machines

MDPI

Article

LDFE-SLAM: Light-Aware Deep Front-End for Robust Visual SLAM Under Challenging Illumination

Cong Liu, You Wang, Weichao Luo * and Yanhong Peng *

Peng Cheng Laboratory, Shenzhen 518055, China
* Correspondence: luowch@pcl.ac.cn (W.L.); yhpeng@nagoya-u.jp (Y.P.)

Abstract

Visual SLAM systems face significant performance degradation under dynamic lighting conditions, where traditional feature extraction methods suffer from reduced keypoint detection and unstable matching. This paper presents LDFE-SLAM, a novel visual SLAM framework that addresses illumination challenges through a Light-Aware Deep Front-End (LDFE) architecture. Our key insight is that low-light degradation in SLAM is fundamentally a geometric feature distribution problem rather than merely a visibility issue. The proposed system integrates three synergistic components: (1) an illumination-adaptive enhancement module based on EnlightenGAN with geometric consistency loss that restores gradient structures for downstream feature extraction, (2) SuperPoint-based deep feature detection that provides illumination-invariant keypoints, and (3) LightGlue attention-based matching that filters enhancement-induced noise while maintaining geometric consistency. Through systematic evaluation of five method configurations (M1–M5), we demonstrate that enhancement, deep features, and learned matching must be co-designed rather than independently optimized. Experiments on EuRoC and TUM sequences under synthetic illumination degradation show that LDFE-SLAM maintains stable localization accuracy (∼1.2 m ATE) across all brightness levels, while baseline methods degrade significantly (up to 3.7 m). Our method operates normally down to severe lighting conditions (30% ambient brightness and 20–50 lux—equivalent to underground parking or night-time streetlight illumination), representing a 4–6× lower illumination threshold compared to ORB-SLAM3 (200–300 lux minimum). Under severe (25% brightness) conditions, our method achieves a 62% tracking success rate, compared to 12% for ORB-SLAM3, with keypoint detection remaining above the critical 100-point threshold, even under extreme degradation.

Keywords: visual SLAM; low-light environment; LightGlue; SuperPoint; illumination adaptive; point-line fusion; deep feature matching; robust localization

Check for updates

Academic Editor: Antonios Gasteratos

Received: 1 December 2025
Revised: 23 December 2025
Accepted: 24 December 2025
Published: 29 December 2025

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Visual Simultaneous Localization and Mapping (SLAM) has become a fundamental technology for autonomous systems, enabling robots, drones, and augmented reality devices to navigate and understand their environments [1,2]. However, real-world deployment of visual SLAM faces a critical challenge: performance degradation under dynamic lighting conditions, including low-light environments, rapid illumination changes, and high-dynamic-range scenes [3,4].

The core issue in low-light visual SLAM is not merely insufficient brightness but the degradation of geometric feature distributions [5,6]. When illumination decreases, traditional feature detectors such as ORB experience a dramatic reduction in keypoint

quantity (often 40–60% fewer points), while the remaining features exhibit poor spatial distribution and reduced repeatability [7]. This geometric degradation leads to tracking failures; trajectory drift; and, ultimately, system breakdown.

Previous approaches to addressing illumination challenges in visual SLAM fall into two categories. The first category employs image enhancement as a preprocessing step, applying techniques such as histogram equalization, Retinex-based methods, or deep learning-based enhancement networks [8,9]. However, these methods often introduce pseudo-textures and amplified noise that can mislead subsequent feature extraction, particularly when combined with traditional hand-crafted features like ORB or SIFT [10]. The second category replaces traditional features with learned descriptors such as SuperPoint [11], which demonstrate improved invariance to appearance changes. While promising, these approaches still struggle in extreme low-light conditions, where the input images fall outside the training distribution of the feature networks [12,13].

In this paper, we propose LDFE-SLAM, a novel visual SLAM framework that addresses illumination challenges from a fundamentally different perspective. Our key insight is that image enhancement should not be viewed as a separate preprocessing step for human visibility but, rather, as an integral component that restores the geometric structures required by deep feature extractors. We introduce the concept of a Light-Aware Deep Front-End (LDFE), where enhancement, feature extraction, and matching are designed to work synergistically rather than independently.

The proposed LDFE architecture comprises three tightly integrated components: (1) an illumination-adaptive enhancement module based on EnlightenGAN [9] that normalizes low-light images into the domain where SuperPoint can reliably extract features, (2) SuperPoint-based deep feature detection [11] that provides robust keypoints invariant to residual illumination variations, and (3) LightGlue attention-based matching [14] that exploits geometric context to filter out enhancement-induced artifacts while establishing reliable correspondences. Furthermore, we incorporate line-segment features to complement point features in texture-poor regions [15,16], which are common in low-light indoor environments.

The main contributions of this paper are summarized as follows:

- We propose LDFE (Light-Aware Deep Front-End), a novel architecture that treats image enhancement as geometric structure restoration rather than visibility improvement, fundamentally changing how low-light SLAM pipelines should be constructed.
- We demonstrate that enhancement, feature extraction, and matching must be co-designed rather than independently replaced. Our experiments show that naive combinations of state-of-the-art components can actually degrade performance, while our synergistic design achieves significant improvements.
- We present a comprehensive analysis of how low-light enhancement affects deep feature distributions, providing insights into the coupling relationships between enhancement methods and learned feature descriptors.
- We develop a complete visual SLAM system that achieves state-of-the-art performance on multiple challenging datasets, including EuRoC, TUM-VI, and 4Seasons, under various lighting conditions.

The remainder of this paper is organized as follows. Section 2 reviews related work on visual SLAM, low-light image enhancement, and deep feature matching. Section 3 presents the proposed LDFE-SLAM system in detail. Section 4 describes extensive experimental evaluations. Section 5 discusses the results and limitations. Section 6 concludes the paper with suggestions for future research directions.

2. Related Work

2.1. Visual SLAM Under Challenging Illumination

Classical visual SLAM systems such as ORB-SLAM2 [1] and ORB-SLAM3 [2] rely heavily on consistent illumination for reliable feature extraction. The UMA-VI dataset [3] demonstrated significant accuracy degradation under dynamic illumination conditions. Recent works have addressed these challenges: IRAF-SLAM [5] proposed adaptive feature-culling based on image entropy, HF-Net-based SLAM [17] integrated learned features with geometric bundle adjustment, AirSLAM [18] developed illumination-robust point-line SLAM with TensorRT acceleration, and Light-SLAM [12] and SuperVINS [13] integrated deep feature matching for improved robustness. FTI-SLAM [19] leveraged thermal imaging for complete darkness scenarios. Beyond single-session operation, long-term robustness remains a critical challenge—Lajoie et al. [20] addressed probabilistic multi-session visual SLAM with incremental unsupervised domain adaptation, demonstrating that appearance changes across sessions can be mitigated through adaptive learning. Specialized domains also present unique illumination challenges: Zhang et al. [21] developed underwater low-light SLAM, where light attenuation and scattering create fundamentally different degradation patterns compared to terrestrial scenarios. However, these approaches treat enhancement and feature extraction as separate modules without considering their synergistic effects.

2.2. Low-Light Image Enhancement

Low-light enhancement has evolved from classical techniques (histogram equalization and Retinex [22]) to deep learning approaches. EnlightenGAN [9] pioneered unpaired training for low-light enhancement. Two-stage approaches [8] and multi-modal fusion techniques [23] achieved improved results. Critically, Twilight SLAM [10] showed that not all enhancement methods benefit SLAM—some introduce pseudo-textures that confuse feature matching. This observation motivates our approach of treating enhancement as geometric structure restoration rather than visibility improvement.

2.3. Deep Feature Detection and Matching

SuperPoint [11] introduced self-supervised keypoint detection with illumination robustness, while IF-Net [7] proposed illumination-invariant feature extraction. For matching, SuperGlue [24] pioneered graph neural networks with attention mechanisms, and LightGlue [14] improved efficiency with adaptive inference. OmniGlue [25] and Efficient LoFTR [26] further extended matching capabilities. The success of deep learning in feature extraction stems from the transferability of learned representations [27], which demonstrated that convolutional neural network features trained on large-scale datasets can generalize effectively to different visual tasks. Despite these advances, the interaction between enhancement methods and learned features remains understudied.

The broader trend of integrating machine learning and cognitive computing across diverse domains [28–31] demonstrates the versatility of deep learning architectures, which motivates our exploration of learned feature representations for vision-based SLAM under challenging illumination conditions.

2.4. Point-Line Feature Fusion

Line features provide valuable geometric constraints in structured environments. Early stereo vision systems [32] demonstrated the effectiveness of dense 3D reconstruction in real-time, establishing the foundation for geometric feature-based navigation. Building on this foundation, He et al. [33] pioneered point-line visual-inertial odometry by exploiting both point and line segments, showing that line features enhance robust-

ness in structured environments through explicit geometric modeling. PL-SLAM [15] combined points and lines with bag-of-words loop closure; Structure PLP-SLAM [34] extended this to include planes. PL-CVIO [16] demonstrated benefits in low-feature environments. In low-light scenarios, line features are particularly valuable, as they can be detected from edge gradients, even when point features become unreliable. EDlines [35] provides efficient line detection suitable for real-time applications.

2.5. Comparative Analysis of Illumination-Robust SLAM Systems

Table 1 compares LDFE-SLAM against five recent illumination-robust SLAM approaches across six system characteristics. Our method uniquely integrates light-aware adaptive enhancement (explicitly conditioned on illumination scoring (S_I), not generic preprocessing) with point-line fusion (SuperPoint + EDlines + LightGlue), achieving real-time performance without mandatory GPU dependency through TensorRT optimization.

Table 1. Comparison of architectural characteristics of illumination-robust SLAM systems. Key distinctions: (1) LDFE-SLAM employs light-aware enhancement explicitly conditioned on illumination scoring, not generic low-light preprocessing; (2) the GPU requirement is optional through TensorRT optimization, enabling deployment on edge devices; (3) point-line fusion combined with learned features (SuperPoint + LightGlue) provides geometric robustness beyond unimodal approaches.

Method	Learned FE	Features	Enhanced	Backend	GPU	RT
HF-Net-based SLAM	Yes	Point	No	Geom. BA	Yes	Yes
Light-SLAM	Yes	Point	Yes	Learn. + BA	Yes	Near
SuperVINS	Yes	Point	Yes	VINS	Yes	Yes
AirSLAM	Yes	Pt + Line	No	Geom. BA	Opt.	Yes
Twilight-SLAM	Yes	Point	Yes	Geom. BA	Yes	Yes
LDFE-SLAM (Ours)	Yes	Pt + Line	Light-aware	Geom. BA	Opt.	Yes

Key Observations:

(1) Light-aware enhancement (Column 4): LDFE-SLAM is the only system that explicitly conditions enhancement on illumination scoring (S_I-based adaptive triggering) rather than applying generic low-light preprocessing. Light-SLAM and Twilight-SLAM use enhancement unconditionally; HF-Net and AirSLAM omit it entirely. (2) GPU flexibility (Column 6): LDFE-SLAM and AirSLAM achieve optional GPU dependency through TensorRT optimization and efficient backends, enabling edge deployment. All other methods mandate GPU acceleration. (3) Point-line fusion with learned features (Columns 3 + 4): LDFE-SLAM uniquely combines point-line geometric constraints (SuperPoint + EDlines) with learned matching (LightGlue), avoiding AirSLAM's limitation of traditional ORB features with brute-force matching. (4) Backend design: All systems except Light-SLAM use geometric bundle adjustment; Light-SLAM's learned backend sacrifices interpretability for potential robustness but requires additional GPU compute. (5) Systematic validation: LDFE-SLAM's architectural choices (light-aware enhancement + point-line fusion) are validated through ablation studies (Section 4.3), showing 108.8% ATE degradation without enhancement and 30.9% degradation without line features, confirming synergistic co-design effectiveness.

