$

### 3.10. Denoising Using Fast Non-Local Means

The final stage of LUMEN applies denoising using Fast Non-Local Means (NLM) to remove residual noise remaining after sharpening. The parameters follow stable configurations reported in the literature [7]:  $h = 7$ , template size = 7, and search window = 15. This approach enables adaptive noise reduction without removing important facial texture details, ensuring that the final image remains sharp, natural, and well-suited for RetinaFace-based face detection.

### 3.11. Perceptual Image Quality Assessment

This study conducts image quality assessment using no-reference image quality assessment (NR-IQA) metrics, namely NIQE (Natural Image Quality Evaluator) and BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator) [15], [46]. These two metrics are widely used in image enhancement research because they can evaluate visual quality without requiring reference images, making them particularly suitable for low-light scenarios where ground truth images are generally unavailable [8].

NIQE evaluates image quality based on deviations of statistical characteristics of image patches from a statistical model of natural images. In general, NIQE models natural images using a Multivariate Gaussian (MVG) constructed from luminance-based statistical features. The NIQE score is computed as the distance between the statistical distribution of the test image and that of natural images, which can be conceptually expressed as:

$$NIQE = \sqrt{(\mu_n - \mu_i)^T \sum_n^{-1} (\mu_n - \mu_i)} \quad (14)$$

Where  $\mu_n$  and  $\sum_n$  denote the mean and covariance of the natural image model, respectively, and,  $\mu_i$  represents the mean of the statistical features of the enhanced image.

Lower NIQE values indicate that the luminance distribution and global structure of the enhanced image are closer to those of natural images [2], [15], [46]. This metric is sensitive to global distortions, illumination instability, and intensity artifacts that often arise from aggressive enhancement, such as those produced by global histogram-based methods [47], [48]. BRISQUE focuses on evaluating spatial structure and local texture using Mean Subtracted Contrast Normalized (MSCN) coefficients. The MSCN coefficients are computed from the luminance image as follows:

$$\hat{I}(x, y) = \frac{I(x, y) - \mu(x, y)}{\sigma(x, y) + C} \quad (15)$$

Where  $I(x, y)$  is the pixel intensity,  $\mu(x, y)$  and  $\sigma(x, y)$  are the local mean and standard deviation, respectively, and  $C$  is a small constant for numerical stability.

The distributions of MSCN coefficients and their neighboring pixel pairs are then extracted as features and mapped to a quality score using a pre-trained Support Vector Regression (SVR) model [49]. Lower BRISQUE values indicate that texture, edges, and noise levels are more natural and freer from artifacts such as edge halos or excessive high-frequency amplification. Therefore, NIQE and BRISQUE are jointly adopted to assess whether the proposed enhancement improves perceptual quality while preserving natural image characteristics in low-light conditions.

### 3.12. Performance Assessment

The performance of the proposed LUMEN pipeline is evaluated from two complementary perspectives: perceptual image quality and face detection robustness under low-light conditions. This dual evaluation is necessary because improvements in visual appearance do not always correspond to improved detection performance, particularly in severely underexposed environments [50]. Accordingly, both aspects are jointly analyzed to assess LUMEN's effectiveness as a preprocessing pipeline rather than as a detector-optimization method. Face detection performance is quantitatively assessed using the detection rate and average confidence score produced by RetinaFace. These metrics are intentionally selected to reflect detection robustness and reliability in scenarios where illumination degradation is the primary challenge. The detection rate measures the proportion of ground-truth faces that are successfully detected across the dataset and is defined as:

$$\text{Detection Rate}(\%) = \frac{\sum_{i=1}^N D_i}{\sum_{i=1}^N G_i} \times 100 \quad (16)$$

Where  $G_i$  denotes the number of ground-truth faces in the  $i$ -th image,  $D_i$  represents the number of correctly detected faces in the same image, and  $N$  is the total number of test images. This metric provides an intuitive, recall-oriented measure of detection robustness, particularly suitable for low-light scenarios where the dominant failure mode is missed detections rather than false positives. Consequently, the detection rate should not be interpreted as classification accuracy, but rather as an indicator of the detector's ability to recover facial instances that are visually degraded or partially obscured under insufficient illumination [51].

In addition to detection rate, the average confidence score is used to evaluate the reliability and stability of detected faces. For each detected face, RetinaFace outputs a confidence value indicating the likelihood that the predicted bounding box corresponds to a true facial instance [17], [52]. The average confidence score is computed as:

$$\text{Average Confidence}(\%) = \frac{1}{M} \sum_{j=1}^M c_j \quad (17)$$

Where  $c_j$  denotes the confidence score of the  $j$ -th detected face, and  $M$  is the total number of detected faces across all test images. A higher average confidence score indicates more stable detector responses and reduced ambiguity in feature extraction, reflecting improved facial visibility after enhancement.

It is important to emphasize that bounding-box-level metrics such as precision, F1-score, and mean Average Precision (mAP) are not reported in this study. This decision is motivated by the characteristics of the evaluation setting, in which the low-light subset used provides only face-count ground truth, without the detailed bounding-box annotations required for accurate computation of localization-based metrics. Moreover, the primary objective of this work is not to benchmark or retrain the RetinaFace detector, but to analyze how low-light enhancement affects detection robustness when the detector is used in a zero-shot, off-the-shelf configuration. Under these conditions, recall-oriented metrics and confidence analysis are more appropriate and methodologically consistent with the research objective.

### 3.13. Algorithmic Description of LUMEN Enhancement

To clarify the implementation and ensure the reproducibility of the proposed method, the LUMEN processing flow is summarized in pseudocode in Algorithm 1. This algorithm describes the complete sequence of preprocessing stages, starting from the raw low-light image to the final image ready for face detection. Each step in the algorithm directly represents the methodological steps described in Sections 3.1–3.10.

