TEACHING LOOKUP TABLES TO SEE IN THE DARK: LEARNING COMPLETE MAPPINGS FROM A DIFFUSION TEACHER

Anonymous authors
Paper under double-blind review

ABSTRACT

Low-light image enhancement (LLIE) aims to restore illumination, color, and structural edges in images captured under poor lighting. Existing LUT-based methods construct image mappings by combining basis LUTs, but are constrained by predefined transformations and supervised primarily by queries induced by the input image, leaving many grid points in high-dimensional LUTs weakly constrained. To address this limitation, this work proposes a diffusion-teacher-guided LUT distillation method for learning complete image-adaptive mappings. In the wavelet domain, the low-frequency LUT corrects brightness, contrast, and color, while the high-frequency residual LUT restores structures and suppresses noise. Specifically, target LUTs are first obtained from paired images through constrained inversion, after which a conditional diffusion teacher learns residual corrections from coarse LUTs to the target LUTs. Teacher and student interpolation outputs are then aligned using both image queries and LUT-grid queries, extending supervision to sparsely accessed regions of the LUT domain. At inference, the student predicts complete LUTs without diffusion sampling; CP decomposition provides a compact representation for lookup. Experiments on multiple benchmarks demonstrate state-of-the-art performance across a range of evaluation metrics.

1 INTRODUCTION

Low-light image enhancement (LLIE) aims to improve the visual quality of images captured under poor illumination. As a fundamental preprocessing task, LLIE should achieve high restoration quality while remaining practical for deployment. Existing methods can be broadly divided into efficiency-oriented and quality-oriented approaches Lu et al. (2025); Zhang et al. (2025); Dong et al. (2025); Pei et al. (2025). LYTNet Brateanu et al. (2025) and DarkIR Feijoo et al. (2025) reduce computational costs through lightweight designs while addressing low illumination, noise, and blur. In contrast, CFWD Xue et al. and LatentDark Hu et al. (2026) further improve texture and color restoration with diffusion models, but still rely on iterative denoising. As shown in Figure 1. **Can we retain the restoration capability of diffusion models while representing enhancement with explicit LUT mappings and avoiding iterative denoising at inference?**

Lookup tables (LUTs) are widely used in image signal processors (ISPs) for their efficient and practical mapping mechanism. LUTs can be either predefined or learned. Predefined LUTs rely on manual tuning and generalize poorly across diverse scenes, whereas learned LUTs Fu et al. (2024); Kim et al. (2025) can model more flexible transformations and have been applied to exposure, contrast, and saturation adjustment. WV-LUT Li et al. (2025) combines hierarchical 4D LUTs with a 3D LUT to model local information and cross-channel color relationships, together with global adjustment for exposure and color correction. DPLUT Lin et al. (2025) introduces light-adjustment and noise-suppression LUTs, using pixel-wise curve mappings for brightening and diffusion priors to suppress amplified noise. Despite these advances, applying LUTs to LLIE remains challenging.

First, low-light images often have low contrast and small pixel values, making conventional color LUTs ineffective at adjusting local contrast and prone to color shifts and artifacts. Second, existing LUT methods mainly focus on exposure and color adjustment, with limited ability to handle noise. Brightening can amplify noise in dark regions, while weak textures are easily confused with noise;

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

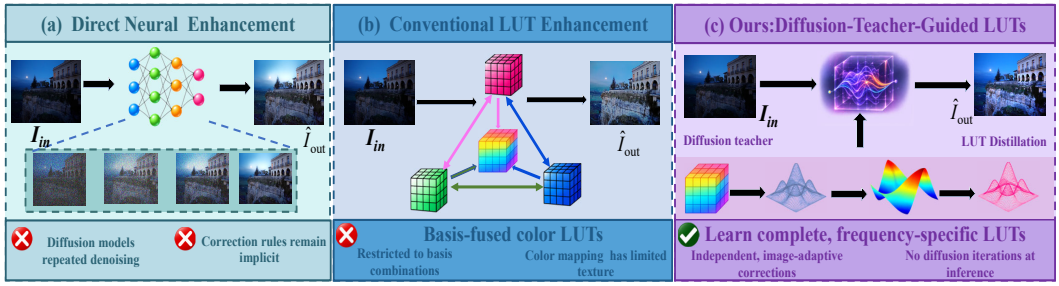


Figure 1: From Direct Neural Enhancement and Conventional LUTs to Diffusion-Teacher-Guided LUT Distillation Framework

excessive denoising may further remove details. Finally, basis-LUT fusion restricts image-adaptive correction to combinations of predefined mappings, limiting independent adjustment across different brightness levels, colors, and textures. Direct LUT prediction avoids this restriction, but supervision based only on enhanced images does not explicitly constrain the complete correction mapping. As shown in Figure 1.

To address these challenges, we propose a diffusion-teacher-guided LUT distillation framework that combines wavelet decomposition with LUT residual diffusion for brightness and color correction, noise suppression, and texture restoration. First, to address the limited local contrast adjustment of conventional color mappings, we construct low-frequency LUTs that incorporate illumination and local contrast information. This enables region-adaptive brightness and color correction. Second, to suppress noise in dark regions while preserving weak textures, we construct high-frequency LUTs that apply residual corrections directly to the high-frequency wavelet components. These LUTs restore texture details while suppressing noise. Finally, to overcome the restrictions of basis LUT fusion, we directly generate complete LUTs conditioned on the input image. This allows corrections for different brightness levels, colors, and textures to be adjusted independently. Specifically, target LUTs are constructed from paired images, and conditional diffusion progressively refines coarse LUTs. We then perform continuous LUT function distillation by aligning teacher and student interpolation outputs at image-derived queries and additional sampled continuous queries. The student ultimately predicts the complete LUTs in a single forward pass.

The main contributions of this work are summarized as follows:

• This work proposes a diffusion teacher for complete LUT mapping learning, which constructs target LUTs through constrained inversion and models residual corrections from coarse LUTs with conditional diffusion, thereby strengthening supervision in regions sparsely covered by image queries.

• This work develops a continuous LUT function distillation method that aligns teacher and student interpolation outputs using both image queries and continuous in-grid queries, thereby extending supervision to query locations unseen in the current image and enabling the student to better preserve the teacher's continuous mapping.

• This work constructs a frequency-specific LUT representation that separately handles illumination and color in the low-frequency components and texture and noise in the high-frequency components, thereby balancing brightness, color, detail, and noise restoration.

2 Methodology

2.1 Overview

Given a low-light RGB image $\mathbf{I}_{\text{in}}$, we formulate enhancement as an image-adaptive LUT inference process based on frequency decomposition. As shown in Figure 2, we first apply the stationary wavelet transform $\mathcal{W}$ independently to the three color channels, producing four spatially aligned subbands $\mathcal{X} = \{\mathbf{X}_b\}_{b\in\mathcal{B}}$, where $\mathcal{B} = \{\text{LL}, \text{LH}, \text{HL}, \text{HH}\}$. A non-uniform 5D LUT is constructed for each subband. For the low-frequency branch, the RGB coefficients, local illumination, and local contrast form the five-dimensional query $\mathbf{q}_{\text{LL}}(\mathbf{p})$, and the LUT directly predicts the corrected low-