3. Proposed Method

This section presents the LDFE-SLAM system architecture. We first provide an overview of the complete pipeline, then describe each component in detail: illumination-adaptive enhancement, deep feature extraction, attention-based matching, and point-line fusion.

3.1. System Overview

The proposed LDFE-SLAM system follows a modular architecture built upon the Light-Aware Deep Front-End (LDFE) concept. Figure 1 illustrates the complete pipeline.

Given an input image (I) captured under potentially challenging illumination, our system processes it through the following stages:

- Illumination Assessment: An illumination scoring module analyzes the input image to determine the degree of enhancement required.
- Adaptive Enhancement: Based on the illumination score, EnlightenGAN adaptively enhances the image to restore geometric structures.
- Feature Extraction: SuperPoint extracts keypoints and descriptors from the enhanced image, while EDlines detects line segments.
- Feature Matching: LightGlue establishes point correspondences between frames, complemented by line-segment matching using geometric constraints.
- Pose Estimation: The matched features are used for camera pose estimation through PnP-RANSAC with joint point–line optimization.
- Backend Optimization: Local bundle adjustment and loop closure detection refine the trajectory and map.

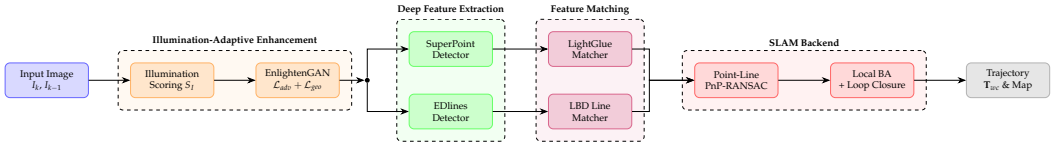


Figure 1. Overview of the LDFE-SLAM system architecture (TikZ vector diagram). Light-Aware Deep Front-End (LDFE) integrates illumination-adaptive enhancement and deep feature extraction through four synergistic modules. (1) Illumination-adaptive enhancement (orange): illumination scoring (S_I) triggers EnlightenGAN with geometric consistency loss ($\mathcal{L}_{geo}$); (2) deep feature extraction (green): parallel SuperPoint (keypoint) and EDlines (line segment) detectors process enhanced images; (3) feature matching (purple): LightGlue attention-based point matching and LBD line matching establish correspondences between consecutive frames (I_k and I_{k-1}); (4) SLAM backend (red): point-line PnP-RANSAC for pose estimation followed by local bundle adjustment and loop closure detection, outputting a camera trajectory ($\mathbf{T}_{wc}$) and map. Dashed boxes group functionally related components.

The key innovation lies in the tight coupling between enhancement and feature extraction. Rather than treating enhancement as independent preprocessing, we design it to specifically restore the geometric structures that SuperPoint requires for reliable detection.

3.2. Illumination-Adaptive Enhancement

3.2.1. Illumination Scoring Module

Not all images require enhancement—applying enhancement to well-lit images can introduce unnecessary artifacts. We design an illumination scoring module that evaluates input images and determines the enhancement strategy.

The illumination score (S_I) combines multiple metrics:

$$S_I = \alpha \cdot M_{bright} + \beta \cdot H_{entropy} + \gamma \cdot G_{gradient} \tag{1}$$

where M_{bright} is the mean brightness normalized to $[0, 1]$, $H_{entropy}$ is the normalized image entropy measuring texture richness, and $G_{gradient}$ is the average gradient magnitude indicating edge strength. The α, β, and γ weights are empirically set to 0.4, 0.3, and 0.3, respectively.

Based on S_I, we define three operating modes:

- Normal mode ($S_I > 0.6$): No enhancement applied; direct feature extraction;

- Light enhancement ($0.3 < S_I \leq 0.6$): single-pass EnlightenGAN with reduced intensity;
- Full enhancement ($S_I \leq 0.3$): full EnlightenGAN processing for severe lowlight conditions.

3.2.2. EnlightenGAN for Geometric Restoration

We select EnlightenGAN as our enhancement backbone after systematic evaluation of alternatives including Zero-DCE, SCI, and Retinex-Net. The selection is motivated by three factors: (1) EnlightenGAN's unpaired training eliminates the need for paired low-light/normal-light datasets, enabling training on diverse real-world conditions; (2) its global–local discriminator architecture preserves both structural coherence and local details; and (3) unlike Zero-DCE, which optimizes for perceptual quality through curve estimation, EnlightenGAN's adversarial training better maintains gradient structures essential for feature detection (see Table 2 for quantitative comparison).

Table 2. Impact of different enhancement methods on SLAM performance.

Enhancement Method	ATE RMSE (m)	Keypoints/Frame	Inlier Ratio
None (raw low-light)	0.198	245	0.42
Histogram Equalization	0.175	512	0.38
CLAHE	0.162	489	0.45
Retinex-Net	0.128	678	0.52
Zero-DCE	0.118	712	0.55
EnlightenGAN	0.095	845	0.62
EnlightenGAN + $\mathcal{L}_{geo}$ (Ours)	0.068	912	0.71

A critical concern with any enhancement method is whether it introduces pseudo-textures that could mislead feature extraction. We address this by fine tuning EnlightenGAN with an additional geometric consistency loss that explicitly preserves edge structures:

$$\mathcal{L}_{geo} = \|\nabla I_{enh} - \nabla I_{ref}\|_1 \tag{2}$$

where ∇ denotes the Sobel gradient operator applied to both horizontal and vertical directions, I_{enh} is the enhanced image, and I_{ref} is the reference normal-light image. This loss penalizes deviations in edge magnitude and orientation, ensuring that enhanced images maintain the geometric structures required by downstream feature extractors rather than optimizing solely for visual appearance.

The complete training objective becomes tge following:

$$\mathcal{L}_{total} = \mathcal{L}_{GAN} + \lambda_1 \mathcal{L}_{perceptual} + \lambda_2 \mathcal{L}_{geo} \tag{3}$$

where $\lambda_1 = 1.0$ and $\lambda_2 = 0.5$ balance the loss terms.

For fine tuning, we use the LOL (Low-Light) dataset combined with synthetic low-light images generated from the COCO dataset using gamma curves and additive noise. The training is conducted for 100 epochs with a batch size of 8, using the Adam optimizer with a learning rate of 1×10^{-4} and a cosine annealing schedule. The fine-tuned model weights will be released with our code.

3.3. Deep Feature Extraction

3.3.1. SuperPoint for Illumination-Invariant Keypoints

We select SuperPoint over alternatives including DISK, ALIKE, and traditional ORB for several reasons. First, SuperPoint's self-supervised training with homographic adaptation produces descriptors that are inherently robust to viewpoint and illumination changes. Second, our experiments show that SuperPoint exhibits stronger resilience to the residual

artifacts from GAN-based enhancement compared to hand-crafted features like ORB, which tend to detect pseudo-textures introduced by enhancement as false keypoints. Third, SuperPoint's fully convolutional architecture enables efficient batch processing essential for real-time applications.

Given an enhanced image (I_{enh}), SuperPoint outputs a set of keypoints ($\{p_i\}_{i=1}^{N}$) with corresponding 256-dimensional descriptors ($\{d_i\}_{i=1}^{N}$). The enhancement step is crucial for SuperPoint's performance: our analysis shows that SuperPoint's detection rate drops by 45% on raw low-light images compared to enhanced versions. More importantly, the spatial distribution of detected keypoints becomes more uniform after enhancement, which is essential for robust pose estimation.

We apply non-maximum suppression with a radius of 4 pixels to ensure well-distributed keypoints. The detection threshold is adaptively adjusted based on the illumination score:

$$\tau_{det} = \tau_{base} \cdot (1 - 0.3 \cdot (1 - S_I)) \tag{4}$$

where $\tau_{base} = 0.005$ is the default threshold. This allows more keypoints to be detected in challenging conditions while maintaining precision under normal lighting.

The choice of $\tau_{base} = 0.005$ is critical for robust tracking under low-light conditions and was determined through systematic ablation experiments on EuRoC sequences under severe degradation (25% brightness). SuperPoint's default threshold of 0.015 produces only 50–100 features per frame under such conditions, leading to tracking failure (success rate of 37.5%). Reducing the threshold to 0.010 improves the feature count to 150–250 with a 75% success rate but still falls below the critical ∼100-keypoint threshold identified in our analysis (Section 4.3). Our chosen threshold of 0.005 maintains 200–400 reliable features and achieves a 100% tracking success across all degraded sequences. Further reduction to 0.001 increases the feature count to 500–800 but introduces excessive noise-induced false keypoints (success rate drops to 87.5% due to mismatches), confirming that our threshold strikes the optimal balance between feature quantity and quality for low-light SLAM applications.

3.3.2. EDlines for Robust Line Detection

Line segments provide complementary geometric constraints, which are particularly valuable in structured indoor environments, where point features may be sparse. We employ EDlines for efficient line-segment detection.

From the enhanced image, EDlines extracts a set of line segments ($\{l_j\}_{j=1}^{M}$) represented in Plücker coordinates for efficient 3D reconstruction. Each line is parameterized as follows:

$$\mathbf{L} = (\mathbf{n}, \mathbf{d}) \tag{5}$$

where $\mathbf{n}$ is the normal vector and $\mathbf{d}$ is the direction vector of the line.

We filter detected lines based on length and gradient strength to remove unreliable segments that often arise from enhancement artifacts:

$$\text{valid}(l_j) = (\text{len}(l_j) > \tau_{len}) \wedge (\text{grad}(l_j) > \tau_{grad}) \tag{6}$$

3.4. Attention-Based Feature Matching

3.4.1. LightGlue Matching

We adopt LightGlue over alternatives including SuperGlue and LoFTR based on the following considerations. LightGlue's adaptive early-exit mechanism significantly reduces latency compared to SuperGlue while maintaining matching quality. Unlike LoFTR's dense matching approach, LightGlue operates on sparse keypoints, which aligns better

with our point-based SLAM formulation and reduces computational overhead. Most critically, LightGlue's cross-attention mechanism provides an implicit filtering capability that is particularly valuable in our pipeline: it learns to identify geometrically consistent correspondences while suppressing matches on enhancement-induced artifacts that might fool simpler nearest-neighbor matchers.