---

#### Algorithm 1. LUMEN Enhancement Pipeline

---

INPUT: `img_raw`: low-light image in BGR color space

OUTPUT: `img_out`: enhanced image

```

1: img ← img_raw.astype(float) + 1 # avoid log(0) in Retinex
2: img_ret ← MSRCR_adaptive(img, scales = {15, 80, 250})
3: yrcb ← convert_color(img_ret, BGR → YCrCb)
4: (Y, Cr, Cb) ← split(yrcb)
5: Y_gamma ← adaptive_gamma_correction(Y)
6: Y_clahe ← CLAHE_Y(Y_gamma, clipLimit = 2.5, gridSize = 8×8)
7: yrcb_enh ← merge(Y_clahe, Cr, Cb)
8: img_enh ← convert_color(yrcb_enh, YCrCb → BGR)
9: img_fusion ←  $0.7 \times \text{img\_enh} + 0.3 \times \text{img\_raw}$ 
10: img_blur ← GaussianBlur(img_fusion,  $\sigma = 2$ )
11: img_sharp ←  $\text{img\_fusion} + 0.3 \times (\text{img\_fusion} - \text{img\_blur})$ 
12: img_out ← FastNonLocalMeans(
13:     img_sharp,
14:     h = 7,
15:     hColor = 7,
16:     templateWindowSize = 7,
17:     searchWindowSize = 15)
18: return img_out

```

---

## 4. Results and Discussion

### 4.1. Experimental Setup

The experiments were conducted on a system equipped with an AMD Ryzen 7 6800H processor, 16 GB of RAM, and an NVIDIA RTX 3050 GPU, running Windows 11 Pro 64-bit. The implementation was developed in Python 3.12 using OpenCV 4.9.0 for all image enhancement stages in the proposed LUMEN pipeline, including adaptive Retinex, adaptive gamma correction, CLAHE, fusion, sharpening, and denoising. NumPy 1.26 was used for numerical operations and array manipulation, while Matplotlib 3.8 was employed for result visualization and graphical analysis.

The RetinaFace model (version 1.1.1) was used as the primary face detector in all experiments. RetinaFace was employed as an off-the-shelf face detection framework using the default pretrained configuration provided by the official implementation. No architectural modifications, parameter tuning, or threshold adjustments were applied, ensuring that all observed

performance differences can be attributed solely to the proposed LUMEN preprocessing pipeline rather than detector configuration. In addition, the Tkinter module was utilized to build a simple graphical user interface (GUI), and the math and time libraries were used for statistical calculations and execution time measurements.

The dataset used in this study is the Human Faces Object Detection Dataset obtained from the Kaggle platform, which contains 2,204 images covering a wide range of lighting conditions, facial poses, numbers of faces per image, and background complexity. As this study focuses on face detection under challenging illumination conditions, only the low-light subset of the dataset was used. The dataset is publicly available at <https://flywh.github.io/CVPRW2019LowLight/>. From the complete dataset, 30 representative low-light images were selected for experimental evaluation, with the number of faces per image ranging from one to five. The selected low-light images are treated as an evaluation set for inference-only analysis, since the RetinaFace detector is applied in its pretrained off-the-shelf configuration without any retraining or parameter adaptation. These selected images are available at <https://github.com/Rivaldisk73/Image-Enhancement-Low-Light.git>.

The selected low-light images exhibit common degradation characteristics, including low average luminance (Y channel values below 70 on an 8-bit scale), noticeable noise in shadow regions, color casts caused by artificial lighting, and multiple faces at varying scales within a single image. To ensure a balanced evaluation across different detection complexities, the image distribution was controlled so that five samples represented each category (one to five faces). Representative examples of the selected low-light images are shown in Figure 3.



**Figure 3.** Representative low-light images from the dataset with varying numbers of faces: (a) one face; (b) two faces; (c) three faces; (d) four faces; (e) five faces.

In developing the LUMEN method, parameter selection was guided by prior image enhancement literature and empirical stability testing. As LUMEN is a deterministic multi-stage

image enhancement pipeline rather than a learning-based model, no loss-based optimization or training was performed. Instead, commonly reported parameter ranges were first identified from previous studies, and representative values were empirically evaluated to ensure stable illumination correction, controlled noise behavior, and reproducible face detection performance. The final parameter settings prioritize visual stability, artifact avoidance, and consistent performance across low-light images rather than dataset-specific tuning.

**Table 1.** Main parameters of the LUMEN enhancement method

Component	Selected Parameters	Recommended Range	Remarks
MSR Scales	[15, 80, 250]	Small $\rightarrow$ large scales	Weights ( $w_k$ ) are positive and satisfy ( $\sum w_k = 1$ ).
CLAHE	clipLimit = 2.5, tile = (8 $\times$ 8)	1.0 – 4.0	Higher values increase local contrast.
Adaptive Gamma	Adaptive $\gamma \in$ [0.5, 0.7, 2.0]	0.5 – 2.0	Discrete adaptive selection based on image brightness to balance enhancement and artifact suppression.
Fusion $\alpha$	$\alpha = 0.7$	0.5 – 0.9	Corresponds to 70% enhanced image and 30% original image.
Gaussian $\sigma$	$\sigma = 2.0$	0.5 – 3.0	Applied for smoothing prior to sharpening.
Unsharp $\kappa$	$\kappa = 0.3$	0.1 – 0.6	Larger values may cause oversharpening.
NL-Means (h / hColor)	7 / 7	3 – 15	Higher values reduce noise but may blur fine details.

#### 4.2. Quantitative and Qualitative Analysis of Visual Enhancement Results

Table 2 and Figure 4 present a combined quantitative and qualitative comparison of different low-light image enhancement methods, including raw input, CLAHE, Histogram Equalization (HE), LIME [20], and the proposed LUMEN pipeline. All methods were independently re-implemented and evaluated under identical experimental conditions to ensure a fair and reproducible comparison. The analysis integrates no-reference image quality metrics (NIQE and BRISQUE) with visual inspection to assess enhancement stability, perceptual naturalness, and suitability for face detection under low-light conditions.