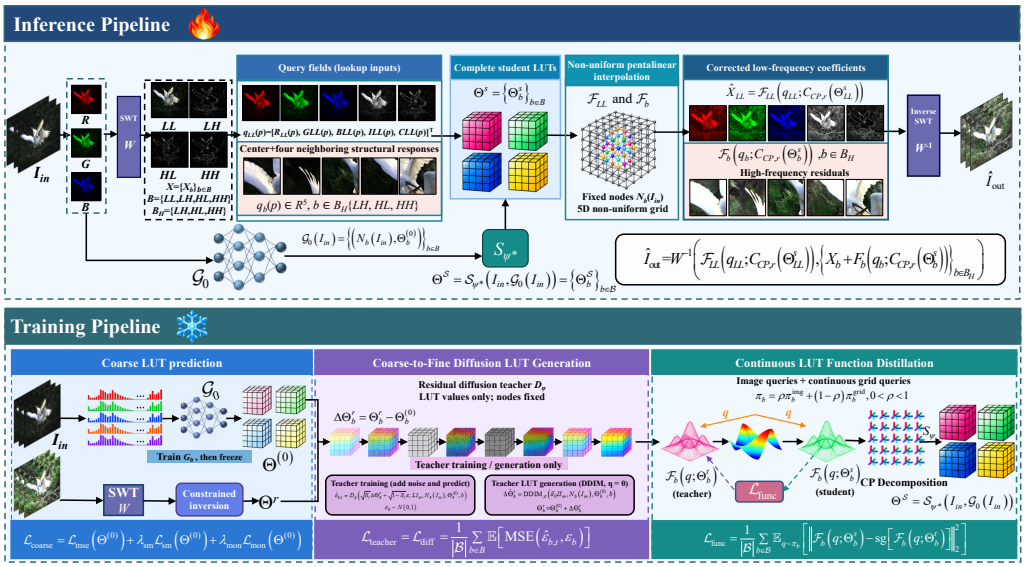


Figure 2: Overview of the Proposed Diffusion-Teacher-Guided LUT Distillation Framework.

frequency coefficients. For each high-frequency branch, $\mathbf{q}_b(\mathbf{p})$ is formed by structural responses at the center and its four neighboring positions, and the corresponding LUT predicts a three-channel residual. The low-frequency branch therefore corrects brightness and color, while the high-frequency branches restore structural details and suppress noise. For each subband, the coarse LUT $\mathbf{\Theta}^{(0)}$, target LUT $\mathbf{\Theta}^{r}$, teacher LUT $\mathbf{\Theta}^{t}$, and student LUT $\mathbf{\Theta}^{s}$ use the same image-adaptive non-uniform nodes and define continuous mappings through non-uniform pentalinear interpolation. During training, the coarse predictor $\mathcal{G}_0$ first predicts the non-uniform nodes and coarse LUT values from the input image. Using paired low-light and normal-light images, we then obtain the target LUT $\mathbf{\Theta}^{r}$ through constrained inversion. Conditioned on the input image, non-uniform nodes, coarse LUT, and subband identity, the diffusion teacher D_ϕ learns the residual between the target and coarse LUTs. The predicted residual is added to $\mathbf{\Theta}^{(0)}$ to form the refined teacher LUT $\mathbf{\Theta}^{t}$. The student model $\mathcal{S}_\psi$ uses the interpolated teacher outputs at continuous query locations as distillation targets and predicts the complete LUTs $\mathbf{\Theta}^{s}$ in a single forward pass. At inference, the diffusion teacher is removed, while $\mathcal{G}_0$ and $\mathcal{S}_\psi$ are retained. The student LUTs are then reparameterized by rank-r CP decomposition and evaluated through non-uniform pentalinear interpolation. The corrected subbands are finally reconstructed into the enhanced image by the inverse wavelet transform.

Let $\mathcal{C}_{\mathrm{CP},r}$ denote the rank-r CP reparameterization, and let $\mathcal{F}_{\mathrm{LL}}$ and $\mathcal{F}_b$ denote the low and high-frequency mapping, respectively. The overall inference process is given by:

$$\widehat{\mathbf{I}}_{\mathrm{out}} = \mathcal{W}^{-1}\Big(\mathcal{F}_{\mathrm{LL}}(\mathbf{q}_{\mathrm{LL}}; \mathcal{C}_{\mathrm{CP},r}(\mathbf{\Theta}^{s}_{\mathrm{LL}})),\, \{\mathbf{X}_b + \mathcal{F}_b(\mathbf{q}_b; \mathcal{C}_{\mathrm{CP},r}(\mathbf{\Theta}^{s}_{b}))\}_{b\in\mathcal{B}_{\mathrm{H}}}\Big). \tag{1}$$

2.2 Frequency-Aware LUT Representation

Motivation. Low-light degradation affects frequency components differently. Brightness and color distortions mainly appear in low-frequency components, whereas texture degradation, edge loss, and noise appear more in high-frequency components. We therefore apply SWT to the RGB channels and construct a low-frequency LUT together with directional high-frequency LUTs. The low-frequency LUT corrects brightness and color, while the high-frequency LUTs restore structures and suppress noise. To better use the limited LUT nodes, we further predict non-uniform nodes for each subband.

Details. At spatial location $\mathbf{p}$, the low-frequency query is defined as $\mathbf{q}_{\mathrm{LL}}(\mathbf{p}) = [R_{\mathrm{LL}}(\mathbf{p}), G_{\mathrm{LL}}(\mathbf{p}), B_{\mathrm{LL}}(\mathbf{p}), I_{\mathrm{LL}}(\mathbf{p}), C_{\mathrm{LL}}(\mathbf{p})]^{\top}$, where I_{LL} and C_{LL} denote local illumination and local contrast, respectively. Together with RGB, these two cues allow similar colors under different local conditions to receive different corrections. The low-frequency LUT directly predicts the cor-

rected RGB coefficients. For each high-frequency subband $b \in \mathcal{B}_{\mathrm{H}}$, we extract a signed gradient s_b from $\mathbf{X}_b$ and construct the five-dimensional query $\mathbf{q}_b(\mathbf{p})$ using the center response and its four axial neighbors. The original high-frequency coefficients $\mathbf{X}_b$ are retained, while the corresponding LUT predicts only a three-channel residual to restore degraded details and suppress noise. To improve node allocation, we introduce a node-prediction branch into $\mathcal{G}_0$. For each query axis, it predicts interval proportions and converts them into a strictly increasing non-uniform node sequence through Softmax normalization and cumulative summation. With the same number of nodes, this assigns denser sampling to regions with concentrated queries or rapidly varying mappings. For signed high-frequency axes, zero is fixed as an anchor, and the remaining nodes are allocated separately to the positive and negative ranges.

Let $\mathcal{N}_{b,d}(\mathbf{I}_{\mathrm{in}})$ denote the node sequence along the d-th query axis of subband b. Their Cartesian product defines $\mathcal{N}_b(\mathbf{I}_{\mathrm{in}}) = \mathcal{N}_{b,1}(\mathbf{I}_{\mathrm{in}}) \times \cdots \times \mathcal{N}_{b,5}(\mathbf{I}_{\mathrm{in}})$. The corresponding LUT is written as $\mathbf{\Theta}_b : \mathcal{N}_b(\mathbf{I}_{\mathrm{in}}) \to \mathbb{R}^3$. For the same input, the coarse, target, teacher, and student LUTs share the same nodes and differ only in their table values. In the following, $\mathbf{\Theta}_b$ denotes both the node domain and its associated table values. The outputs of the frequency-specific LUTs are written as

$$\widehat{\mathcal{X}} = \left(\mathcal{F}_{\mathrm{LL}}(\mathbf{q}_{\mathrm{LL}}; \mathbf{\Theta}_{\mathrm{LL}}), \{\mathbf{X}_b + \mathcal{F}_b(\mathbf{q}_b; \mathbf{\Theta}_b)\}_{b \in \mathcal{B}_{\mathrm{H}}}\right). \tag{2}$$

2.3 Coarse-to-Fine Diffusion LUT Generation

Motivation. A complete 5D LUT is difficult to determine directly from training samples. Training queries typically cover only part of the LUT domain, and each interpolation accesses only a local grid cell, leaving sparsely queried regions weakly constrained. Direct regression provides only a one-step prediction, while basis-LUT fusion is restricted by predefined mappings. We therefore adopt a coarse-to-fine strategy: a coarse LUT first captures the main correction, and conditional diffusion then predicts residuals to refine the initial mapping, particularly in sparsely queried regions and for corrections that are difficult to learn directly.