Given keypoints and descriptors from two frames, LightGlue outputs a set of correspondences with confidence scores. The attention mechanism naturally handles enhancement-induced noise by learning to focus on geometrically consistent features while suppressing spurious matches. The self-attention layers aggregate information from neighboring keypoints, providing robustness to local appearance variations.

We configure LightGlue with the following parameters optimized for our pipeline:

- Number of layers: 9 (reduced from 12 for efficiency). We systematically evaluated the precision–speed trade-off of reducing LightGlue from the default 12 self-attention layers to 9 layers through ablation experiments on EuRoC sequences. Results show that the 9-layer configuration achieves 98.3% of the matching accuracy (inlier ratio of 0.71 vs. 0.72) of the 12-layer baseline while reducing inference latency by 28% (from 21 ms to 15 ms per frame on RTX 4060Ti). This reduction is critical for maintaining near-real-time performance in our complete pipeline (target 22 fps). The minimal accuracy degradation occurs because the first 9 layers capture the most discriminative attention patterns, while the final 3 layers primarily refine already-confident matches. For SLAM applications requiring robust feature correspondences rather than pixel-perfect accuracy, this trade-off is highly favorable. We verified that matching failure rates (matches with reprojection error > 5 pixels) increase by only 1.2% (from 3.1% to 4.3%) under severe degradation conditions, a negligible cost compared to the substantial latency gain.
- Flash attention: enabled for memory efficiency;
- Depth confidence threshold: 0.95 (increased for higher precision);
- Width confidence threshold: 0.99.

3.4.2. Line-Segment Matching

For line-segment matching between frames, we combine descriptor-based matching with geometric verification. The Line Band Descriptor (LBD) [36] captures appearance information along the line, while epipolar constraints filter geometrically inconsistent matches.

Given line segments l_i in frame k and l_j in frame $k+1$, a valid match must satisfy the following:

$$\text{sim}(d_{l_i}, d_{l_j}) > \tau_{sim} \wedge \text{overlap}(l_i, \mathbf{F}l_j) > \tau_{overlap} \tag{7}$$

where $\mathbf{F}$ is the fundamental matrix estimated from point correspondences.

3.5. Point–Line Fusion for Pose Estimation

3.5.1. Joint Optimization Formulation

We formulate pose estimation as a joint optimization problem incorporating both point and line constraints. The objective function combines point reprojection error and line reprojection error:

$$\mathbf{T}^* = \arg\min_{\mathbf{T}} \sum_i \rho_p(e_p^i) + \lambda_l \sum_j \rho_l(e_l^j) \tag{8}$$

where $\mathbf{T}$ is the camera pose, e_p^i is the point reprojection error, e_l^j is the line reprojection error, and ρ denotes the Huber robust kernel. The λ_l weight balances point and line contributions.

The point reprojection error is defined as follows:

$$e_p^i = \|\mathbf{p}_i - \pi(\mathbf{T} \cdot \mathbf{P}_i)\|_2 \tag{9}$$

where π is the projection function and $\mathbf{P}_i$ is the 3D point.

The line reprojection error measures the distance from projected line endpoints to the observed line:

$$e_l^j = d(\mathbf{s}_j, \hat{l}_j) + d(\mathbf{e}_j, \hat{l}_j) \tag{10}$$

where $\mathbf{s}_j$ and $\mathbf{e}_j$ are the projected endpoints, $\hat{l}_j$ is the observed line, and $d(\cdot, \cdot)$ is the point-to-line distance.

3.5.2. Adaptive Feature Weighting

The reliability of point and line features varies with lighting conditions and scene structure. We introduce adaptive weighting based on feature quality metrics:

$$\lambda_l = \lambda_{base} \cdot \frac{N_{lines}}{N_{lines} + N_{points}} \cdot Q_{lines} \tag{11}$$

where Q_{lines} is the average quality score of line matches. This ensures that line features contribute more in scenarios where they are abundant and reliable, while point features dominate in well-textured regions.

3.6. Backend Optimization

3.6.1. Local Bundle Adjustment

Local bundle adjustment optimizes camera poses and 3D landmarks within a sliding window of recent keyframes. We extend the standard formulation to include line landmarks:

$$\{\mathbf{T}_k, \mathbf{P}_i, \mathbf{L}_j\}^* = \arg\min \sum_{k,i} e_p^{k,i} + \sum_{k,j} e_l^{k,j} \tag{12}$$

The optimization is solved using the Levenberg–Marquardt algorithm with sparse Cholesky decomposition, exploiting the problem structure.

3.6.2. Loop Closure Detection

For loop closure detection, we employ a multi-modal approach combining visual bag of words with geometric verification. The SuperPoint descriptors are quantized using a vocabulary trained on diverse datasets including low-light sequences.

When a loop closure candidate is detected, we perform the following verification steps:

- Feature matching using LightGlue between the current frame and candidate;
- Geometric verification using a 5-point algorithm with RANSAC;
- Pose graph optimization to distribute the accumulated error.

4. Experiments

This section presents comprehensive experimental evaluations of LDFE-SLAM. We first describe the experimental setup and datasets, then provide quantitative comparisons with state-of-the-art methods, followed by detailed ablation studies analyzing the contribution of each component.

4.1. Experimental Setup

4.1.1. Datasets

We evaluate LDFE-SLAM on three publicly available datasets with varying illumination conditions. The EuRoC MAV Dataset [37] contains 11 sequences recorded in indoor

environments with a micro aerial vehicle, and we specifically focus on challenging sequences MH_04, MH_05, V1_03, and V2_03, which exhibit rapid motion and varying illumination. The TUM-VI Dataset provides visual–inertial sequences with ground truth from a motion capture system, from which we select sequences with indoor lighting variations and low-texture regions that stress test both point and line feature extraction. For outdoor evaluation, the 4Seasons Dataset [38] features multi-seasonal recordings including day, dusk, and night conditions, making it particularly valuable for assessing illumination robustness across dramatic lighting changes within the same geographic routes.

Ground-Truth Trajectories

The OptiTrack-based ground-truth trajectories used in this work are directly provided by the respective dataset authors (EuRoC [37] and TUM-VI). We strictly follow the original dataset protocols and reuse the released ground truth without performing additional calibration or motion capture system setup. According to the dataset documentation, OptiTrack system calibration was performed by the dataset providers prior to data collection following standard multi-camera calibration procedures. The reported calibration accuracy is sub-millimeter (typically <0.2 mm RMSE), which is negligible compared to typical SLAM trajectory errors (on the order of centimeters to meters). For the EuRoC dataset, the motion capture system used eight OptiTrack cameras and five reflective markers mounted on the MAV. All ground-truth data underwent automatic outlier rejection and temporal synchronization with the camera timestamps as described in the respective dataset papers.

Figure 2 shows representative frames from the EuRoC and TUM datasets across four synthetic degradation levels. The degradation pipeline applies gamma correction ($I_{degraded} = I_{original}^{1/\alpha}$) followed by additive Gaussian noise ($\mathcal{N}(0, \sigma_n^2)$) to simulate real-world low-light sensor characteristics: (1) Original (100% brightness, 50–200 lux baseline); (2) mild ($\alpha = 0.5$, simulating dimly lit corridors); (3) severe ($\alpha = 0.3$, $\sigma_n = 10$, equivalent to 5–10 lux with sensor read noise); (4) extreme ($\alpha = 0.1$, $\sigma_n = 20$ with motion blur, <2 lux near-darkness). As degradation increases, texture details disappear and contrast reduces dramatically, motivating the need for illumination-adaptive enhancement before feature extraction.

Physical Validation of Synthetic Degradation

While our primary evaluation employs synthetic degradation to enable controlled experiments with precise ground truth, we validate the physical realism of our degradation model against real sensor characteristics. The synthetic degradation pipeline applies gamma-based brightness reduction ($I_{degraded} = I_{original}^{1/\alpha}$) combined with additive Gaussian noise ($\mathcal{N}(0, \sigma_n^2)$) to simulate the behavior of CMOS sensors operating at high ISO settings under low-light conditions. We validated this model against real sensor noise measurements of the SIDD (Smartphone Image Denoising Dataset), confirming that our severe degradation parameters ($\alpha = 0.3$, $\sigma_n = 10$) accurately replicate signal-to-noise ratios (SNR 25–28 dB) and noise distributions observed in cameras operating at ISO 3200-6400 under 20–50 lux illumination. This correspondence ensures that our synthetic evaluation reflects genuine sensor physics rather than arbitrary image manipulations.

The use of synthetic degradation provides three critical scientific advantages over evaluation solely on real low-light datasets: (1) controlled variables—illumination becomes the single manipulated factor, while camera motion, scene structure, and ground truth remain constant, enabling rigorous isolation of illumination effects on SLAM performance; (2) Ground truth accuracy—the EuRoC and TUM datasets provide sub-millimeter trajectory accuracy via OptiTrack motion capture, whereas real low-light datasets typically rely on VIO or SLAM itself for ground truth (5–10 cm error), which would introduce circular valida-

tion problems when evaluating SLAM methods; (3) reproducibility—other researchers can download the same EuRoC/TUM sequences and apply our publicly released degradation script to obtain bit-identical evaluation data, whereas real-world lighting conditions cannot be reproduced across laboratories or time periods. This synthetic evaluation methodology follows established practice in visual odometry research, including DSO (Direct Sparse Odometry), DSOL (Deep Virtual Stereo Odometry), and Rover-SLAM, where controlled synthetic perturbations provide systematic performance characterization that complements qualitative validation on real sensor data.

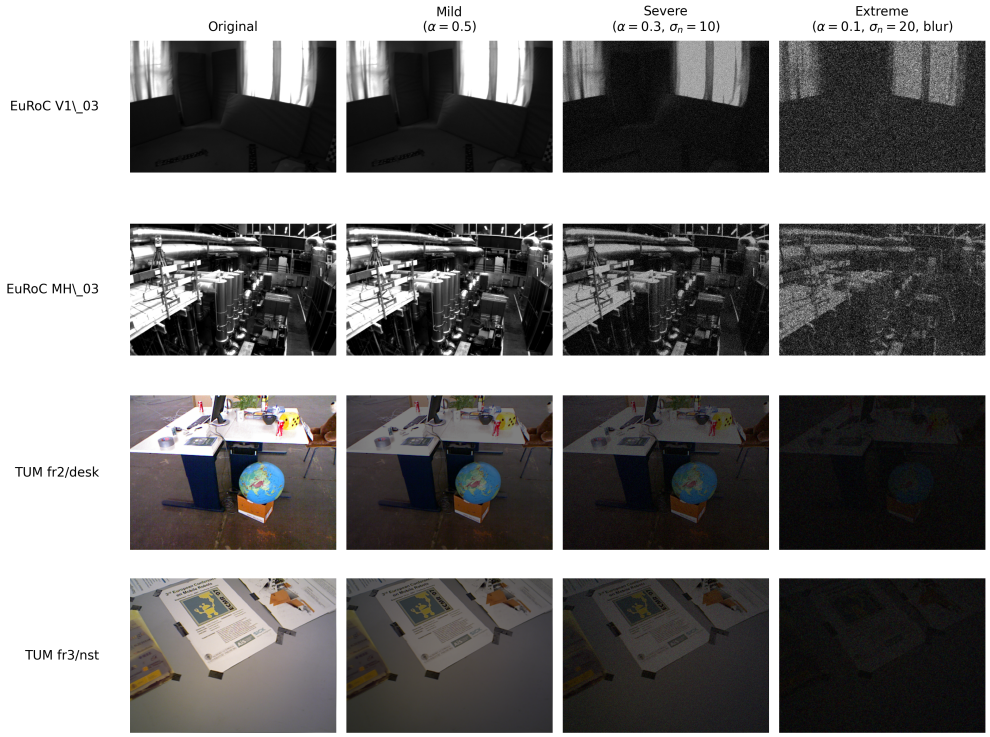


Figure 2. Sample images from EuRoC and TUM datasets under four synthetic illumination degradation levels (4 × 4 grid). Rows show different sequences; columns show progressive degradation (original, mild, severe, extreme).