**Table 2.** NIQE and BRISQUE of LUMEN results and comparison.

Method	NIQE $\downarrow$	BRISQUE $\downarrow$
Raw	31.209	7.394
CLAHE	4.271	7.225
HE	2860.526	13.798
LIME	863.405	17.216
LUMEN	45.656	6.856

The raw low-light images are severely underexposed and noisy, resulting in obscured facial structures and unreliable visual information. This degradation is reflected by a high NIQE value (31.209), indicating substantial deviation from natural image statistics, confirming that raw images are unsuitable for robust face detection. Among histogram-based approaches, CLAHE demonstrates the most stable performance. Its low NIQE (4.271) and BRISQUE (7.225) values indicate well-preserved global luminance distribution and local texture structure. Visual inspection confirms that CLAHE enhances facial contrast in a controlled manner while avoiding severe color distortion, making it a strong baseline for low-light enhancement.

In contrast, HE and LIME exhibit significant instability. HE produces extreme global contrast distortion, as evidenced by an exceptionally high NIQE score (2860.526) and visible over-amplification artifacts. LIME improves brightness more aggressively but introduces strong color casts, halo effects, and amplified noise, resulting in high NIQE (863.405) and BRISQUE (17.216) values. These results indicate that single-stage aggressive enhancement

strategies are insufficient to maintain perceptual quality in complex low-light scenes. The proposed LUMEN pipeline achieves the most balanced performance across all evaluation criteria. Although its NIQE value (45.656) is higher than that of CLAHE, it remains substantially lower than those of HE and LIME, indicating controlled global luminance deviation. More importantly, LUMEN attains the lowest BRISQUE score (6.856), demonstrating superior preservation of local textures, effective noise suppression, and minimal artifact introduction. Visual results further confirm that LUMEN enhances facial regions while maintaining natural color tones and stable background appearance.



**Figure 4.** Visual comparison of low-light face image enhancement methods (a) raw image; (b) CLAHE; (c) HE; (d) LIME; (e) LUMEN

Overall, the results reveal a consistent pattern: aggressive single-stage methods tend to improve brightness at the expense of perceptual quality, whereas adaptive multi-stage approaches provide more stable and natural enhancement. While CLAHE excels in luminance statistical stability, LUMEN offers the most balanced trade-off between illumination enhancement, texture preservation, and robustness for face detection, making it well-suited for practical low-light face recognition and surveillance applications.

#### 4.3. Detection Performance Analysis after Enhancement

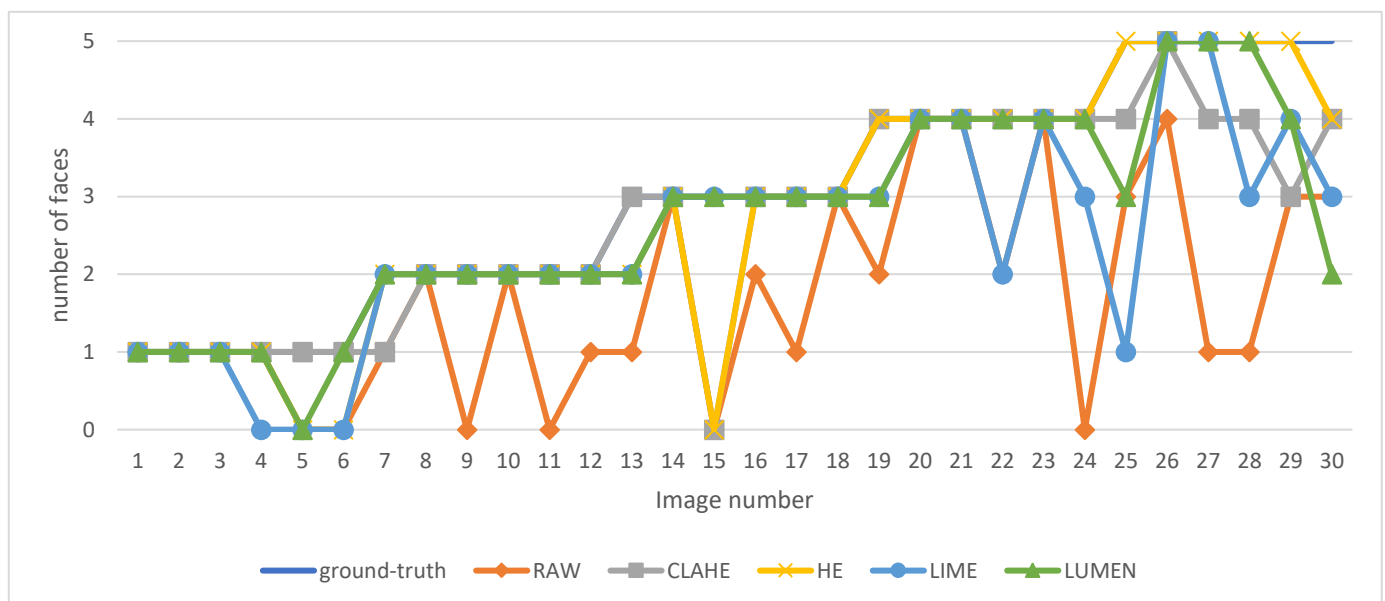
This subsection evaluates the impact of different enhancement methods on face detection performance under low-light conditions using detection-oriented metrics, namely detection rate and average confidence score produced by RetinaFace. The quantitative comparison of detection performance is summarized in Table 3, while Figures 5 and 6 illustrate the detection behavior across images with varying numbers of faces.

As shown in Table 3, raw low-light images yield the poorest detection performance, achieving a detection rate of only 58% with a relatively low average confidence score of

0.7811. This result confirms that insufficient illumination severely limits the detector's ability to extract reliable facial features, underscoring the need for image enhancement as a preprocessing step in low-light scenarios.