Details. We train the coarse LUT predictor $\mathcal{G}_0$ and freeze its parameters. Given an input $\mathbf{I}_{\mathrm{in}}$, $\mathcal{G}_0$ predicts the non-uniform nodes and coarse LUTs for the four subbands as $\mathcal{G}_0(\mathbf{I}_{\mathrm{in}}) = \{(\mathcal{N}_b(\mathbf{I}_{\mathrm{in}}), \mathbf{\Theta}_b^{(0)})\}_{b \in \mathcal{B}}$. For a paired sample $(\mathbf{I}_{\mathrm{in}}, \mathbf{I}_{\mathrm{gt}})$, the normal-light SWT representation is denoted by $\mathcal{Y} = \{\mathbf{Y}_b\}_{b \in \mathcal{B}} = \mathcal{W}(\mathbf{I}_{\mathrm{gt}})$. The target LUT is obtained through constrained inversion:

$$\mathbf{\Theta}^r = \arg\min_{\mathbf{\Theta}} \sum_{b \in \mathcal{B}} \left[\frac{\|\widehat{\mathbf{X}}_b - \mathbf{Y}_b\|_F^2}{\|\mathbf{Y}_b\|_F^2 + \varepsilon_{\mathrm{f}}} + \lambda_{\mathrm{a}} \|\mathbf{\Lambda}_b^{1/2} \odot (\mathbf{\Theta}_b - \mathbf{\Theta}_b^{(0)})\|_F^2 + \lambda_{\mathrm{s}} \mathcal{R}_{\mathrm{sm}}\Big(\mathbf{\Theta}_b - \mathbf{\Theta}_b^{(0)}\Big) \right]. \tag{3}$$

The first term aligns the restored subbands with their normal-light counterparts. $\mathbf{\Lambda}_b$ places stronger constraints on sparsely queried nodes, keeping them close to the coarse LUT, while $\mathcal{R}_{\mathrm{sm}}$ encourages smooth residual variations between neighboring nodes. During inversion, the nodes and coarse LUTs are fixed, and only the target LUT values $\mathbf{\Theta}^r$ are optimized.

Given the target LUT, diffusion teacher learns its residual relative to the coarse LUT. For each $b \in \mathcal{B}$,

$$\widehat{\boldsymbol{\epsilon}}_{b,\tau} = D_\phi\Big(\sqrt{\bar{\alpha}_\tau}\, \Delta\mathbf{\Theta}_b^r + \sqrt{1-\bar{\alpha}_\tau}\, \boldsymbol{\epsilon}_b, \tau \mid \mathbf{I}_{\mathrm{in}}, \mathcal{N}_b(\mathbf{I}_{\mathrm{in}}), \mathbf{\Theta}_b^{(0)}, b\Big), \quad \boldsymbol{\epsilon}_b \sim \mathcal{N}(\mathbf{0}, \mathbf{I}). \tag{4}$$

Here, $\Delta\mathbf{\Theta}_b^r = \mathbf{\Theta}_b^r - \mathbf{\Theta}_b^{(0)}$ and $\bar{\alpha}_\tau = \prod_{s=1}^{\tau}(1-\beta_s)$. The denoiser D_ϕ predicts the injected noise from the noisy residual, input image, non-uniform nodes, coarse LUT, and subband identity. Its parameters are shared across the four subbands. Diffusion is applied only to the LUT values, while the node positions remain fixed. To generate the teacher LUT, DDIM starts from a fixed Gaussian state $\mathbf{Z}^\star_{b,\tau_{\max}}$ and progressively recovers the correction residual, then, the recovered residual is added to the coarse LUT to obtain the teacher LUT $\mathbf{\Theta}_b^t$.

$$\widehat{\Delta\mathbf{\Theta}}_b^t = \mathrm{DDIM}_{\phi,\eta=0}\left(\mathbf{Z}^\star_{b,\tau_{\max}} \mid \mathbf{I}_{\mathrm{in}}, \mathcal{N}_b(\mathbf{I}_{\mathrm{in}}), \mathbf{\Theta}_b^{(0)}, b\right), \qquad \mathbf{\Theta}_b^t = \mathbf{\Theta}_b^{(0)} + \widehat{\Delta\mathbf{\Theta}}_b^t. \tag{5}$$

The student LUTs are CP-decomposed and evaluated through pentalinear interpolation (see Appendix for details) to obtain the enhanced subbands, which are then reconstructed by inverse SWT.

2.4 CONTINUOUS LUT FUNCTION DISTILLATION

Motivation. Image-adaptive LUTs must be generated for each input, while the diffusion teacher requires multi-step sampling and thus increases inference cost. We therefore distill the continuous mapping of the teacher LUT $\mathbf{\Theta}^t$ into a student model, allowing LUT prediction without running the diffusion process. Rather than directly matching LUT values, we align teacher and student interpolation outputs at the same query locations, directly constraining the resulting mappings.

Details. We fix the teacher LUTs and use their interpolated outputs to supervise the student model $\mathcal{S}_\psi$, where ψ denotes the trainable student parameters. A shared image encoder extracts input features, and four subband branches combine these features with the corresponding coarse LUTs and node coordinates. The student predicts only the LUT values and retains the non-uniform nodes generated by $\mathcal{G}_0$. Distillation queries are drawn from both the input image and LUT grid cells. Image queries are sampled from $\mathbf{q}_b(\mathbf{p})$ of the current input, forming π_b^{img}, while grid queries are continuously sampled within LUT cells, forming π_b^{grid}. The two are combined as

$$\pi_b = \rho\, \pi_b^{\text{img}} + (1-\rho)\, \pi_b^{\text{grid}}, \qquad 0 < \rho < 1. \tag{6}$$

Here, ρ and $1-\rho$ are the probabilities of selecting image and grid queries, respectively. Image queries follow the actual query distribution of the current input, while grid queries cover continuous locations not sampled by the image.

Let $\mathbf{\Theta}_b^t$ denote the teacher LUT of subband b, and let $\mathcal{F}_b(\mathbf{q}; \mathbf{\Theta}_b)$ denote the three-channel output obtained by non-uniform pentalinear interpolation at a five-dimensional query $\mathbf{q} \in \mathbb{R}^5$. For each query $\mathbf{q}$ sampled from π_b, we evaluate the teacher and student at the same location and define the continuous function distillation loss as

$$\mathcal{L}_{\text{func}} = \frac{1}{|\mathcal{B}|} \sum_{b \in \mathcal{B}} \mathbb{E}_{\mathbf{q} \sim \pi_b} \left[\left\| \mathcal{F}_b(\mathbf{q}; \mathbf{\Theta}_b^s) - \text{sg}\big[\mathcal{F}_b(\mathbf{q}; \mathbf{\Theta}_b^t)\big] \right\|_2^2 \right]. \tag{7}$$

Here, $|\mathcal{B}| = 4$ denotes the four subbands, and the expectation is estimated by averaging over sampled queries during training. The term $\|\cdot\|_2^2$ measures the squared error over the three output channels, while sg stops gradients through the teacher output, treating it as a fixed distillation target. The teacher LUTs and node positions remain fixed, and only the student parameters ψ are updated.