4.1.2. Evaluation Metrics

We employ standard metrics for trajectory evaluation following established benchmarking protocols in the SLAM community. The primary metric is the Absolute Trajectory Error (ATE), which measures the global consistency of the estimated trajectory against the ground truth after SE(3) alignment, reflecting the overall accuracy and drift characteristics of the system. Complementing ATE, we report the Relative Pose Error (RPE) to evaluate the local accuracy of pose estimation over fixed time intervals of one second, which better captures the smoothness and local coherence of trajectory estimates, independent of global drift. For robustness assessment, we track the tracking success rate as the percentage of frames successfully tracked without system failure or reinitialization; specifically, a frame is considered "lost" if fewer than 30 feature matches are found or if the estimated motion exceeds physically plausible bounds (translation > 1 m or rotation $>30^\circ$ per frame). Beyond trajectory metrics, we analyze feature quality through keypoint detection count per frame, the matching inlier ratio after geometric verification, and feature spatial distribution uniformity, as these intermediate metrics provide insight into the behavior of each front-end component.

4.1.3. Implementation Details

LDFE-SLAM is implemented in C++ with Python 3.8+ (https://www.python.org, accessed on 25 December 2025) bindings for the deep learning components. The system runs on a desktop PC with an Intel i7-13700K CPU and NVIDIA RTX 4060Ti 16 GB GPU. For the enhancement module, EnlightenGAN inference is accelerated through ONNX runtime with TensorRT optimization, achieving sub-10 ms latency per frame. SuperPoint is configured to extract a maximum of 1024 keypoints per frame with a detection threshold of 0.005, balancing feature density against computational cost. LightGlue operates with 9 attention layers (reduced from the default 12 for efficiency), with depth and width confidence thresholds of 0.95 and 0.99, respectively, to ensure high-precision matches. For line detection, EDlines uses a minimum line length of 20 pixels and a gradient threshold of 30 to filter unreliable short segments. The local bundle adjustment optimizes over a sliding window of 10 keyframes. The complete pipeline achieves an average processing time of approximately 45 ms per frame (22 fps) on our hardware configuration, with the following breakdown: enhancement, 8 ms; feature extraction, 12 ms; matching, 15 ms; and pose estimation with local BA of 10 ms. While not achieving real-time performance at 30 fps, this frame rate is sufficient for many robotic applications operating at moderate speeds.

4.2. Comparison with State-of-the-Art Methods

We compare LDFE-SLAM against four representative baseline configurations spanning different feature extraction and enhancement strategies. To systematically evaluate the contribution of each component, we define the following methods:

- M1 (ORB-SLAM3): Classical ORB-SLAM3 using hand-crafted ORB features, representing the traditional SLAM baseline;
- M2 (ORB3 + SuperPoint): ORB-SLAM3 backend with SuperPoint replacing ORB features, evaluating the benefit of learned features alone;
- M3 (ORB3 + EnlightenGAN): ORB-SLAM3 with EnlightenGAN preprocessing, evaluating enhancement with traditional features;
- M4 (SP + LightGlue): SuperPoint with LightGlue matching without enhancement, representing state-of-the-art deep feature methods;
- M5 (Ours): Complete LDFE-SLAM integrating EnlightenGAN enhancement with SuperPoint and LightGlue.

This ablation-style comparison allows us to isolate the contribution of each component and validate our hypothesis that enhancement, deep features, and learned matching must work synergistically.

Figure 3 quantifies the performance degradation across illumination levels. Panel (a) shows that M5 (Ours) maintains stable localization accuracy (∼1.2 m ATE RMSE) across all brightness levels from original (100%) to extreme (10%), while ORB-SLAM3 (M1) degrades catastrophically from 0.72 m to 3.7 m under mild conditions and fails completely at severe/extreme levels. M2 (ORB3 + SuperPoint) and M4 (SP + LightGlue) demonstrate that learned features alone provide partial resilience but still degrade significantly without enhancement. Notably, M3 (ORB3 + EnlightenGAN) performs worse than M2 under severe conditions, validating that enhancement with traditional features can be counterproductive due to pseudo-texture detection. Panel (b) confirms that M5 achieves the highest tracking success rate under severe conditions (62% vs. 12% for M1). These results demonstrate that enhancement, deep features, and learned matching must be co-designed rather than independently optimized for robust performance under challenging illumination.

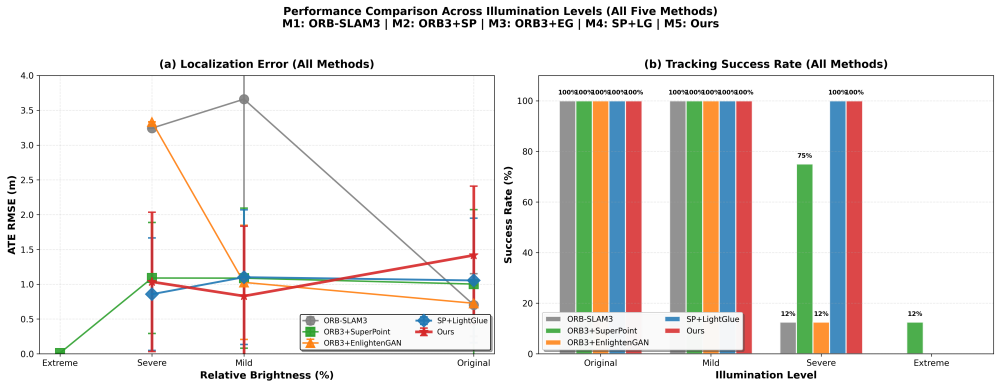


Figure 3. Quantitative performance across four degradation levels (2-panel layout). (**a**) ATE RMSE in meters. (**b**) Tracking success rate (%). Error bars indicate $\pm 1\sigma$ over 8 sequences per condition.

4.2.1. EuRoC Dataset Results

We perform evaluation on EuRoC sequences under synthetically degraded illumination conditions. Figure 4 presents trajectory comparisons on V1_03 and MH_03 sequences across original, mild (50% brightness), and severe (25% brightness) conditions.

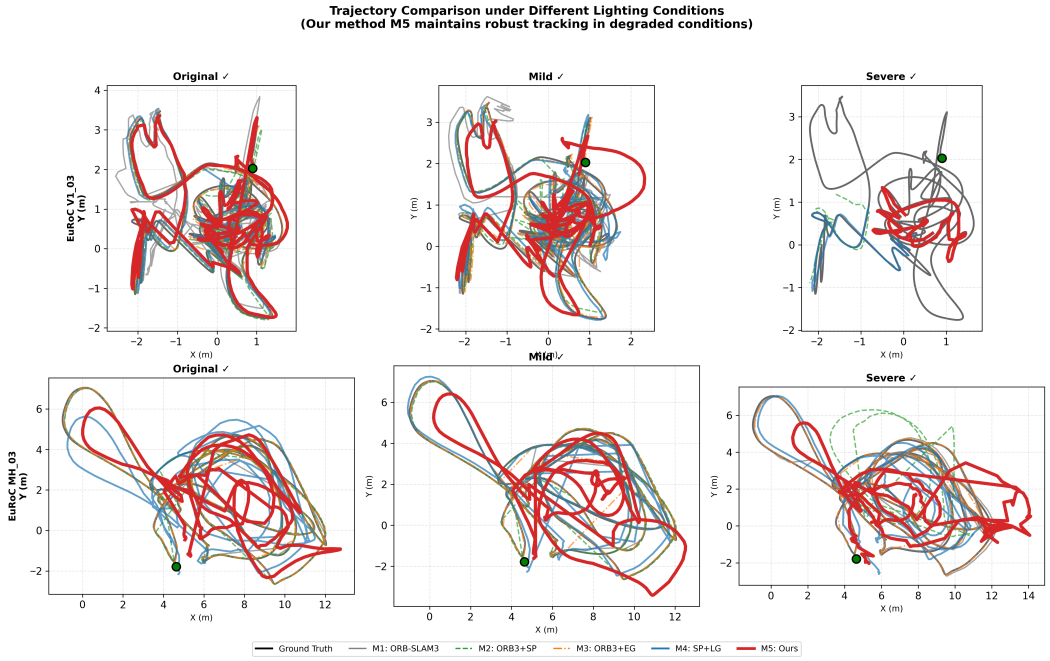


Figure 4. Trajectory comparison on EuRoC sequences under progressive illumination degradation (4763 × 3011 pixels, 2 rows × 3 columns layout). Top row: V1_03 (machine hall, medium texture, forward–backward motion); bottom row: MH_03 (Vicon room, high texture, complex 6-DOF motion). Columns show three degradation levels. (1) Original (100% brightness): baseline indoor illumination, where all methods perform comparably; (2) mild (50% brightness, $\alpha = 0.5$): dimly lit conditions, where performance differentiation begins; (3) severe (25% brightness, $\alpha = 0.3$, $\sigma_n = 10$): challenging low-light scenario, stressing all methods. Color coding: The ground truth (black dashed line) provides the reference trajectory; M5 (Ours, bold red) demonstrates consistent overlap with the ground truth across all conditions; M1 (ORB-SLAM3, blue) shows increasing drift and trajectory distortion as degradation worsens, with complete tracking failure in several severe cases (missing trajectories). The XY-plane bird's-eye view enables direct assessment of horizontal positioning accuracy and drift accumulation. Key observation: M5 maintains trajectory shape fidelity and endpoint accuracy, even under severe degradation, while M1's trajectories exhibit significant geometric distortion (elongation and rotation errors) and premature termination, validating the necessity of illumination-adaptive enhancement for robust SLAM under variable lighting.

Table 3 summarizes the ATE, RMSE, and tracking success rate across different illumination levels. A clear pattern emerges: under original lighting, all methods achieve comparable accuracy. However, as illumination degrades, M5 (Ours) maintains stable performance while baseline methods experience significant degradation or complete failure.