**Table 3.** Detection performance comparison.

Method	Detection Rate (%)	Confidence
Raw	58	0.7811
CLAHE	91	0.9565
HE	88	0.8926
LIME	80	0.8895
LUMEN	91	0.9545



**Figure 5.** Face detection performance comparison across different enhancement methods (Raw, CLAHE, HE, LIME, and LUMEN) against the ground-truth number of faces per image.

Among the enhancement methods, CLAHE achieves the strongest overall detection performance, reaching the highest detection rate of 91% and the highest confidence score (0.9565). This indicates that locally adaptive contrast enhancement effectively improves facial visibility while preserving discriminative features required by the detector. Histogram Equalization (HE) also improves the detection rate to 88%, but its lower confidence score (0.8926) suggests less stable feature reliability. Similarly, LIME achieves a moderate detection rate of 80% with a confidence score of 0.8895, indicating that illumination-map-based enhancement alone does not consistently support robust face detection under complex low-light conditions.

Notably, the proposed LUMEN pipeline matches the highest detection rate achieved by CLAHE (91%) while maintaining a comparably high confidence score of 0.9545. This demonstrates that the multi-stage enhancement strategy effectively enhances detection performance without compromising detector stability. As illustrated in Figure 5, both LUMEN and CLAHE exhibit consistent detection behavior across images with varying numbers of faces, closely matching the ground-truth face counts. In contrast, raw images exhibit large detection fluctuations, while HE and LIME show increased variability as scene complexity increases.

Overall, the results in Table 3 and Figures 5–6 reveal a clear trend: enhancement methods based on aggressive global transformations or single-stage illumination correction can improve visibility but do not consistently translate into stable detection performance. In contrast, CLAHE and the proposed LUMEN pipeline provide the most reliable detection outcomes. While CLAHE achieves the highest confidence score, LUMEN offers a more balanced solution by delivering comparable detection accuracy with improved robustness across

varying face counts and illumination conditions, reinforcing its suitability for practical low-light face detection applications.



**Figure 6.** Detection with bounding box comparison of low-light face image enhancement methods (a) raw image; (b) CLAHE; (c) HE; (d) LIME; (e) LUMEN

#### 4.4. Ablation Study

An ablation study was conducted to analyze the contribution of each component in the LUMEN pipeline to face detection performance and visual quality under low-light conditions. In each experiment, one major module was removed from the full configuration (Full LUMEN), and the resulting system was evaluated using detection rate, confidence score, and no-reference image quality metrics, namely NIQE and BRISQUE. A summary of the quantitative results is presented in Table 4, while more detailed per-image comparisons are illustrated in the graph shown in Figure 7.

##### 4.5.1. Quantitative Ablation Results

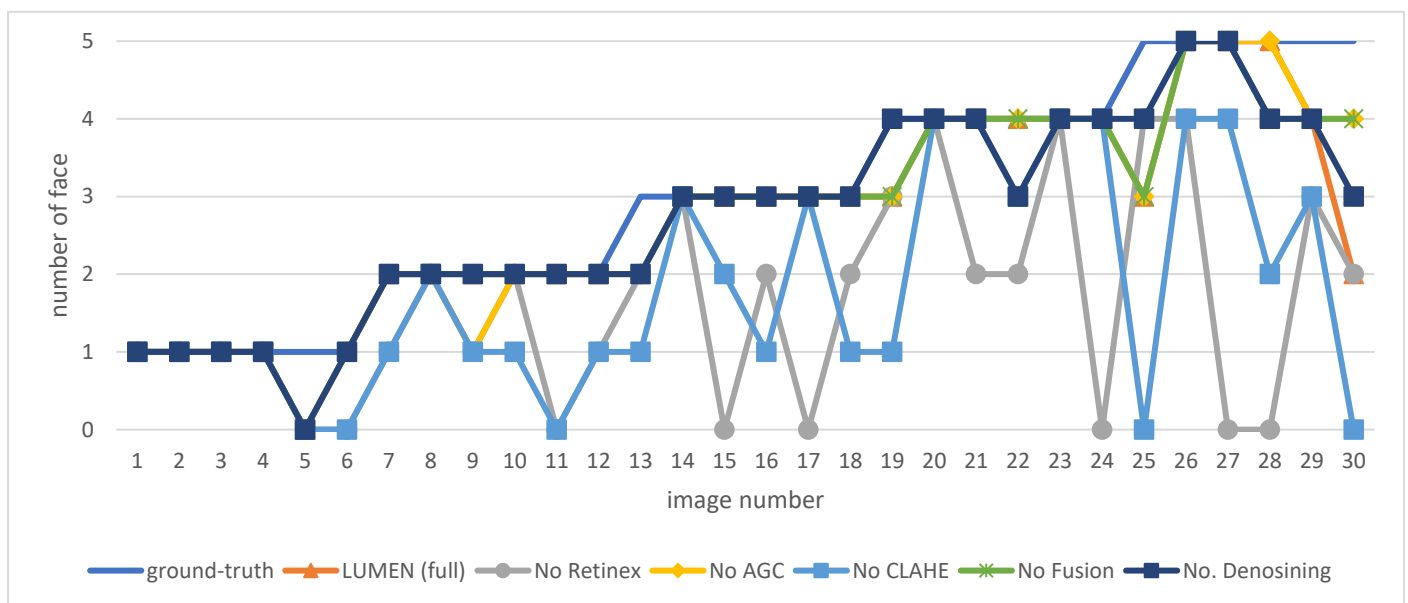
The results in Table 4 indicate that Retinex and CLAHE are the two most critical components in the LUMEN pipeline. When Retinex is removed, the detection rate drops drastically to 56%, accompanied by a significant decrease in confidence. This finding confirms that global illumination normalization is a fundamental prerequisite for enabling CNN backbones to recognize facial structures under low-light conditions.

Removing CLAHE also results in a clear performance degradation, with the detection rate decreasing to 61%. This suggests that local contrast enhancement plays a significant role in enhancing micro-texture and edge cues that are important to the detector. In contrast, removing Adaptive Gamma Correction (AGC) does not significantly affect the detection rate or confidence score, but it causes a substantial increase in the NIQE value. This suggests that

AGC primarily contributes to stabilizing luminance distribution and perceptual quality rather than directly improving detection accuracy.