After distillation, let $\psi^\star$ denote the trained student parameters. At inference, the student predicts the complete LUTs for all four subbands in a single forward pass without running the diffusion teacher:

$$\mathbf{\Theta}^s = \mathcal{S}_{\psi^\star}(\mathbf{I}_{\text{in}}, \mathcal{G}_0(\mathbf{I}_{\text{in}})) = \{\mathbf{\Theta}_b^s\}_{b \in \mathcal{B}}. \tag{8}$$

2.5 TRAINING OBJECTIVES

Following Cai et al. (2026), we adopt an image reconstruction MSE loss $\mathcal{L}_{\text{mse}}$ and adapt its smoothness and monotonicity regularizers. The smoothness loss $\mathcal{L}_{\text{sm}}$ accounts for non-uniform node spacing, while monotonicity loss $\mathcal{L}_{\text{mon}}$ is restricted to matching color axes in the low-frequency LUT. The coarse LUT objective is

$$\mathcal{L}_{\text{coarse}} = \mathcal{L}_{\text{mse}}(\mathbf{\Theta}^{(0)}) + \lambda_{\text{sm}} \mathcal{L}_{\text{sm}}(\mathbf{\Theta}^{(0)}) + \lambda_{\text{mon}} \mathcal{L}_{\text{mon}}(\mathbf{\Theta}^{(0)}). \tag{9}$$

We freeze $\mathcal{G}_0$ and obtain the target LUTs using Eq. 3. Following Shin et al. (2025), we optimize the teacher parameters ϕ with a noise-prediction loss:

$$\mathcal{L}_{\text{teacher}} = \mathcal{L}_{\text{diff}} = \frac{1}{|\mathcal{B}|} \sum_{b \in \mathcal{B}} \mathbb{E}\left[\text{MSE}\left(\widehat{\boldsymbol{\epsilon}}_{b,\tau}, \boldsymbol{\epsilon}_b\right)\right]. \tag{10}$$

The expectation is taken over training samples, diffusion timesteps, and Gaussian noise. We then fix the teacher LUTs and adopt the output-distillation principle of Yi et al. (2025). Using the continuous-query distillation loss in Eq. 7, we optimize only the student parameters ψ:

$$\mathcal{L}_{\text{student}} = \mathcal{L}_{\text{mse}}(\mathbf{\Theta}^s) + \lambda_{\text{func}}\mathcal{L}_{\text{func}} + \lambda_{\text{sm}}\mathcal{L}_{\text{sm}}(\mathbf{\Theta}^s) + \lambda_{\text{mon}}\mathcal{L}_{\text{mon}}(\mathbf{\Theta}^s). \tag{11}$$

3 EXPERIMENTS

3.1 EXPERIMENTAL SETTINGS

Implementation Details. All experiments use PyTorch on an NVIDIA RTX 4060 Ti GPU. We train to convergence over 1×10^5 iterations using Adam with $\beta_1 = 0.9$, $\beta_2 = 0.999$, a batch size of 8, and an initial learning rate of 1×10^{-4}.

Datasets and Metrics. We evaluate our method on paired and unpaired benchmarks: FiveK, LOLv1, LOLv2-real, and LOLv2-synthetic cover expert-retouched, real, and synthetic paired low-light images. DICM, MEF, LIME, NPE, and VV provide real unpaired indoor/outdoor scenes, while SICE provides paired multi-exposure images for ablation Lin et al. (2026); Zhu et al. (2025); Liu et al. (2026). PSNR, SSIM, and ΔE assess pixel accuracy, structure, and color on paired data; NIQE, BRISQUE (BRI), and PI assess no-reference perceptual quality Jung et al. (2025); Huaqiu & Wang (2025); Zhang et al. (2026).

Comparison Methods. To evaluate our method against recent state-of-the-art (SOTA) methods, we compare it with NiLUT Conde et al. (2024), AttLUT Fu et al. (2024), SvdLUT Kim et al. (2025), Fastllve Li et al. (2023), WLUT He et al. (2026), DPLUT Lin et al. (2025), FLUT Hu et al. (2025), PanLUT Cai et al. (2026), DnLUT Yang et al. (2025), Darkir Feijoo et al. (2025), DpSPG Lu et al. (2025), LatentD Hu et al. (2026), TSCnet Zhang et al. (2025), Adair Cui et al. (2025), CIDNet Yan et al. (2024), SG-LLIE Dong et al. (2025), LYTNET Brateanu et al. (2025), FFTF Pei et al. (2025), CFWD Xue et al., and WV-LUT Li et al. (2025).

Quantitative Comparisons. On FiveK (Table 1), our method achieves the best PSNR, SSIM, ΔE, and LPIPS at both full resolution and 480p. At full resolution, it exceeds NiLUT by $0.90\,\text{dB}$ in PSNR, improves SSIM by 0.008 and reduces ΔE by 0.65 over SvdLUT, and lowers LPIPS by 0.02 relative to PanLUT. At 480p, it improves PSNR and SSIM over NiLUT by $0.57\,\text{dB}$ and 0.008, while reducing ΔE and LPIPS by 0.53 and 0.02 over SvdLUT and PanLUT, respectively. On the LOL benchmarks (Table 2), our method achieves the best PSNR, SSIM, and LPIPS on all three datasets. Its PSNR exceeds the second-best results by $1.90\,\text{dB}$, $1.68\,\text{dB}$, and $1.26\,\text{dB}$ on LOLv1, LOLv2-real, and LOLv2-synthetic, respectively; the corresponding SSIM gains are 0.012, 0.011, and 0.005, while LPIPS is reduced by 0.02, 0.03, and 0.03. Across the five no-reference benchmarks (Table 3), our method ranks first in 13 of the 15 comparisons and achieves the best results on all three metrics for DICM, MEF, and LIME. In particular, BRISQUE is reduced by 0.245 over PanLUT on DICM, while NIQE is reduced by 0.264 and 0.122 on LIME and NPE, respectively. The only second-best results are BRISQUE on NPE and VV. These results are consistent with the intended roles of frequency-specific correction and continuous mapping supervision; individual contributions are examined in the ablation study.

Qualitative Comparisons. Figures 3–5 compare different methods in terms of visual quality, reconstruction error, color distribution, and residual noise. In Figures 3 and 4, the pseudo-color maps encode increasing deviation from the GT from blue to red, while bright regions in the SSIM difference maps indicate larger local structural errors. For the noise statistics, lower medians and tighter distributions indicate less residual noise. Figure 5 further uses RGB histograms and HSV scatter plots to examine intensity and color distributions. On FiveK (Figure 3), our method recovers natural exposure and color across indoor, nighttime street, and outdoor scenes, while exhibiting lower reconstruction and structural errors in the pseudo-color and SSIM difference maps. On LOL (Figure 4), it improves dark-region visibility while preserving color and structure, notably reducing indoor color shifts and retaining warm illumination and edge contours in the architectural scene. On the unpaired datasets (Figure 5), our method also recovers shadow details while preserving facial and background textures, with fewer overexposure and color-shift artifacts. The RGB, HSV, and noise statistics further show more reasonable intensity, color, and residual-noise distributions.

Table 1: Quantitative comparisons on FiveK test datasets. The best results are highlighted in bold.

Methods	#Para(K)	Runtime(ms)	Full Resolution(4k)				480p			
			PSNR↑	SSIM↑	ΔE↓	LPIPS↓	PSNR↑	SSIM↑	ΔE↓	LPIPS↓
NiLUT	133	8.3	26.35	0.925	7.32	0.22	26.87	0.944	7.28	0.18
AttLUT	593	6.8	25.56	0.926	7.53	0.20	25.61	0.932	7.48	0.24
SvdLUT	160.5	4.2	25.69	0.938	7.27	0.18	25.76	0.931	7.26	0.18
Fastllve	11000	13.5	26.22	0.861	7.44	0.23	26.54	0.878	7.39	0.21
WLUT	10742	10.7	25.75	0.879	7.47	0.15	25.96	0.885	7.43	0.10
DPLUT	**78**	6.9	20.56	0.738	7.97	0.17	21.33	0.822	7.92	0.14
PanLUT	662.6	3.8	26.24	0.911	7.58	0.09	26.52	0.938	7.36	0.07
Ours	795	**3.5**	**27.25**	**0.946**	**6.62**	**0.07**	**27.44**	**0.952**	**6.73**	**0.05**

Table 2: Quantitative comparisons on LOL test datasets.