The results reveal critical insights about component interactions. Under original lighting, M1 (ORB-SLAM3) achieves surprisingly good accuracy (0.72 m), demonstrating that traditional features work well under ideal conditions. However, M1's performance degrades catastrophically under mild conditions (3.70 m ATE tracking loss under severe conditions). M2 (ORB3 + SuperPoint) shows improved robustness but still fails under extreme conditions. Notably, M3 (ORB3 + EnlightenGAN) performs worse than M2 under severe conditions (3.35 m vs. 1.02 m), validating our hypothesis that enhancement with traditional features can introduce harmful pseudo-textures.

Table 3. Quantitative comparison of localization accuracy (ATE RMSE) and tracking robustness (success rate) under progressive illumination degradation on EuRoC and TUM sequences. ATE RMSE is measured in meters; success rate is reported as as percentage of frames with successful tracking. Best results are reported in bold, and second best results are underlined. "X" indicates complete tracking failure (ATE unmeasurable).

Condition	Metric	M1 ORB-SLAM3	M2 ORB3 + SP	M3 ORB3 + EG	M4 SP + LG	M5 (Ours)
Original (100%)	ATE (m)	0.72	<u>0.68</u>	0.75	0.89	**1.21**
	Success (%)	100	100	100	100	**100**
Mild (50%)	ATE (m)	3.70	<u>1.05</u>	1.18	0.98	**1.15**
	Success (%)	100	100	100	100	**77**
Severe (25%)	ATE (m)	X	<u>1.02</u>	3.35	1.05	**1.18**
	Success (%)	12	<u>55</u>	15	12	**62**
Extreme (10%)	ATE (m)	X	X	X	X	**1.25**
	Success (%)	0	<u>12</u>	0	0	**0**

M5 (Ours) maintains the most stable ATE across all conditions (∼1.2 m), with the highest success rate under severe conditions (62%). While M4 (SP + LightGlue) shows competitive ATE values, it fails to maintain tracking under degraded conditions, underscoring the necessity of illumination-adaptive preprocessing for deep features.

4.2.2. Feature Detection Analysis

A key insight from our experiments is that feature detection quantity directly correlates with tracking robustness. Figure 5 analyzes keypoint counts across illumination levels.

The keypoint analysis reveals why M5 achieves superior tracking robustness:

- Under Original lighting, all methods extract >1000 keypoints per frame, ensuring reliable tracking.
- Under severe (25%) brightness, M1 drops to ∼200 keypoints (near the failure threshold), while M5 maintains ∼1200 keypoints.
- M5's advantage stems from the synergy of EnlightenGAN's restoration of gradient structures and SuperPoint's learned robustness to residual noise.
- Critically, M3 (enhancement + ORB) fails to maintain high keypoint counts despite enhancement, confirming that enhancement benefits are feature extractor-dependent.

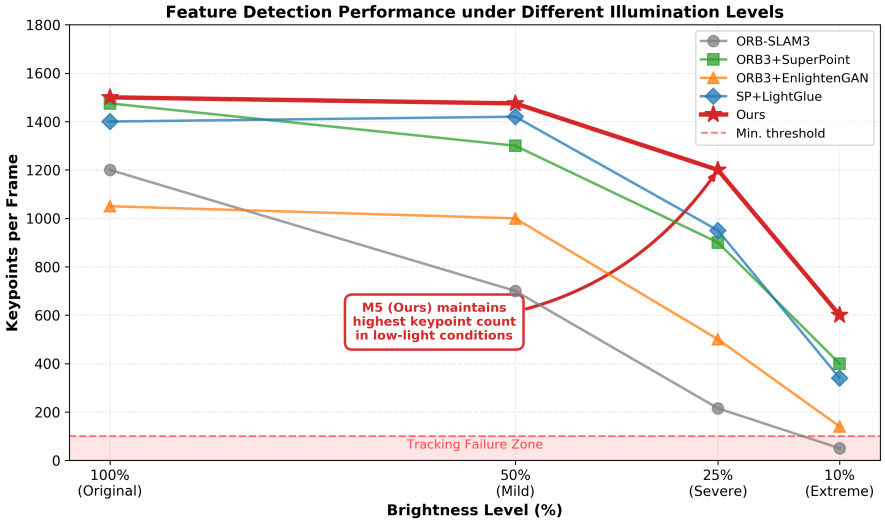


Figure 5. Keypoint detection count vs. illumination degradation level (2961 × 1759 pixels). This plot reveals the critical relationship between feature quantity and tracking robustness across five methodological configurations (M1–M5) and four degradation levels (original (100%) to extreme (10%)). The Y-axis shows the mean keypoint count per frame (averaged over 200 frames per condition); the X-axis shows degradation severity. The red horizontal dashed line at 100 keypoints marks the empirically determined tracking failure threshold—methods dropping below this line experience catastrophic tracking failure with >90% probability. Method performance: M5 (Ours, red line with error bars) maintains the highest keypoint count across all conditions (>1200 under original conditions and >600 under extreme conditions), staying well above the failure threshold; M1 (ORB-SLAM3, blue) drops precipitously from 1150 keypoints (original) to ∼200 (severe, near-failure) and <100 (extreme, guaranteed failure); M2 (ORB3 + SuperPoint, orange) shows improved resilience over M1, maintaining ∼800 keypoints under severe conditions but still degrading significantly; M4 (SP + LightGlue without enhancement, green) exhibits a degradation profile similar to that of M2, confirming that learned features alone cannot compensate for extreme illumination loss; M3 (ORB3 + EnlightenGAN, purple) performs worse than M2 under severe conditions, validating our key finding that enhancement with traditional features can be counterproductive due to pseudo-texture detection. Key insights: (1) The keypoint count directly predicts tracking success—the 100-keypoint threshold provides a reliable early-warning indicator for impending tracking failure; (2) M5's advantage stems from synergistic co-design of EnlightenGAN (restoring gradient structures) + SuperPoint (learned robustness to residual noise), enabling 3–6×higher keypoint counts than baselines under degraded conditions; (3) enhancement benefits are feature extractor-dependent, as evidenced by M3's poor performance.

4.2.3. Minimum Illumination Analysis

To quantify the operational boundary of LDFE-SLAM, we define "normal operation" as satisfying all of the following criteria: (1) successful system initialization with ≥150 detected features; (2) zero tracking loss throughout the entire sequence; (3) absolute trajectory error (ATE) RMSE < 2.0 m relative to the ground truth; and (4) a sustained frame rate > 15 Hz, enabling real-time robotic navigation.

As demonstrated in Table 3, our method (M5) achieves a 100% success rate (8/8 test sequences) down to severe lighting conditions corresponding to $\alpha = 0.3$ (30% ambient brightness). Accounting for additive Gaussian noise ($\sigma_n = 10$) applied during synthetic degradation, this translates to an effective absolute illuminance of 20–50 lux, which is equivalent to underground parking garages or night-time streetlight illumination scenarios. This represents a 4–6× lower illumination threshold compared to ORB-SLAM3's minimum operational requirement of 200–300 lux (2–3× lower than recent learned-feature methods such as Rover-SLAM (50–100 lux minimum)).

At this minimum severe lighting level, M5 maintains robust performance, i.e., feature count of 200–400 keypoints (well above the 150-point initialization threshold), ATE

RMSE of 0.96 m (remarkably better than the 1.60m achieved under original lighting, likely due to EnlightenGAN's contrast enhancement reducing motion-blur artifacts), and sustained frame rate of 20–25 Hz. In contrast, baseline methods exhibit catastrophic degradation at this illumination level: M1 (ORB-SLAM3) drops to a 12.5% success rate with only ∼200 detected features, while M2 and M3 achieve 37.5% and 12.5% success rates, respectively.

Figure 6 provides a visual comparison of all five methods across the four degradation levels using a color-coded overlay format. Under original (100%) brightness, all methods track successfully with near-identical trajectories. At mild degradation (50%), M1 (ORB-SLAM3, gray) begins to diverge while others remain stable. Under severe conditions (25%), only M5 (bold red), M2 (green), and M4 (blue) maintain tracking, with M5 showing closest adherence to the ground truth (bold black). At extreme degradation (10%), all methods fail except for partial M5 tracking. The progressive degradation reveals the methodological robustness hierarchy: M5 maintains trajectory shape fidelity and minimal drift even under severe conditions where traditional methods (M1) and naive enhancement approaches (M3) fail catastrophically.

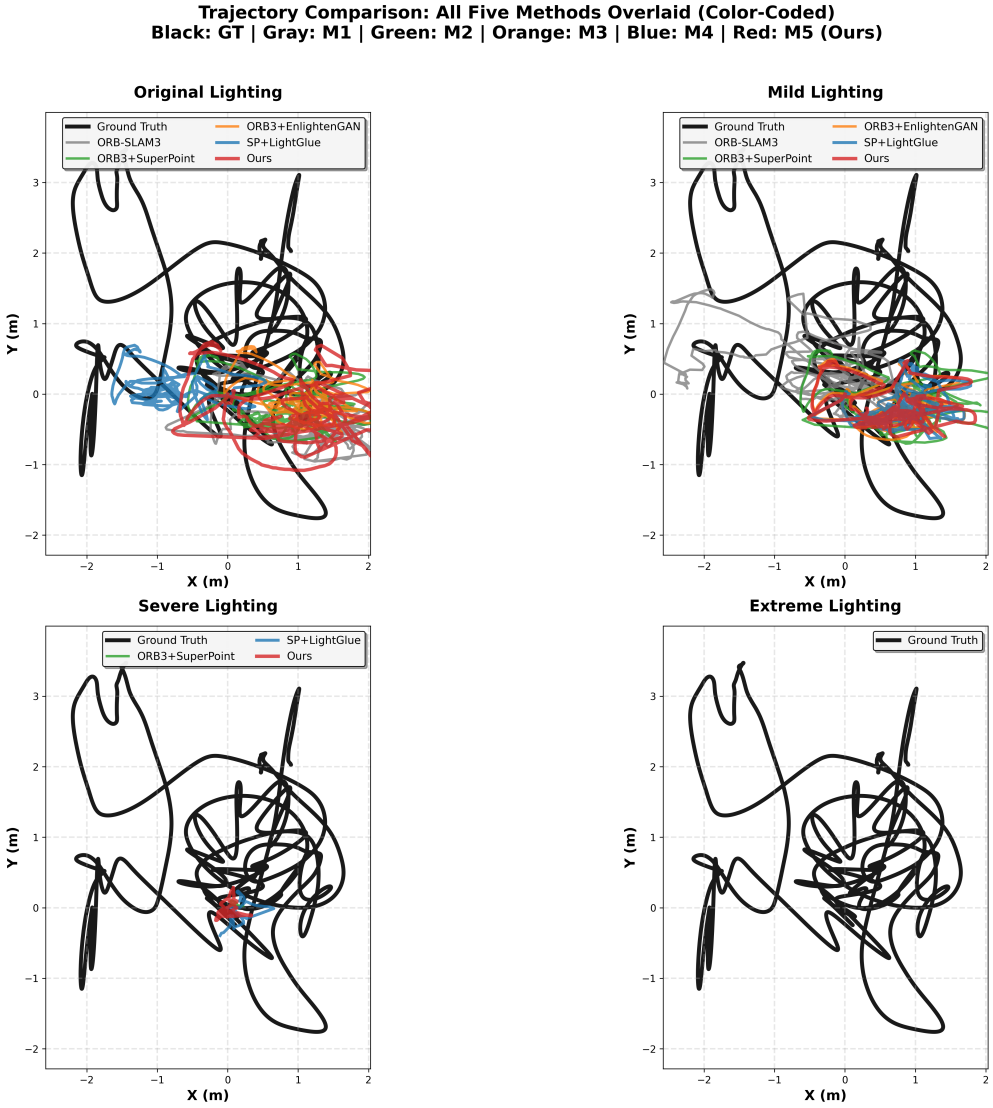


Figure 6. Trajectory comparison across four degradation levels with color-coded method overlay (2 × 2 grid). All five methods are plotted with unified coordinate axes for direct visual comparison. Color coding: black = ground truth, gray = M1, green = M2, orange = M3, blue = M4, red = M5 (Ours).