**Table 4.** Ablation performance results

Configuration	Detection Rate (%)	Confidence	NIQE ↓	BRISQUE ↓
No Retinex	56	0.7206	467.576	4.443
No AGC	90	0.9548	186.567	7.056
No CLAHE	61	0.8201	97.719	5.915
No Fusion	91	0.9557	153.201	6.964
No Denoising	91	0.9563	8.528	7.655
LUMEN (Full)	91	0.9545	45.656	6.856



**Figure 7.** Face detection performance comparison across ablation against the ground-truth number of faces per image.

Eliminating the fusion stage does not affect detection performance quantitatively; however, both NIQE and BRISQUE values increase. This confirms that fusion functions as a regularization mechanism to preserve naturalness and prevent artifacts caused by aggressive enhancement, rather than acting as a feature-strengthening component for detection. In the No Denoising configuration, both the detection rate and confidence remain high, while the NIQE score drops to an extremely low level. However, the increase in BRISQUE indicates that textures become rougher. This finding highlights that denoising is not critical for detection accuracy but plays an important role as a visual stabilizer, controlling residual noise and maintaining perceptual image quality.

#### 4.5.2. Visual Analysis of Ablation Configurations

The visual analysis in Figure 8(a–f) provides insight into the functional role of each module in the LUMEN pipeline. The full configuration (Figure 8(f)) exhibits balanced illumination, well-defined facial contours, and controlled noise levels, indicating an effective balance between enhancement strength and perceptual naturalness.

Removing the Retinex module (Figure 8(a)) results in severe visual degradation, with a dark appearance and limited dynamic range that suppresses facial gradients. This explains the detection failures observed quantitatively and confirms that global illumination normalization is a fundamental component of LUMEN. Excluding AGC (Figure 8(b)) results in a less natural tonal distribution, with locally over-darkened or over-brightened regions, although facial structures remain recognizable. This indicates that AGC primarily contributes to luminance balancing and perceptual consistency rather than detection-critical enhancement.

Without CLAHE (Figure 8(c)), local contrast is noticeably reduced, particularly in fine facial textures, resulting in flatter and less distinct facial structures. This observation aligns with the reduced detection performance and highlights the importance of adaptive contrast enhancement for reinforcing facial micro-structures. Removing the fusion stage (Figure 8(d)) results in more aggressive enhancement and localized artifacts, thereby reducing visual naturalness. This confirms fusion's role as a regularization mechanism that controls enhancement intensity and preserves coherence with the original image.



**Figure 8.** Visual comparison of low-light face image enhancement ablation (a) no retinex; (b) no AGC; (c) no CLAHE; (d) no fusion; (e) no denoising; (f) full configuration (LUMEN)

When denoising is excluded (Figure 8(e)), the global appearance remains comparable to the full configuration (Figure 8(f)); however, higher levels of high-frequency noise are clearly visible in shadowed background regions. While face detection performance remains largely unchanged between the two configurations, the full LUMEN pipeline in Figure 8(f) exhibits smoother textures and reduced noise artifacts. This difference is reflected in lower BRISQUE scores for the full configuration, confirming that denoising primarily contributes to perceptual stability rather than detection accuracy. Although its absence does not significantly affect face detection outcomes, denoising plays an important role in improving visual smoothness and practical usability for human-centered applications such as surveillance and monitoring.

#### 4.5.3. Discussion of Ablation Findings

The ablation study results demonstrate that LUMEN's effectiveness does not rely on any single component, but rather on the synergistic interactions among its modules. Table 5 summarizes the primary functional role of each component, the observed impact when removed, and the corresponding interpretation, providing a concise overview of each stage's contribution within the pipeline.

**Table 5.** Summary of ablation findings and component contributions.

Component	Primary role in pipeline	Impact when removed	Interpretation
Retinex	Global illumination normalization and dynamic range recovery	Detection collapses; facial gradients fail to form	Most critical component; foundational for all subsequent enhancement stages
CLAHE	Local contrast enhancement and micro-structure amplification	Significant drop in detection robustness; flat facial textures	Essential for strengthening discriminative local facial features in low-light
AGC	Luminance redistribution and tonal balancing	Minor impact on detection; perceptual inconsistency increases	Supports perceptual naturalness rather than direct detection accuracy
Fusion	Regularization of enhancement intensity and artifact suppression	Over-processed appearance; tonal imbalance	Acts as a visual regularizer to prevent excessive enhancement
Denosing	High-frequency noise suppression and texture smoothing	Detection unaffected; visual roughness increases	Serves as a perceptual stabilizer rather than a detection-critical module

Retinex emerges as the most essential component because it directly restores global illumination and expands the image's dynamic range. When this module is removed, the pipeline loses its ability to form meaningful facial gradients, leading CNN-based detectors to fail to recognize facial structures consistently. This finding confirms that global illumination normalization constitutes an absolute foundation for the entire enhancement process under low-light conditions. CLAHE ranks second in importance, particularly for enhancing local contrast and reinforcing facial microstructures. Although removing CLAHE does not always result in total detection failure, the observed degradation in stability and visual quality indicates that adaptive contrast plays a critical role in improving the discriminability of facial features in low-light scenarios with weak textures.

The AGC and fusion exert a stronger influence on perceptual quality than on detection accuracy. AGC stabilizes the luminance distribution to prevent regions from becoming excessively dark or bright, while fusion controls the aggressiveness of enhancement. As summarized in Table 5, neither module directly increases the detection rate; however, both are crucial for maintaining visual naturalness and consistency in the enhanced output. The denosing module demonstrates a distinct characteristic. Quantitatively, its removal does not reduce detection performance; however, visually, it produces images with rougher textures and more pronounced high-frequency noise. This observation indicates that certain metrics, such as NIQE, may yield lower values due to sharpened micro-structures, yet do not necessarily reflect superior visual quality. Consequently, denosing serves as an important visual stabilizer, enhancing perceptual comfort and practical acceptability in real-world applications.