Methods	LOLv1			LOLv2-real			LOLv2-syn		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
NiLUT	23.23	0.794	0.21	25.62	0.836	0.23	25.44	0.928	0.21
WV-LUT	21.15	0.854	0.24	19.36	0.705	0.25	24.81	0.871	0.23
SvdLUT	24.25	0.878	0.16	24.68	0.844	0.18	21.41	0.848	0.18
Darkir	20.63	0.823	0.20	19.94	0.827	0.25	23.71	0.914	0.22
Adair	24.36	0.827	0.16	24.45	0.895	0.18	25.28	0.877	0.23
WLUT	25.32	0.922	0.15	25.49	0.889	0.18	24.68	0.919	0.12
CIDNet	23.81	0.857	0.21	24.11	0.868	0.18	25.13	0.939	0.15
LatentD	24.22	0.866	0.09	22.65	0.821	0.12	23.21	0.842	0.08
PanLUT	25.29	0.926	0.08	24.52	0.922	0.15	24.55	0.875	0.10
FLUT	23.73	0.854	0.08	24.26	0.863	0.08	25.15	0.920	0.09
DnLUT	24.88	0.886	0.11	25.55	0.875	0.09	25.68	0.925	0.14
Ours	**27.22**	**0.938**	**0.06**	**27.30**	**0.933**	**0.05**	**26.94**	**0.944**	**0.05**

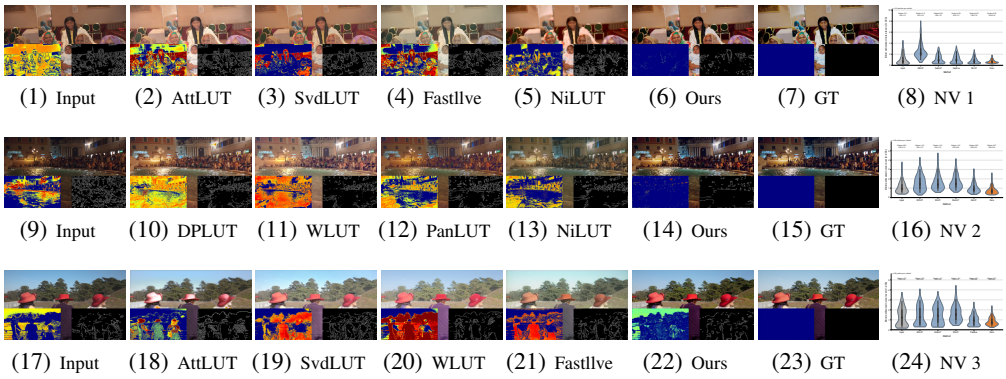


Figure 3: Visual comparison with recent SOTA methods on FiveK datasets. Best viewed zoomed in.

3.2 Ablation Study

Loss Functions. Table 4 reports the loss ablation results. Using only the reconstruction loss $\mathcal{L}_{\text{mse}}$ (S1) reduces PSNR and SSIM by $0.97\,\text{dB}$ and 0.019, respectively, and increases ΔE by 0.68 relative to the full objective (S8), indicating that image reconstruction alone does not sufficiently constrain the continuous LUT mapping. Adding the smoothness regularizer $\mathcal{L}_{\text{sm}}$ (S2) improves PSNR by $0.12\,\text{dB}$. Removing $\mathcal{L}_{\text{mse}}$ (S4) or the function distillation loss $\mathcal{L}_{\text{func}}$ (S3) decreases PSNR by $0.60\,\text{dB}$ and $0.79\,\text{dB}$, respectively. Replacing $\mathcal{L}_{\text{func}}$ with the node-value matching loss $\mathcal{L}_{\text{node}}$ (S5) reduces PSNR and SSIM by $0.40\,\text{dB}$ and 0.007, showing that matching discrete LUT values does not ensure consistent interpolation between nodes. Removing $\mathcal{L}_{\text{sm}}$ (S6) decreases PSNR by $0.31\,\text{dB}$, while removing the monotonicity regularizer $\mathcal{L}_{\text{mon}}$ (S7) decreases SSIM by 0.009 and increases ΔE by 0.34. The coarse LUT exhibits the same trend. Using only $\mathcal{L}_{\text{mse}}$ (C1) reduces the final-student PSNR and SSIM by $0.52\,\text{dB}$ and 0.008 relative to the full objective (C4). Overall, $\mathcal{L}_{\text{mse}}$ provides reconstruction supervision, $\mathcal{L}_{\text{func}}$ constrains the continuous mapping, and $\mathcal{L}_{\text{sm}}$ and $\mathcal{L}_{\text{mon}}$ stabilize

Table 3: Quantitative comparisons on DICM,MEF,LIME,NPE and VV test datasets.

Methods	DICM			MEF			LIME			NPE			VV		
	NIQE↓	BRI↓	PI↓	NIQE↓	BRI↓	PI↓	NIQE↓	BRI↓	PI↓	NIQE↓	BRI↓	PI↓	NIQE↓	BRI↓	PI↓
NiLUT	3.878	19.436	3.256	3.825	26.248	3.477	3.626	21.232	3.358	3.611	20.55	2.883	4.565	19.534	2.828
DPLUT	3.832	19.129	3.256	3.556	26.799	3.270	4.249	22.664	3.381	3.775	21.157	2.953	3.689	**14.153**	2.749
DnLUT	3.824	19.295	3.663	3.626	20.277	3.325	3.881	21.687	3.331	3.858	18.444	2.788	3.639	19.854	3.503
FFTF	4.519	27.922	3.700	3.608	26.716	3.286	4.463	25.170	3.376	4.124	28.887	3.534	5.312	22.800	3.648
PanLUT	3.545	18.878	3.344	3.658	24.404	3.367	4.133	20.451	3.441	3.546	22.235	3.425	3.339	22.24	3.374
CFWD	4.774	24.544	3.565	4.231	34.720	3.665	4.694	29.022	3.713	4.028	26.094	3.153	3.851	22.457	2.891
DpSPG	4.422	20.662	3.826	4.745	28.577	4.233	4.575	27.425	4.874	4.365	26.542	3.768	3.835	24.411	2.770
TSCnet	4.277	22.369	3.538	3.969	22.775	3.216	4.110	21.528	3.562	4.211	22.563	3.525	3.824	22.663	3.467
LYTNET	4.418	20.598	3.723	4.406	27.984	3.974	4.742	26.972	4.452	4.724	27.844	3.866	3.985	23.625	3.384
SG-LLIE	4.571	20.979	4.540	4.475	28.779	4.830	4.550	26.744	4.669	4.645	27.656	4.533	4.553	22.877	2.967
Ours	**3.535**	**18.633**	**3.119**	**3.425**	**19.677**	**3.155**	**3.362**	**20.226**	**2.884**	**3.424**	**19.655**	**2.538**	**3.311**	19.212	**2.656**

Table 4: Loss Functions ablation.