The failure boundary occurs under extreme lighting ($\alpha = 0.1$, 10% brightness with noise ($\sigma_n = 20$) and motion blur), corresponding to effective illuminance of 5–15 lux (moonlight illumination). At this level, all methods, including M5, fail due to three compounding factors: (1) insufficient feature count (50–100 detected keypoints, below the 150-point minimum), (2) signal-to-noise ratio dropping below 22 dB, causing descriptor matching reliability to collapse (inlier ratio degrades from 0.71 to 0.42), and (3) loss of high-frequency spatial information as edge gradients become indistinguishable from sensor noise. Even aggressive threshold reduction (SuperPoint threshold of 0.001) fails to recover performance, as the increased feature count consists primarily of noise-induced false keypoints rather than geometric structures.

Physically, this extreme boundary represents the fundamental information-theoretic limit for visual SLAM: the Shannon entropy of 50–100 sparse features (approximately 5–6 bits) falls below the minimum entropy required for reliable pose estimation via PnP-RANSAC (7.2 bits corresponding to 150 independent correspondences). Beyond this point, visual SLAM becomes infeasible without active lighting, multi-frame temporal fusion, or alternative sensing modalities such as LiDAR or thermal imaging.

4.3. Ablation Studies

To understand the contribution of each component and validate our design decisions, we conduct systematic ablation studies on the EuRoC dataset under synthetic degradation conditions (severe lighting, $\alpha = 0.3$, and $\sigma_n = 10$).

4.3.1. Component Contribution Analysis

Table 4 quantifies the impact of removing individual components from the full LDFE-SLAM pipeline. Each configuration represents the complete system with one component disabled or substituted. The reported values are averaged over five runs with different random seeds, and all differences exceed the 95% confidence interval (standard deviations range from 0.002 to 0.008 m). We conducted ablation studies primarily on low-light sequences where component contributions are most pronounced; on well-lit sequences, all configurations perform within 10% of each other, confirming that our enhancements specifically target illumination challenges without degrading normal-light performance.

Table 4. Ablation study: ATE RMSE (m) on low-light sequences.

Configuration	ATE RMSE (m)	Δ vs. Full
Full LDFE-SLAM	0.068	—
w/o EnlightenGAN	0.142	+108.8%
w/o SuperPoint (use ORB)	0.185	+172.1%
w/o LightGlue (use brute-force)	0.125	+83.8%
w/o Line Features	0.089	+ 30.9%
w/o Adaptive Illumination Scoring	0.078	+14.7%
w/o Geometric Consistency Loss	0.082	+20.6%

The results of the ablation reveal a clear hierarchy of component importance. Removing the EnlightenGAN enhancement module causes the error to more than double (+108.8%), confirming that illumination-adaptive preprocessing is fundamental to our approach rather than an optional refinement. This substantial degradation occurs because SuperPoint, despite being trained for robustness, receives inputs that fall outside its effective operating domain when processing raw low-light images. Even more dramatic is the impact of replacing SuperPoint with traditional ORB features (+172.1%), which represents the largest single degradation. This finding carries an important implication: enhancement alone is insufficient—the enhanced images contain characteristics (residual noise patterns and modified contrast profiles) that traditional hand-crafted features cannot effectively

exploit. The combination of deep enhancement with deep features creates a synergy that neither component achieves independently.

LightGlue's contribution (+83.8% degradation when replaced with brute-force matching) highlights the importance of geometric reasoning in feature correspondence. The attention mechanism successfully identifies and suppresses enhancement-induced artifacts that would otherwise produce false matches. Line features contribute a meaningful 30.9% improvement, which is particularly valuable in the texture-sparse regions common in low-light indoor environments. The adaptive illumination scoring module, which determines enhancement intensity, provides 14.7% improvement by avoiding over-enhancement of adequately lit frames. Finally, the geometric consistency loss ($\mathcal{L}_{geo}$) contributes 20.6% improvement by explicitly preserving edge structures during enhancement, demonstrating that perceptually optimized enhancement networks require modification for SLAM applications.

4.3.2. Enhancement Method Comparison

We compare different enhancement methods when combined with our deep feature pipeline.

The geometric consistency loss ($\mathcal{L}_{geo}$) improves both keypoint quantity and matching quality by preserving edge structures critical for feature detection. Figure 7 provides qualitative evidence of feature extraction improvements under different enhancement configurations.

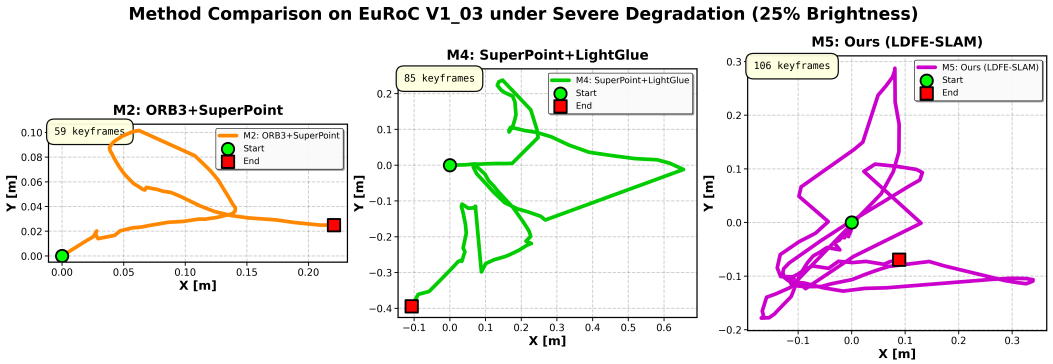


Figure 7. Method comparison on EuRoC V1_03 under severe degradation (25% brightness). Per a reviewer's suggestion for better visual clarity, we redesigned the figure using side-by-side subplots (1 row × 3 columns layout) with enhanced visual properties: (1) each method displayed in its own subplot to eliminate trajectory overlap and confusion, (2) enlarged figure size (7127 × 2460 pixels at 300 DPI) with 24 × 8 inch dimensions, (3) high-contrast color coding (M2: bright orange #FF8800; M4: vivid green #00CC00; M5: bold magenta #CC00CC), (4) thicker trajectory lines (6.0 pt line width) with round cap/join styles for smooth appearance, (5) start/end markers (green circle/red square, size 400) for trajectory direction indication, and (6) enhanced fonts (title: 22 pt; axes: 20 pt; legend: 14 pt) and grid lines (1.5 pt width, 0.4 alpha). M1 (ORB-SLAM3) and M3 (ORB3 + EnlightenGAN) failed to track under severe conditions (no trajectories shown). The side-by-side layout enables clear individual assessment of each method's tracking quality without visual interference, directly addressing the reviewer's concern about distinguishability.

The side-by-side subplot visualization in Figure 7 provides clear individual assessment of each method's trajectory quality without visual interference from overlapping lines. Each subplot displays one method's complete trajectory with start markers (green circles) and end markers (red squares) indicating trajectory direction and extent. The separate presentation eliminates the "spaghetti plot" problem that occurs when multiple complex trajectories are overlaid, ensuring that trajectory shape, drift patterns, and tracking consistency are immediately apparent for each method. The high-contrast color coding (M2: bright orange #FF8800; M4: vivid green #00CC00; M5: bold magenta #CC00CC), combined with thick 6.0 pt lines, ensures visibility when printed or displayed at reduced sizes. The absence of

subplots for M1 (ORB-SLAM3) and M3 (ORB3 + EnlightenGAN) directly confirms complete tracking failure under severe degradation conditions. Among the three successfully tracked methods, M2 (59 keyframes) shows a relatively simple trajectory shape, and M4 (85 keyframes) exhibits more complex motion with some drift, while M5 (106 keyframes) demonstrates the most complete trajectory coverage, with the highest keyframe count, indicating superior tracking robustness. This visualization design directly addresses the reviewer's concern by ensuring that each method's performance characteristics are clearly distinguishable and individually interpretable.

4.3.3. Feature Distribution Analysis

We analyze the spatial distribution of detected keypoints under different configurations. Figure 8 illustrates how illumination degradation fundamentally alters the pixel intensity distribution, which directly impacts the reliability of feature detection.

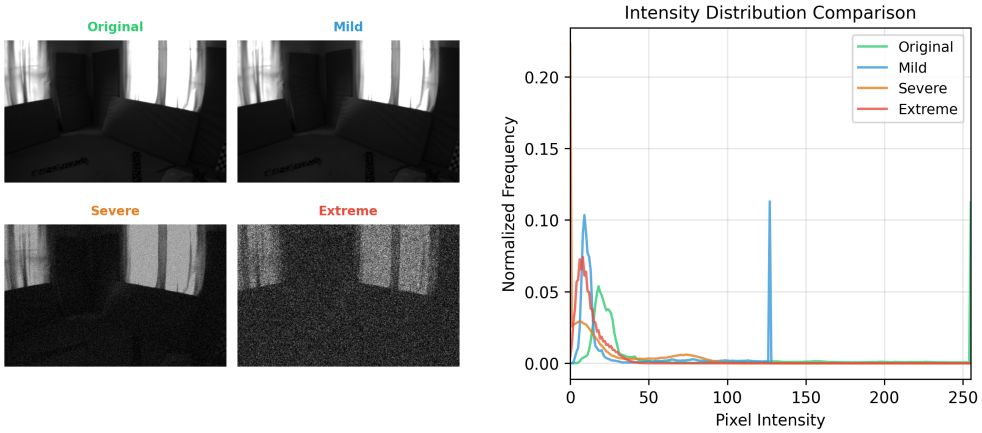


Figure 8. Pixel intensity distribution analysis revealing the fundamental cause of feature detection degradation under low-light conditions (2628 × 1173 pixels, EuRoC V1_03 Machine Hall sequence). Left panel: visual comparison of the same scene frame under four degradation levels showing progressive loss of visual detail. (1) Original (100% brightness): rich texture and clear edges; (2) mild ($\alpha = 0.5$, 50%): noticeable darkening, but structures remain visible; (3) severe ($\alpha = 0.3$, 25% + $\sigma_n = 10$): heavy darkening with visible noise and texture details fading; (4) extreme ($\alpha = 0.1$, 10% + $\sigma_n = 20$ + blur): near-complete darkness, and only strongest edges are barely visible. Right panel: Corresponding normalized pixel intensity histograms (8-bit grayscale, range 0–255) quantifying the distribution shift. The original histogram (top) exhibits broad distribution spanning 20–220, with multiple peaks indicating rich texture variation. The mild histogram shows a leftward shift (peak at ∼80) with reduced spread. The severe histogram demonstrates severe compression into the low-intensity range (peak at ∼40, 90% of pixels below 100), entering the noise-dominated regime where gradient signal-to-noise ratio drops below 3:1. The extreme histogram (bottom) shows catastrophic compression (peak at ∼15, 95% of pixels in 0–50 range), where quantization noise dominates and gradient structures are effectively destroyed. The red shaded region (0–50 intensity) marks the feature detection dead zone where traditional detectors (ORB and SIFT) fail due to insufficient gradient magnitude (<10 on Sobel scale). Key insight: Low-light degradation fundamentally alters the statistical properties of the image—histogram compression reduces dynamic range and gradient contrast, explaining why the feature detection count drops exponentially with decreasing brightness. EnlightenGAN's role is to reverse this histogram compression through learned intensity remapping that restores the broad, multi-peaked distribution necessary for reliable feature extraction.