Overall, these ablation findings confirm that LUMEN's primary strength lies in the full integration of its components. No single module can completely substitute for another. Retinex and CLAHE determine detection success, while AGC, fusion, and denosing ensure that the enhancement results remain natural, stable, and free from disturbing artifacts. This integrated multi-stage approach positions LUMEN as a reliable, practical preprocessing solution for modern face detection systems operating under low-light conditions.

## 5. Conclusions

This study proposed LUMEN, a multi-stage image enhancement pipeline designed to address the challenges of face detection under low-light conditions. The pipeline integrates global illumination normalization (Retinex), adaptive gamma correction, local contrast enhancement (CLAHE), controlled fusion, and noise suppression to produce visually stable and detector-friendly images. The primary objective was to improve face detection reliability while preserving perceptual naturalness, which is often compromised by aggressive enhancement techniques. Experimental results demonstrate that LUMEN consistently improves face detection performance, achieving a detection rate of 91% with a high average confidence score (0.9545) when evaluated using RetinaFace. Compared to raw images and baseline enhancement methods such as HE and LIME, LUMEN provides more stable detection outcomes

while avoiding excessive noise amplification and color distortion. These findings directly support the research objective of balancing illumination enhancement with structural and perceptual stability.

Ablation experiments further reveal that LUMEN's effectiveness arises from the synergistic interaction of its components rather than any single module in isolation. Retinex and CLAHE are identified as the most critical stages for successful detection, as they restore global illumination and reinforce local facial structures. Meanwhile, adaptive gamma correction, fusion, and denoising primarily contribute to perceptual consistency and artifact suppression, ensuring that enhancement remains natural and visually acceptable. This modular analysis highlights the importance of an integrated, multi-stage design for low-light preprocessing. From a broader perspective, this work contributes to the field of low-light computer vision by demonstrating that carefully designed traditional enhancement pipelines remain highly competitive and practical for supporting modern deep-learning-based face detectors. LUMEN provides a model-agnostic preprocessing solution that can be readily applied to real-world scenarios such as surveillance, access control, and outdoor monitoring under suboptimal lighting.

Despite its effectiveness, this study has several limitations. The evaluation was conducted on a relatively small subset of low-light images, and experiments focused on a single face detection model. Future work may explore adaptive parameter tuning, validation on larger and multi-sensor datasets, integration with video-based enhancement, and hybrid approaches combining classical enhancement with learning-based or generative models. Extending LUMEN to multi-class object detection and real-time video processing also represents a promising direction. In conclusion, LUMEN offers a robust, balanced, and practical image enhancement pipeline for low-light face detection, demonstrating that multi-stage integration of illumination, contrast, and noise control is essential for achieving reliable detection performance without sacrificing visual quality.

**Author Contributions:** Conceptualization: R.R. and D.R.I.M.S.; Methodology: R.R. and D.R.I.M.S.; Software: R.R. and I.H.; Validation: I.H., K.D.W., and C.A.; Formal analysis: K.D.W. and C.A.; Investigation: D.R.I.M.S. and C.A.; Resources: R.R.; Writing—original draft preparation: R.R.; Writing—review and editing: D.R.I.M.S.; Visualization: I.H.; Supervision: D.R.I.M.S.; Project administration: R.R.; Funding acquisition: All. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The dataset used and the source code in this study are publicly available and can be accessed via the following repository: <https://github.com/Rivaldisky73/Image-Enhancement-Low-Light>. This repository contains the Human Faces Object Detection Dataset and related resources utilized for low-light image enhancement and face detection experiments. All data supporting the findings of this research are included within this repository, and no additional proprietary data were used.

**Acknowledgments:** The authors acknowledge the limited use of AI-based language assistance tools to improve grammar and readability during manuscript preparation. All ideas, methodology, experiments, analyses, and conclusions were conceived, conducted, and validated solely by the authors. The use of AI tools did not influence the scientific content or originality of this work.

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

## References

- [1] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, "Face recognition," *ACM Comput. Surv.*, vol. 35, no. 4, pp. 399–458, Dec. 2003, doi: 10.1145/954339.954342.
- [2] W. Liu, P. Zhao, X. Song, and B. Zhang, "A Survey of Low-light Image Enhancement," *Front. Comput. Intell. Syst.*, vol. 1, no. 3, pp. 88–92, Oct. 2022, doi: 10.54097/fcis.v1i3.2242.
- [3] L. R. Zuama, D. R. I. M. Setiadi, A. Susanto, S. Santosa, H.-S. Gan, and A. A. Ojugo, "High-Performance Face Spoofing Detection using Feature Fusion of FaceNet and Tuned DenseNet201," *J. Futur. Artif. Intell. Technol.*, vol. 1, no. 4, pp. 385–400, Feb. 2025, doi: 10.62411/faith.3048-3719-62.