Stage	ID	Setting	$\mathcal{L}_{\text{mse}}$	$\mathcal{L}_{\text{func}}$	$\mathcal{L}_{\text{node}}$	$\mathcal{L}_{\text{sm}}$	$\mathcal{L}_{\text{mon}}$	PSNR↑	SSIM↑	$\Delta E\downarrow$
Student LUT Θ^s	S1	MSE only	✓	–	–	–	–	26.46	0.916	7.43
	S2	MSE + smoothness	✓	–	–	✓	–	26.58	0.920	7.31
	S3	W/O function distillation	✓	–	–	✓	✓	26.64	0.922	7.22
	S4	W/O MSE	–	✓	–	✓	✓	26.83	0.925	7.15
	S5	Node-value matching	✓	–	✓	✓	✓	27.03	0.928	7.03
	S6	W/O smoothness	✓	✓	–	–	✓	27.12	0.931	6.98
	S7	W/O monotonicity	✓	✓	–	✓	–	27.24	0.926	7.09
	S8	Full student objective	✓	✓	–	✓	✓	27.43	0.935	6.75
Coarse LUT $\Theta^{(0)}$	C1	MSE only	✓	–	–	–	–	26.91	0.927	7.13
	C2	W/O smoothness	✓	–	–	–	✓	27.08	0.929	6.94
	C3	W/O monotonicity	✓	–	–	✓	–	27.18	0.930	6.99
	C4	Full coarse objective	✓	–	–	✓	✓	27.43	0.935	6.75

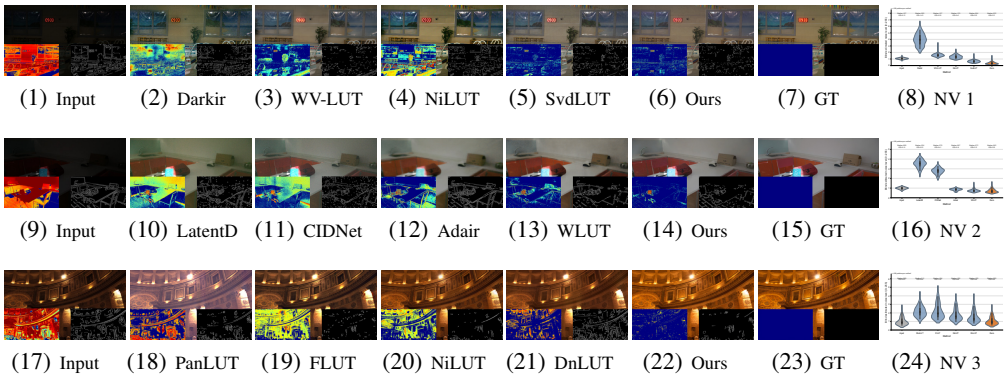


Figure 4: Visual comparison with recent SOTA methods on LOL datasets. Best viewed zoomed in.

local LUT variations and preserve color ordering, respectively. Together, these terms improve reconstruction accuracy and color fidelity.

Network Architecture. Table 5 reports the architecture ablation results. HFC, LC, TL, HFR, NU, GQ, RD, SP, and CP denote high-frequency correction, local illumination and contrast conditioning, teacher LUTs, high-frequency residual prediction, non-uniform nodes, grid queries, residual diffusion, student prediction, and CP compression, respectively. A1–A9 remove or replace these key designs, while A10 is the full configuration. High-frequency correction and local conditioning have the largest effects. Removing HFC (A1) reduces PSNR and SSIM by $0.89\,\mathrm{dB}$ and 0.021, respectively, while replacing local conditions with image-wise means (A2) lowers PSNR by $0.64\,\mathrm{dB}$ and increases ΔE by 0.46. Variants A3–A7, which modify teacher supervision, high-frequency residual

Table 5: Network Architecture ablation.

ID	HFC	LC	TL	HFR	NU	GQ	RD	SP	CP	PSNR↑	SSIM↑	$\Delta E\downarrow$
A1	–	✓	✓	N/A	✓	✓	✓	✓	✓	26.54	0.914	7.13
A2	✓	–	✓	✓	✓	✓	✓	✓	✓	26.79	0.925	7.21
A3	✓	✓	–	✓	✓	✓	N/A	✓	✓	27.04	0.928	7.05
A4	✓	✓	✓	–	✓	✓	✓	✓	✓	27.06	0.928	7.04
A5	✓	✓	✓	✓	–	✓	✓	✓	✓	27.11	0.930	7.01
A6	✓	✓	✓	✓	✓	–	✓	✓	✓	27.14	0.930	6.98
A7	✓	✓	✓	✓	✓	✓	–	✓	✓	27.17	0.931	6.96
A8	✓	✓	✓	✓	✓	N/A	✓	–	–	27.32	0.912	6.65
A9	✓	✓	✓	✓	✓	✓	✓	✓	–	27.43	0.935	6.75
A10	✓	✓	✓	✓	✓	✓	✓	✓	✓	27.43	0.935	6.75

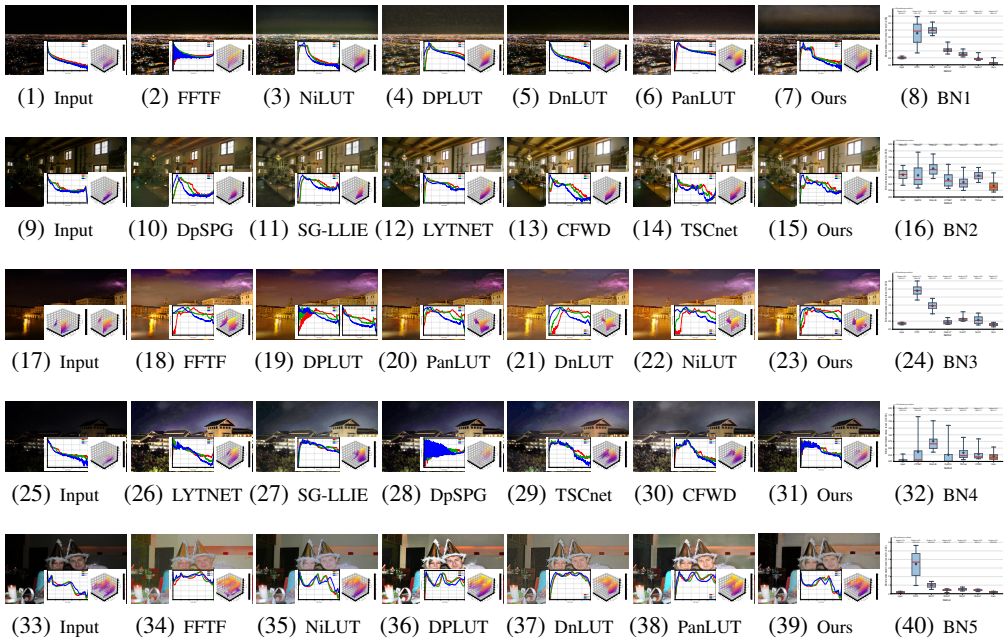


Figure 5: Visual comparison with SOTA methods on DICM,MEF,LIME,NPE and VV datasets.

prediction, non-uniform nodes, grid queries, or residual diffusion, reduce PSNR by 0.26–0.39 dB. Among them, A3 and A6 incur drops of 0.39 dB and 0.29 dB, confirming the value of teacher refinement and grid queries for LUT learning. Compared with directly applying the teacher LUTs in A8, the uncompressed student in A9 improves PSNR and SSIM by 0.11 dB and 0.023, respectively, at the cost of a 0.10 increase in ΔE. A10 matches A9 on all three metrics at the reported precision, indicating no observable performance loss from CP compression.

4 Conclusion

This work explores a low-light enhancement approach that transfers the correction capability of diffusion models into explicit LUT mappings. Frequency-domain decomposition is used to handle illumination, color, texture, and noise separately, while a diffusion teacher supplements mapping relations that are difficult to learn from limited observations. The resulting restoration capability is distilled into student LUTs that can be executed directly. Extensive experiments demonstrate stable performance gains. However, the current method still relies on paired data to construct target LUTs. Future work will investigate weakly supervised or unsupervised target LUT construction and extend the method to low-light video enhancement.

AI USE STATEMENT

Generative AI was not used in the conception of this study, method design, experimental analysis, or code implementation. Its use was limited to language assistance, including grammar correction, sentence refinement, and academic polishing. All AI-assisted revisions were reviewed and confirmed by the authors. The research ideas, technical methods, experimental results, analyses, and conclusions presented in this work were independently developed and verified by the authors. The authors take full responsibility for the scientific validity, accuracy, and final content of the manuscript, including any text refined with the assistance of generative AI.

REFERENCES

Alexandru Brateanu, Raul Balmez, Adrian Avram, Ciprian Orhei, and Cosmin Ancuti. Lyt-net: Lightweight yuv transformer-based network for low-light image enhancement. *IEEE Signal Processing Letters*, 32:2065–2069, 2025.