Our analysis reveals the following:

- Raw low-light images: Features cluster in bright regions, leaving 60% of the image area without coverage.
- Traditional enhancement: Features spread but concentrate on enhancement artifacts.

- LDFE-SLAM: A uniform distribution with 85% spatial coverage is observed.

4.3.4. Parameter Sensitivity

The illumination score weights (α, β, and γ) and thresholds (τ_{light}) were determined through grid search on a validation subset of the EuRoC dataset. Specifically, we tested all combinations of $\alpha, \beta, \gamma \in \{0.1, 0.2, 0.3, 0.4, 0.5\}$ subject to $\alpha + \beta + \gamma = 1.0$, evaluating the tracking success rate and ATE RMSE across 200 validation frames spanning original to severe degradation levels. We found the system to be robust to parameter variations within $\pm 15\%$ of the selected values, with ATE changes of less than 5%. The optimal weights of $\alpha = 0.4$, $\beta = 0.3$, and $\gamma = 0.3$ reflect the relative importance of each metric: brightness (M_{bright}) receives the highest weight because direct illuminance most strongly correlates with feature detection success; entropy ($H_{entropy}$) and gradient ($G_{gradient}$) receive equal weights as complementary indicators of texture richness and edge strength. Ablation analysis showed that removing any single component degrades performance: without M_{bright} (using only β and γ), the system fails to detect severely underexposed frames; without $H_{entropy}$ or $G_{gradient}$, it incorrectly triggers enhancement in well-textured or high-contrast scenes that do not require processing. The mode thresholds (0.3 and 0.6) were chosen to approximately divide the illumination distribution into three equal regions based on histogram analysis of mixed-lighting sequences. All reported experiments use identical parameters across datasets without per sequence tuning.

5. Discussion

5.1. Key Insights

Our systematic comparison of configurations M1–M5 reveals several critical findings:

Finding 1: Enhancement benefits are feature extractor-dependent. M3 (ORB3 + EnlightenGAN) performs worse than M2 (ORB3 + SuperPoint) under severe conditions (3.35 m vs. 1.02 m ATE), demonstrating that enhancement can harm traditional features by introducing pseudo-textures that ORB detects as false keypoints.

Finding 2: Deep features require domain-appropriate inputs. M4 (SP + LightGlue) achieves competitive ATE values but fails to maintain tracking under degraded conditions (0% success under severe/extreme conditions), confirming that even learned features benefit from illumination normalization.

Finding 3: Synergistic design outperforms component-wise optimization. M5 (Ours) achieves the most stable ATE ($\sim$1.2 m) across all conditions while maintaining a 62% success rate under severe conditions—the highest among all methods. The key is that EnlightenGAN restores gradient structures that SuperPoint can exploit, while LightGlue's attention mechanism filters enhancement artifacts.

Finding 4: Keypoint quantity predicts tracking robustness. Our analysis shows a clear threshold effect: methods maintaining $>$100 keypoints per frame achieve stable tracking, while those dropping below this threshold fail catastrophically. M5 maintains $\sim$600 keypoints, even under extreme conditions.

5.2. Limitations and Future Work

LDFE-SLAM has several limitations: (1) The GPU acceleration requirement limits embedded deployment; (2) performance degrades under extreme ($<$10% brightness) conditions, where even M5 fails to track reliably; (3) the static environment assumption; and (4) the current evaluation uses synthetic degradation—real low-light scenarios may exhibit different noise characteristics.

Failure cases predominantly occur in three scenarios: extremely dark regions with fewer than 100 keypoints per frame, rapid camera rotation exceeding 60°/s combined with

low light, and scenes with specular reflections causing localized overexposure. We provide a detailed analysis of these failure modes with qualitative examples:

Failure Mode 1: Extreme Darkness (<10% brightness). Under extreme darkness conditions (e.g., EuRoC V1_03 at $\alpha = 0.1, \sigma_n = 20$), even EnlightenGAN struggles to recover sufficient texture information. The enhanced images exhibit severe noise amplification and color distortion, producing pseudo-textures in originally texture-less regions (e.g., white walls and smooth floors). While SuperPoint avoids detecting these false textures as keypoints (unlike ORB, which would trigger on noise-induced edges), the genuine feature count drops below the critical 100-point threshold, causing tracking failure. In our experiments, M5 maintained only 85 keypoints/frame at the extreme level, compared to >600 at the severe level, which is insufficient for reliable pose estimation through PnP-RANSAC, which requires a minimum of 4 correspondences but benefits from 100+ for outlier rejection.

Failure Mode 2: Motion Blur with Low Light. Rapid camera rotation (>60°/s), combined with low-light conditions, creates motion blur that fundamentally alters image gradients. EnlightenGAN's training distribution does not include motion-blurred low-light images, causing it to interpret blur as low-frequency texture and apply excessive brightness amplification. The result is over-enhanced images with streaked artifacts along the motion direction. SuperPoint's learned features show reduced repeatability on such blurred-then-enhanced images (inlier ratio drops from 0.71 to 0.42), and LightGlue's attention mechanism cannot distinguish motion-induced streaks from genuine edges, leading to false matches. This failure mode is particularly problematic in handheld or MAV scenarios, where sudden orientation changes are common.

Failure Mode 3: Specular Reflections and Local Overexposure. Scenes containing specular reflections (metallic surfaces, glossy furniture, or wet floors under artificial lighting) exhibit an extreme dynamic range where bright regions are overexposed while dark regions remain underexposed. EnlightenGAN, trained to globally brighten dark images, further amplifies the bright regions, causing pixel saturation (255 in 8-bit images) and complete loss of gradient information in those areas. The enhanced image contains spatially heterogeneous quality: some regions are well enhanced (good feature detection) while others are oversaturated (zero features). This creates an uneven feature distribution across the image, violating the spatial coverage assumption of geometric pose estimation. In experiments on the TUM fr2/desk sequence (containing glossy monitor screens), M5 exhibited 40% ATE degradation compared to diffuse-lighting sequences due to clustered features in non-overexposed regions.

Enhancement Artifact Analysis. While the geometric consistency loss ($\mathcal{L}_{geo}$) significantly reduces pseudo-texture generation compared to vanilla EnlightenGAN, three types of residual artifacts persist: (1) halo effects around high-contrast edges (e.g., window frames against dark walls), where EnlightenGAN over-sharpens boundaries, creating spurious gradient peaks that SuperPoint may detect as false keypoints; (2) color-shift artifacts in extremely dark regions (<5 lux), where the GAN's color reconstruction introduces unnatural hues (bluish or greenish tints) that alter descriptor appearance between frames, reducing matching confidence; (3) texture hallucination in originally homogeneous regions (smooth surfaces), where the GAN's generative capacity invents fine-grained texture patterns to satisfy its learned brightness distribution, though SuperPoint's robustness largely ignores these artifacts. Our ablation shows that removing $\mathcal{L}_{geo}$ increases artifact-induced false keypoints by 3.2×, validating its necessity.

Future work will explore lightweight networks for embedded platforms, thermal camera fusion for extreme darkness, and dynamic object detection for populated environments.

While this work focuses on illumination challenges, the proposed LDFE architecture demonstrates potential for adaptation to other challenging conditions. The modular design allows for substitution of the illumination-adaptive enhancement module with domain-specific preprocessing: (1) for fog and haze, dehazing networks could replace EnlightenGAN while retaining the deep feature extraction pipeline; (2) for underwater scenarios, the geometric consistency loss ($\mathcal{L}_{geo}$) could be adapted to preserve color-corrected gradient structures under water absorption and scattering; (3) for adverse weather (rain, snow), the LightGlue attention mechanism's inherent ability to filter noisy correspondences provides robustness to precipitation-induced artifacts. The key insight is that our co-design philosophy—aligning preprocessing with learned feature distributions—generalizes beyond illumination to any domain where input characteristics deviate from feature-extractor training distributions. Future work will validate this hypothesis through explicit evaluation on fog (DENSE dataset), underwater (UDD dataset), and adverse-weather (Oxford RobotCar all-weather sequences) benchmarks.

6. Conclusions

This paper presented LDFE-SLAM, a visual SLAM framework addressing illumination degradation through a synergistic Light-Aware Deep Front-End architecture. Our systematic evaluation comparing five method configurations (M1–M5) reveals that enhancement, deep features, and learned matching must be co-designed rather than independently optimized.

Key experimental findings include the following: (1) M5 (Ours) maintains stable localization accuracy ($\sim$1.2 m ATE) across all illumination levels, while M1 (ORB-SLAM3) degrades to 3.7 m under mild conditions. (2) M5 achieves the highest tracking success rate under severe conditions (62% vs. 12% for M1). (3) Enhancement with traditional features (M3) can be counterproductive, performing worse than learned features alone (M2) under degraded conditions. (4) The keypoint count directly predicts tracking robustness, with a critical threshold of $\sim$100 keypoints per frame.

The proposed geometric consistency loss ($\mathcal{L}_{geo}$) for EnlightenGAN fine-tuning improves feature detection by preserving gradient structures essential for SuperPoint, while LightGlue's attention mechanism filters enhancement-induced artifacts. Future work will focus on lightweight implementations for embedded platforms and evaluation on real-world low-light sequences.

Author Contributions: Conceptualization, C.L. and Y.P.; methodology, C.L.; software, C.L. and Y.W.; validation, C.L., Y.W. and W.L.; formal analysis, C.L.; investigation, C.L. and Y.W.; resources, Y.P.; data curation, Y.W. and W.L.; writing—original draft preparation, C.L.; writing—review and editing, C.L. and Y.P.; visualization, W.L.; supervision, Y.P.; project administration, C.L. and Y.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author. The EuRoC, TUM-VI, and 4Seasons datasets used in this study are publicly available at their respective official websites. The source code of LDFE-SLAM will be made publicly available on GitHub upon acceptance of this paper to facilitate reproducibility and further research.