- [4] R. M. Azeez, I. A. Alshabeeb, and W. M. R. Shakir, "Enhanced Face Recognition Using Dolphin Swarm Optimization with Euclidean Classification and PCA," *J. Futur. Artif. Intell. Technol.*, vol. 2, no. 3, pp. 405–416, Sep. 2025, doi: 10.62411/faith.3048-3719-127.
- [5] M. Alansari, O. A. Hay, S. Javed, A. Shoufan, Y. Zweiri, and N. Werghe, "GhostFaceNets: Lightweight Face Recognition Model From Cheap Operations," *IEEE Access*, vol. 11, pp. 35429–35446, 2023, doi: 10.1109/ACCESS.2023.3266068.
- [6] D. J. Jobson, Z. Rahman, and G. A. Woodell, "A multiscale retinex for bridging the gap between color images and the human observation of scenes," *IEEE Trans. Image Process.*, vol. 6, no. 7, pp. 965–976, Jul. 1997, doi: 10.1109/83.597272.
- [7] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017, doi: 10.1109/TIP.2017.2662206.
- [8] W. Yang, S. Wang, Y. Fang, Y. Wang, and J. Liu, "From Fidelity to Perceptual Quality: A Semi-Supervised Approach for Low-Light Image Enhancement," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020, pp. 3060–3069. doi: 10.1109/CVPR42600.2020.00313.
- [9] X. Guo, "LIME: A Method for Low-light Image Enhancement," in *Proceedings of the 24th ACM international conference on Multimedia*, Oct. 2016, pp. 87–91. doi: 10.1145/2964284.2967188.
- [10] D. Tete and C. S. Panda, "A Review Paper on LIME: Low-Light Image Enhancement," *Int. J. Nov. Res. Dev.*, vol. 7, no. 12, 2022.
- [11] S. M. Pizer *et al.*, "Adaptive histogram equalization and its variations," *Comput. Vision, Graph. Image Process.*, vol. 39, no. 3, pp. 355–368, Sep. 1987, doi: 10.1016/S0734-189X(87)80186-X.
- [12] H. Han, S. Shan, L. Qing, X. Chen, and W. Gao, "Lighting Aware Preprocessing for Face Recognition across Varying Illumination," in *Lecture Notes in Computer Science*, 2010, pp. 308–321. doi: 10.1007/978-3-642-15552-9\_23.
- [13] S. Rahman, M. M. Rahman, M. Abdullah-Al-Wadud, G. D. Al-Quaderi, and M. Shoyab, "An adaptive gamma correction for image enhancement," *EURASIP J. Image Video Process.*, vol. 2016, no. 1, p. 35, Dec. 2016, doi: 10.1186/s13640-016-0138-1.
- [14] Z. Zhang, H. Zheng, R. Hong, M. Xu, S. Yan, and M. Wang, "Deep Color Consistent Network for Low-Light Image Enhancement," in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2022, pp. 1889–1898. doi: 10.1109/CVPR52688.2022.00194.
- [15] H. Ackar, A. A. Almisreb, and M. A. Saleh, "A Review on Image Enhancement Techniques," *Southeast Eur. J. Soft Comput.*, vol. 8, no. 1, Apr. 2019, doi: 10.21533/scjournal.v8i1.175.
- [16] J. Wang, H. Wang, Y. Sun, and J. Yang, "Improved Retinex-Theory-Based Low-Light Image Enhancement Algorithm," *Appl. Sci.*, vol. 13, no. 14, p. 8148, Jul. 2023, doi: 10.3390/app13148148.
- [17] J. Deng, J. Guo, Y. Zhou, J. Yu, I. Kotsia, and S. Zafeiriou, "RetinaFace: Single-stage Dense Face Localisation in the Wild," *arXiv*. May 04, 2019. [Online]. Available: <http://arxiv.org/abs/1905.00641>
- [18] A. M. Reza, "Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement," *J. VLSI signal Process. Syst. signal, image video Technol.*, vol. 38, no. 1, pp. 35–44, Aug. 2004, doi: 10.1023/B:VLSI.0000028532.53893.82.
- [19] E. H. Land, "The Retinex Theory of Color Vision," *Sci. Am.*, vol. 237, no. 6, pp. 108–128, Dec. 1977, doi: 10.1038/scientificamerican1277-108.
- [20] X. Guo, Y. Li, and H. Ling, "LIME: Low-Light Image Enhancement via Illumination Map Estimation," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 982–993, Feb. 2017, doi: 10.1109/TIP.2016.2639450.
- [21] X. Pan, C. Li, Z. Pan, J. Yan, S. Tang, and X. Yin, "Low-Light Image Enhancement Method Based on Retinex Theory by Improving Illumination Map," *Appl. Sci.*, vol. 12, no. 10, 2022, doi: 10.3390/app12105257.
- [22] Z. Gu, M. Ju, and D. Zhang, "A novel Retinex image enhancement approach via brightness channel prior and change of detail prior," *Pattern Recognit. Image Anal.*, vol. 27, no. 2, pp. 234–242, Apr. 2017, doi: 10.1134/S1054661817020055.
- [23] C. Guo *et al.*, "Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020, pp. 1777–1786. doi: 10.1109/CVPR42600.2020.00185.
- [24] I. Majid Mohammed and N. Ashidi Mat Isa, "Contrast Limited Adaptive Local Histogram Equalization Method for Poor Contrast Image Enhancement," *IEEE Access*, vol. 13, pp. 62600–62632, 2025, doi: 10.1109/ACCESS.2025.3558506.
- [25] J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou, "RetinaFace: Single-Shot Multi-Level Face Localisation in the Wild," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020, pp. 5202–5211. doi: 10.1109/CVPR42600.2020.00525.
- [26] B. B. Singh and S. Patel, "Efficient Medical Image Enhancement using CLAHE Enhancement and Wavelet Fusion," *Int. J. Comput. Appl.*, vol. 167, no. 5, pp. 1–5, Jun. 2017, doi: 10.5120/ijca2017913277.
- [27] J. Zhong, H. Xie, and F. L. Wang, "The research trends in recommender systems for e-learning," *Asian Assoc. Open Univ. J.*, vol. 14, no. 1, pp. 12–27, Nov. 2019, doi: 10.1108/AAOUJ-03-2019-0015.
- [28] C. Li, J. Zhu, L. Bi, W. Zhang, and Y. Liu, "A low-light image enhancement method with brightness balance and detail preservation," *PLoS One*, vol. 17, no. 5, p. e0262478, May 2022, doi: 10.1371/journal.pone.0262478.
- [29] W. Kim, "Low-Light Image Enhancement: A Comparative Review and Prospects," *IEEE Access*, vol. 10, pp. 84535–84557, 2022, doi: 10.1109/ACCESS.2022.3197629.
- [30] D. Mishra and B. Palkar, "Image Fusion Techniques: A Review," *Int. J. Comput. Appl.*, vol. 130, no. 9, pp. 7–13, Nov. 2015, doi: 10.5120/ijca2015907084.
- [31] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, "GhostNet: More Features From Cheap Operations," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020, pp. 1577–1586. doi: 10.1109/CVPR42600.2020.00165.
- [32] S. Fu, Q. Zhao, H. Liu, Q. Tao, and D. Liu, "Low-light object detection via adaptive enhancement and dynamic feature fusion," *Alexandria Eng. J.*, vol. 126, pp. 60–69, Jul. 2025, doi: 10.1016/j.aej.2025.04.047.
- [33] M. F. Khan, D. Goyal, M. M. Nofal, E. Khan, R. Al-Hmouz, and E. Herrera-Viedma, "Fuzzy-Based Histogram Partitioning for Bi-Histogram Equalisation of Low Contrast Images," *IEEE Access*, vol. 8, pp. 11595–11614, 2020, doi: 10.1109/ACCESS.2020.2965174.
- [34] H. Wang and Y. Song, "Adaptive Low Light Enhancement via Joint Global-Local Illumination Adjustment," *arXiv*. Apr. 01, 2025. [Online]. Available: <http://arxiv.org/abs/2504.00400>