Zhongnan Cai, Yingying Wang, Hui Zheng, Panwang Pan, ZiXu Lin, Ge Meng, Chenxin Li, Chunming He, Jiaxin Xie, Yunlong Lin, et al. Pan-lut: Efficient pan-sharpening via learnable look-up tables. *Advances in Neural Information Processing Systems*, 38:126459–126483, 2026.

Marcos V Conde, Javier Vazquez-Corral, Michael S Brown, and Radu Timofte. Nilut: Conditional neural implicit 3d lookup tables for image enhancement. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 1371–1379, 2024.

Yuning Cui, Syed Waqas Zamir, Salman Khan, Alois Knoll, Mubarak Shah, and Fahad Khan. Adair: Adaptive all-in-one image restoration via frequency mining and modulation. In *International Conference on Learning Representations*, volume 2025, pp. 101306–101327, 2025.

Wei Dong, Yan Min, Han Zhou, and Jun Chen. Towards scale-aware low-light enhancement via structure-guided transformer design. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1454–1461. IEEE, 2025.

Daniel Feijoo, Juan C Benito, Alvaro Garcia, and Marcos V Conde. Darkir: Robust low-light image restoration. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10879–10889. IEEE, 2025.

Kang Fu, Yicong Peng, Zicheng Zhang, Qihang Xu, Xiaohong Liu, Jia Wang, and Guangtao Zhai. Attentionlut: Attention fusion-based canonical polyadic lut for real-time image enhancement. In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3255–3259. IEEE, 2024.

Jinhong He, Minglong Xue, Wenhai Wang, and Mingliang Zhou. Optimizing a 4d lookup table for low-light video enhancement via wavelet priori. *IEEE Transactions on Multimedia*, 2026.

Liubing Hu, Chen Wu, Anrui Wang, Dianjie Lu, Guijuan Zhang, and Zhuoran Zheng. Flowlut: Efficient image enhancement via differentiable luts and iterative flow matching. *arXiv preprint arXiv:2509.23608*, 2025.

Renzhi Hu, Ting Luo, Gangyi Jiang, Leiming Liu, Yeyao Chen, Haiyong Xu, and Zhouyan He. Latentdark: Reflectance guided latent diffusion model for low-light image enhancement. *Signal Processing*, 238:110125, 2026.

Li Huaqiu and Haoqian Wang. Interpretable unsupervised joint denoising and enhancement for real-world low-light scenarios. In *International Conference on Learning Representations*, volume 2025, pp. 39525–39537, 2025.

Donggoo Jung, Daehyun Kim, and Tae Hyun Kim. Continuous exposure learning for low-light image enhancement using neural odes. In *International Conference on Learning Representations*, volume 2025, pp. 37376–37401, 2025.

Wontae Kim, Keuntek Lee, and Nam Ik Cho. Lightweight and fast real-time image enhancement via decomposition of the spatial-aware lookup tables. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 11895–11905. IEEE, 2025.

Canlin Li, Haowen Su, Xin Tan, Xiangfei Zhang, and Lizhuang Ma. Wv-lut: Wide vision lookup tables for real-time low-light image enhancement. *IEEE Transactions on Multimedia*, 27:4441–4453, 2025.

Wenhao Li, Guangyang Wu, Wenyi Wang, Peiran Ren, and Xiaohong Liu. Fastllve: Real-time low-light video enhancement with intensity-aware look-up table. In *Proceedings of the 31st ACM International Conference on Multimedia*, pp. 8134–8144, 2023.

Jingbo Lin, Zhilu Zhang, Wenbo Li, Renjing Pei, Hang Xu, Hongzhi Zhang, and Wangmeng Zuo. Unirestorer: Universal image restoration via adaptively estimating image degradation at proper granularity. In *International Conference on Learning Representations*, volume 2026, pp. 2369–2409, 2026.

Yunlong Lin, Zhenqi Fu, Kairun Wen, Tian Ye, Sixiang Chen, Ge Meng, Yingying Wang, Chui Kong, Yue Huang, Xiaotong Tu, et al. Dplut: Unsupervised low-light image enhancement with lookup tables and diffusion priors. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 5316–5324, 2025.

Xiaoyang Liu, Zhengyan Zhou, Zihang Xu, Jiezhang Cao, Zheng Chen, and Yulun Zhang. Fidediff: Efficient diffusion model for high-fidelity image motion deblurring. In *International Conference on Learning Representations*, volume 2026, pp. 57231–57243, 2026.

Jialang Lu, Huayu Zhao, Huiyu Zhai, Xingxing Yang, and Shini Han. Deepspg: Exploring deep semantic prior guidance for low-light image enhancement with multimodal learning. In *Proceedings of the 2025 International Conference on Multimedia Retrieval*, pp. 935–943, 2025.

Xiao Pei, Yongdong Huang, Weijian Su, Fengjuan Zhu, and Qiang Liu. Fftformer: A spatial-frequency noise aware cnn-transformer for low light image enhancement. *Knowledge-Based Systems*, 314:113055, 2025.

Seunghyun Shin, Dongmin Shin, Jisu Shin, Hae-Gon Jeon, and Joon-Young Lee. Video color grading via look-up table generation. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 19141–19152. IEEE, 2025.

M Xue, J He, W Wang, and M Zhou. Low-light image enhancement via clip-fourier guided wavelet diffusion. arxiv 2024. *arXiv preprint arXiv:2401.03788*.

Qingsen Yan, Yixu Feng, Cheng Zhang, Pei Wang, Peng Wu, Wei Dong, Jinqiu Sun, and Yanning Zhang. You only need one color space: An efficient network for low-light image enhancement. *arXiv preprint arXiv:2402.05809*, 2024.

Sidi Yang, Binxiao Huang, Yulun Zhang, Dahai Yu, Yujiu Yang, and Ngai Wong. Dnlut: Ultra-efficient color image denoising via channel-aware lookup tables. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 7582–7591, 2025.

Xunpeng Yi, Yibing Zhang, Xinyu Xiang, Qinglong Yan, Han Xu, and Jiayi Ma. Lut-fuse: Towards extremely fast infrared and visible image fusion via distillation to learnable look-up tables. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 14559–14568. IEEE, 2025.

Deheng Zhang, Yuqian Fu, Runyi Yang, Yang Miao, Tianwen Qian, Xu Zheng, Guolei Sun, Ajad Chhatkuli, Xuanjing Huang, Yu-Gang Jiang, et al. Egonight: Towards egocentric vision understanding at night with a challenging benchmark. In *International Conference on Learning Representations*, volume 2026, pp. 887–901, 2026.

Miao Zhang, Jun Yin, Pengyu Zeng, Yiqing Shen, Shuai Lu, and Xueqian Wang. Tscnet: A text-driven semantic-level controllable framework for customized low-light image enhancement. *Neurocomputing*, 625:129509, 2025.

Yixuan Zhu, Haolin Wang, Ao Li, Wenliang Zhao, Yansong Tang, Jingxuan Niu, Lei Chen, Jie Zhou, and Jiwen Lu. Instarevive: One-step image enhancement via dynamic score matching. In *International Conference on Learning Representations*, volume 2025, pp. 26965–26988, 2025.

TEACHING LOOKUP TABLES TO SEE IN THE DARK: LEARNING COMPLETE MAPPINGS FROM A DIFFUSION TEACHER

Anonymous authors
Paper under double-blind review

1 COMPACT LUT REPARAMETERIZATION

Motivation. Distillation removes the iterative sampling of the diffusion teacher, but the student still generates four complete 5D LUTs, resulting in considerable storage overhead. Non-uniform sampling improves node allocation without reducing the number of nodes. To keep compression from affecting mapping learning, we separate LUT learning from compression: the student first predicts complete LUTs, which are then reparameterized by CP decomposition. This preserves the original query grids while reducing the number of table parameters.