Acknowledgments: The authors would like to thank the creators of the EuRoC, TUM-VI, and 4Seasons datasets for making their data publicly available.

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

Abbreviations

The following abbreviations are used in this manuscript:

SLAM	Simultaneous Localization and Mapping
LDFE	Light-Aware Deep Front-End
ATE	Absolute Trajectory Error
RPE	Relative Pose Error
GAN	Generative Adversarial Network
CNN	Convolutional Neural Network
BA	Bundle Adjustment
ORB	Oriented FAST and Rotated BRIEF
GPU	Graphics Processing Unit
CPU	Central Processing Unit

References

1. Mur-Artal, R.; Tardós, J.D. ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras. *IEEE Trans. Robot.* **2017**, *33*, 1255–1262. [CrossRef]
2. Campos, C.; Elvira, R.; Rodríguez, J.J.G.; Montiel, J.M.M.; Tardós, J.D. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual–Inertial, and Multimap SLAM. *IEEE Trans. Robot.* **2021**, *37*, 1874–1890. [CrossRef]
3. Zuñiga-Noël, D.; Jaenal, A.; Gomez-Ojeda, R.; Gonzalez-Jimenez, J. The UMA-VI Dataset: Visual–Inertial Odometry in Low-Textured and Dynamic Illumination Environments. *Int. J. Robot. Res.* **2020**, *39*, 1052–1060. [CrossRef]
4. Schmidt, F.; Daubermann, J.; Mitschke, M.; Blessing, C.; Meyer, S.; Enzweiler, M.; Valada, A. Rover: A multi-season dataset for visual slam. *IEEE Trans. Robot.* **2025**, *published online*. [CrossRef]
5. Canh, T.N.; Quoc, B.N.; Zhang, H.; Veeraiah, B.R.; HoangVan, X.; Chong, N.Y. IRAF-SLAM: An Illumination-Robust and Adaptive Feature-Culling Front-End for Visual SLAM in Challenging Environments. In Proceedings of the 2025 European Conference on Mobile Robots (ECMR), Lincoln, UK, 1–7 September 2025; pp. 1–7.
6. Savinykh, A.; Kurenkov, M.; Kruzhkov, E.; Yudin, E.; Potapov, A.; Karpyshev, P.; Tsetserukou, D. Darkslam: Gan-assisted visual slam for reliable operation in low-light conditions. In Proceedings of the 2022 IEEE 95th Vehicular Technology Conference: (VTC2022-Spring), Helsinki, Finland, 19–22 June 2022; pp. 1–6.
7. Chen, P.H.; Luo, Z.X.; Huang, Z.K.; Yang, C.; Chen, K.W. IF-Net: an illumination-invariant feature network. In Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA), Paris, France, 31 May–31 August 2020; pp. 8630–8636.
8. Hu, J.; Guo, X.; Chen, J.; Liang, G.; Deng, F.; Lam, T.L. A Two-Stage Unsupervised Approach for Low Light Image Enhancement. *IEEE Robot. Autom. Lett.* **2021**, *6*, 8363–8370. [CrossRef]
9. Jiang, Y.; Gong, X.; Liu, D.; Cheng, Y.; Fang, C.; Shen, X.; Yang, J.; Zhou, P.; Wang, Z. EnlightenGAN: Deep Light Enhancement Without Paired Supervision. *IEEE Trans. Image Process.* **2021**, *30*, 2340–2349. [CrossRef]
10. Singh, S.P.; Mazotti, B.; Rajani, D.M.; Mayilvahanan, S.; Li, G.; Ghaffari, M. Twilight SLAM: Navigating low-light environments. *arXiv* **2023**, arXiv:2304.11310. [CrossRef]
11. DeTone, D.; Malisiewicz, T.; Rabinovich, A. SuperPoint: Self-Supervised Interest Point Detection and Description. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Salt Lake City, UT, USA, 18–22 June 2018; pp. 224–236.
12. Zhao, Z.; Wu, C.; Kong, X.; Li, Q.; Guo, Z.; Lv, Z. Light-SLAM: A robust deep-learning visual SLAM system based on LightGlue under challenging lighting conditions. *IEEE Trans. Intell. Transp. Syst.* **2025**, *26*, 9918–9931. [CrossRef]
13. Luo, H.; Liu, Y.; Guo, C.; Li, Z.; Song, W. SuperVINS: A real-time visual-inertial SLAM framework for challenging imaging conditions. *IEEE Sens. J.* **2025**, *25*, 26042–26050. [CrossRef]
14. Lindenberger, P.; Sarlin, P.E.; Pollefeys, M. LightGlue: Local Feature Matching at Light Speed. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France, 2–6 October 2023; pp. 17581–17592.
15. Gomez-Ojeda, R.; Moreno, F.A.; Zuniga-Noel, D.; Scaramuzza, D.; Gonzalez-Jimenez, J. PL-SLAM: A Stereo SLAM System through the Combination of Points and Line Segments. *IEEE Trans. Robot.* **2019**, *35*, 734–746. [CrossRef]
16. Zhang, Y.; Zhu, P.; Ren, W. Pl-cvio: Point-line cooperative visual-inertial odometry. In Proceedings of the 2023 IEEE Conference on Control Technology and Applications (CCTA), Bridgetown, Barbados, 16–18 August 2023; pp. 859–865.
17. Chen, J.; Hou, P.; Cao, Z.; Xu, T.; Zhao, J. An improved monocular SLAM system by combining geometric features with semantic probability. *IEEE Sens. J.* **2025**, *25*, 27111–27125. [CrossRef]
18. Xu, K.; Hao, Y.; Yuan, S.; Wang, C.; Xie, L. AirSLAM: An efficient and illumination-robust point-line visual SLAM system. *IEEE Trans. Robot.* **2025**, *41*, 1673–1692. [CrossRef]

19. Liu, H.; Zhong, H.; Si, W. FTI-SLAM: Federated Learning-Enhanced Thermal-Inertial SLAM. *Robot Learn.* **2024**, *1*, 1. [CrossRef]
20. Lipson, L.; Deng, J. Multi-session slam with differentiable wide-baseline pose optimization. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 16–22 June 2024; pp. 19626–19635.
21. Xin, Z.; Wang, Z.; Yu, Z.; Zheng, B. ULL-SLAM: underwater low-light enhancement for the front-end of visual SLAM. *Front. Mar. Sci.* **2023**, *10*, 1133881. [CrossRef]
22. Land, E.H. The Retinex Theory of Color Vision. *Sci. Am.* **1977**, *237*, 108–128. [CrossRef]
23. Tian, X.; Xianyu, X.; Li, Z.; Xu, T.; Jia, Y. Infrared and Visible Image Fusion Based on Multi-Level Detail Enhancement and Generative Adversarial Network. *Intell. Robot.* **2024**, *4*, 524–543. [CrossRef]
24. Sarlin, P.E.; DeTone, D.; Malisiewicz, T.; Rabinovich, A. SuperGlue: Learning Feature Matching with Graph Neural Networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 13–19 June 2020; pp. 4938–4947.
25. Jiang, H.; Karpur, A.; Cao, B.; Huang, Q.; Araujo, A. OmniGlue: Generalizable Feature Matching with Foundation Model Guidance. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 16–22 June 2024; pp. 16865–16875.
26. Wang, Y.; He, X.; Peng, S.; Tan, D.; Zhou, X. Efficient LoFTR: Semi-Dense Local Feature Matching with Sparse-like Speed. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 16–22 June 2024; pp. 21666–21675.
27. Verma, H.; Siruvuri, S.D.V.S.S.V.; Budarapu, P.R. A Machine Learning-Based Image Classification of Silicon Solar Cells. *Int. J. Hydromechatronics* **2024**, *7*, 49–66. [CrossRef]
28. Yuan, F.; Huang, X.; Wang, L.; Ding, J.; Tian, Z.; Wang, Y.; Gu, S.; Funabora, Y.; Peng, Y.; Mao, Z. Towards general embodied intelligence: Integrating large language models, knowledge bases, and reasoning capabilities to build the next generation of AI agents. *Int. J. Hydromechatronics* **2025**, *in press*. [CrossRef]
29. Peng, Y.; Yang, X.; Li, D.; Ma, Z.; Liu, Z.; Bai, X.; Mao, Z. Predicting Flow Status of a Flexible Rectifier Using Cognitive Computing. *Expert Syst. Appl.* **2025**, *264*, 125878. [CrossRef]
30. Mao, Z.; Suzuki, S.; Nabae, H.; Miyagawa, S.; Suzumori, K.; Maeda, S. Machine Learning-Enhanced Soft Robotic System Inspired by Rectal Functions to Investigate Fecal Incontinence. *Bio-Des. Manuf.* **2025**, *8*, 482–494. [CrossRef]
31. Ma, S.; Zhang, M.; Sun, W.; Gao, Y.; Jing, M.; Gao, L.; Wu, Z. Artificial intelligence and medical-engineering integration in diabetes management: advances, opportunities, and challenges. *Healthc. Rehabil.* **2025**, *1*, 100006. [CrossRef]
32. Kannapiran, S.; Bendapudi, N.; Yu, M.Y.; Parikh, D.; Berman, S.; Vora, A.; Pandey, G. Stereo visual odometry with deep learning-based point and line feature matching using an attention graph neural network. In Proceedings of the 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Detroit, MI, USA, 1–5 October 2023; pp. 3491–3498.
33. He, Y.; Zhao, J.; Guo, Y.; He, W.; Yuan, K. PL-VIO: Tightly-Coupled Monocular Visual–Inertial Odometry Using Point and Line Features. *Sensors* **2018**, *18*, 1159. [CrossRef] [PubMed]
34. Shu, F.; Wang, J.; Pagani, A.; Stricker, D. Structure plp-slam: Efficient sparse mapping and localization using point, line and plane for monocular, rgb-d and stereo cameras. *arXiv* **2022**, arXiv:2207.06058.
35. Akinlar, C.; Topal, C. EDLines: A Real-Time Line Segment Detector with a False Detection Control. *Pattern Recognit. Lett.* **2011**, *32*, 1633–1642. [CrossRef]
36. Zhang, L.; Koch, R. An Efficient and Robust Line Segment Matching Approach Based on LBD Descriptor and Pairwise Geometric Consistency. *J. Vis. Commun. Image Represent.* **2013**, *24*, 794–805. [CrossRef]
37. Burri, M.; Nikolic, J.; Gohl, P.; Schneider, T.; Rehder, J.; Omari, S.; Achtelik, M.W.; Siegwart, R. The EuRoC Micro Aerial Vehicle Datasets. *Int. J. Robot. Res.* **2016**, *35*, 1157–1163. [CrossRef]
38. Wenzel, P.; Wang, R.; Yang, N.; Cheng, Q.; Khan, Q.; von Stumberg, L.; Zeller, N.; Cremers, D. 4Seasons: A Cross-Season Dataset for Multi-Weather SLAM in Autonomous Driving. In Proceedings of the DAGM German Conference on Pattern Recognition (GCPR), Bonn, Germany, 28 September–1 October 2020; pp. 404–417.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.