- [35] X. Liang, X. Chen, K. Ren, X. Miao, Z. Chen, and Y. Jin, "Low-light image enhancement via adaptive frequency decomposition network," *Sci. Rep.*, vol. 13, no. 1, p. 14107, Aug. 2023, doi: 10.1038/s41598-023-40899-8.
- [36] S. Fu, J. Zhang, and Y. Luo, "A Retinex-based variational model for low-light image enhancement with noise transformation," *Pattern Recognit.*, vol. 172, p. 112781, Apr. 2026, doi: 10.1016/j.patcog.2025.112781.
- [37] Z. Liu *et al.*, "Illuminating the Shadows: Enhanced Low-Light Image via a Retinex-based Model with Color Equalization," *Expert Syst. Appl.*, vol. 296, p. 129039, Jan. 2026, doi: 10.1016/j.eswa.2025.129039.
- [38] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 4th ed. Pearson, 2018.
- [39] A. Setyono and D. Setiadi, "Robust Video Watermarking using Tchebichef Transform and Singular Value Decomposition on the Selected Frame Based YCbCr Color Space," *Int. J. Intell. Eng. Syst.*, vol. 13, no. 6, pp. 432–441, Dec. 2020, doi: 10.22266/ijies2020.1231.38.
- [40] Z. Hong *et al.*, "Efficient Global Color, Luminance, and Contrast Consistency Optimization for Multiple Remote Sensing Images," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 16, pp. 622–637, 2023, doi: 10.1109/JSTARS.2022.3229392.
- [41] M. S. Ragavendirane and S. Dhanasekar, "Low-Light Image Enhancement via New Intuitionistic Fuzzy Generator-Based Retinex Approach," *IEEE Access*, vol. 13, pp. 38454–38469, 2025, doi: 10.1109/ACCESS.2025.3545258.
- [42] J. J. Jeon, J. Y. Park, and I. K. Eom, "Low-light image enhancement using gamma correction prior in mixed color spaces," *Pattern Recognit.*, vol. 146, p. 110001, Feb. 2024, doi: 10.1016/j.patcog.2023.110001.
- [43] S. A. Mim, J. Cheng, and J. Zhang, "Fusion-based framework for low-light image enhancement using BM3D filter and adaptive gamma correction," *Opt. Laser Technol.*, vol. 192, p. 113662, Dec. 2025, doi: 10.1016/j.optlastec.2025.113662.
- [44] L. Pi, Y. Mu, and L. Pi, "Multiscale processing of image gradient domain-based convolutional neural networks for visual feature enhancement," *J. Comb. Math. Comb. Comput.*, vol. 127a, Apr. 2025, doi: 10.61091/jcmcc127a-269.
- [45] D. Mamieva, A. B. Abdusalomov, M. Mukhiddinov, and T. K. Whangbo, "Improved Face Detection Method via Learning Small Faces on Hard Images Based on a Deep Learning Approach," *Sensors*, vol. 23, no. 1, p. 502, Jan. 2023, doi: 10.3390/s23010502.
- [46] S. Santoso, D. R. I. M. Setiadi, and R. A. Pramunendar, "Comparative No-Reference Evaluation of Classical Image Sharpening Techniques under Varying Degradation Conditions," *J. Appl. Informatics Comput.*, vol. 9, no. 6, pp. 3442–3453, 2025, doi: 10.30871/jaic.v9i6.11430.
- [47] N. Yuan *et al.*, "Low-Light Image Enhancement by Combining Transformer and Convolutional Neural Network," *Mathematics*, vol. 11, no. 7, p. 1657, Mar. 2023, doi: 10.3390/math11071657.
- [48] J. Vazquez-Corral, G. D. Finlayson, and L. Herranz, "Improving the perception of low-light enhanced images," *Opt. Express*, vol. 32, no. 4, p. 5174, Feb. 2024, doi: 10.1364/OE.509713.
- [49] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-Reference Image Quality Assessment in the Spatial Domain," *IEEE Trans. Image Process.*, vol. 21, no. 12, pp. 4695–4708, Dec. 2012, doi: 10.1109/TIP.2012.2214050.
- [50] X. Wu *et al.*, "Low-Light Enhancement Effect on Classification and Detection: An Empirical Study," *arXiv*. Sep. 22, 2024. [Online]. Available: <http://arxiv.org/abs/2409.14461>
- [51] J. Zhang, C. Hou, X. Yang, X. Yang, W. Yang, and H. Cui, "Advancing face detection efficiency: Utilizing classification networks for lowering false positive incidences," *Array*, vol. 22, p. 100347, Jul. 2024, doi: 10.1016/j.array.2024.100347.
- [52] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal Loss for Dense Object Detection," *ArXiv*. Feb. 07, 2018. [Online]. Available: <http://arxiv.org/abs/1708.02002>