Details. We apply CP decomposition to the three-channel LUT values of each subband independently. Let $K_{b,d}$ denote the number of nodes along query axis d of subband b. The corresponding CP factors are represented as $\mathbf{A}_{b,d} \in \mathbb{R}^{K_{b,d}\times r\times 3}$, where r is the CP rank for each output channel. The approximated LUT values are defined as

$$\widetilde{\mathbf{\Theta}}_b[k_1,\dots,k_5,c] = \sum_{\ell=1}^{r}\prod_{d=1}^{5}\mathbf{A}_{b,d}[k_d,\ell,c]. \tag{1}$$

Here, $k_d \in \{1,\dots,K_{b,d}\}$ indexes the node along query axis d, $c \in \{1,2,3\}$ indexes the output channel, and ℓ indexes the rank-one component. The CP factors are obtained by fitting the complete student LUT:

$$\min_{\{\mathbf{A}_{b,d}\}_{d=1}^{5}} \left\|\mathbf{\Theta}_b^s - \widetilde{\mathbf{\Theta}}_b\right\|_F^2, \qquad b\in\mathcal{B}. \tag{2}$$

We approximately solve this problem using a finite number of alternating least-squares iterations. For each subband, the complete LUT contains $3\prod_{d=1}^{5}K_{b,d}$ scalar parameters, whereas the CP representation requires only $3r\sum_{d=1}^{5}K_{b,d}$. The node coordinates remain unchanged after factorization. Let $\mathbf{\Gamma}_b$ denote the compact LUT composed of the original non-uniform nodes and the fitted CP factors. The resulting representation is

$$\mathbf{\Gamma}_b = \mathcal{C}_{\mathrm{CP},r}(\mathbf{\Theta}_b^s) = \left(\mathcal{N}_b(\mathbf{I}_{\mathrm{in}}), \{\mathbf{A}_{b,d}\}_{d=1}^{5}\right), \qquad b\in\mathcal{B}. \tag{3}$$

The pseudocode for Compact LUT Reparameterization is presented in Algorithm 1

2 NON-UNIFORM PENTALINEAR INTERPOLATION

After LUT generation and CP reparameterization, we correct each frequency subband using non-uniform pentalinear interpolation. For the d-th component $q_{b,d}$ of $\mathbf{q}_b$, binary search locates the adjacent nodes $\nu_{b,d}^{(k_d)}$ and $\nu_{b,d}^{(k_d+1)}$ in $\mathcal{N}_{b,d}(\mathbf{I}_{\mathrm{in}})$, where $k_d \in \{1,\dots,K_{b,d}-1\}$ indexes the left node. The normalized position is

$$t_{b,d} = \frac{q_{b,d}-\nu_{b,d}^{(k_d)}}{\nu_{b,d}^{(k_d+1)}-\nu_{b,d}^{(k_d)}}, \qquad d=1,\dots,5. \tag{4}$$

Queries outside the node range are clipped accordingly, and queries at the maximum node use the final interval. Since the nodes are strictly increasing, the denominator is positive and $t_{b,d}\in[0,1]$.

Algorithm 1 Compact LUT Reparameterization

Input: Student LUT values $\{\mathbf{\Theta}_b^s\}_{b\in\mathcal{B}}$; original non-uniform nodes $\{\mathcal{N}_b(\mathbf{I}_{\text{in}})\}_{b\in\mathcal{B}}$; number of rank-one terms per output channel r.

Output: Compact LUTs $\{\mathbf{\Gamma}_b\}_{b\in\mathcal{B}}$.

for each $b \in \mathcal{B}$ **do**
  Retain the original nodes $\mathcal{N}_b(\mathbf{I}_{\text{in}})$.
  Initialize $\mathbf{A}_{b,d} \in \mathbb{R}^{K_{b,d}\times r\times 3}$, $d = 1, \ldots, 5$.
  Represent the approximate table values using Eq. 1:

$$\widetilde{\mathbf{\Theta}}_b[k_1, \ldots, k_5, c] = \sum_{\ell=1}^{r}\prod_{d=1}^{5}\mathbf{A}_{b,d}[k_d, \ell, c].$$

  Fit the student table using Eq. 2:

$$\min_{\{\mathbf{A}_{b,d}\}_{d=1}^{5}} \left\|\mathbf{\Theta}_b^s - \widetilde{\mathbf{\Theta}}_b\right\|_F^2.$$

  for $c = 1, 2, 3$ **do**
    repeat a finite number of ALS iterations
      for $d = 1, \ldots, 5$ **do**
        Fix the factors on all other axes.
        Update the channel-c factor in $\mathbf{A}_{b,d}$
        by least squares for Eq. 2.
      end for
    end repeat
  end for
  Form the compact LUT using Eq. 3:

$$\mathbf{\Gamma}_b = \mathcal{C}_{\text{CP},r}(\mathbf{\Theta}_b^s) = \left(\mathcal{N}_b(\mathbf{I}_{\text{in}}), \{\mathbf{A}_{b,d}\}_{d=1}^{5}\right).$$

end for
return $\{\mathbf{\Gamma}_b\}_{b\in\mathcal{B}}$.

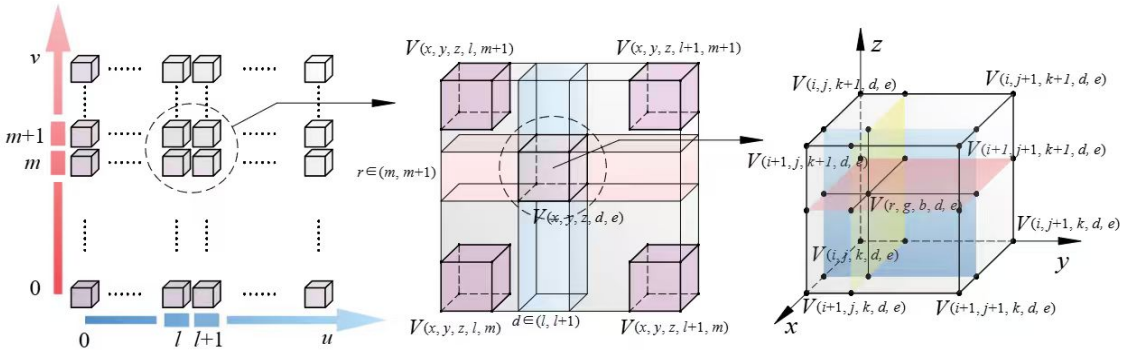


Figure 1: Illustration of Pentalinear Interpolation.

The adjacent nodes along the five query axes define a five-dimensional cell with 32 corners. Let $\boldsymbol{a} = (a_1, \ldots, a_5) \in \{0, 1\}^5$ index a corner, where $a_d = 0$ and $a_d = 1$ select the left and right nodes along axis d, respectively. The corresponding weight is

$$w_b(\boldsymbol{a} \mid \mathbf{q}_b) = \prod_{d=1}^{5} t_{b,d}^{a_d}(1 - t_{b,d})^{1-a_d}. \tag{5}$$

For the compact LUT $\mathbf{\Gamma}_b = \mathcal{C}_{\text{CP},r}(\mathbf{\Theta}_b^s)$ in Eq. 3, $\widetilde{\mathbf{\Theta}}_b[k_1 + a_1, \ldots, k_5 + a_5] \in \mathbb{R}^3$ denotes the three-channel node value defined in Eq. 1, with the output-channel index c omitted for brevity. The non-uniform pentalinear interpolation is

$$\mathcal{F}_b(\mathbf{q}_b; \mathbf{\Gamma}_b) = \sum_{\boldsymbol{a}\in\{0,1\}^5} w_b(\boldsymbol{a} \mid \mathbf{q}_b)\, \widetilde{\mathbf{\Theta}}_b[k_1 + a_1, \ldots, k_5 + a_5], \qquad b \in \mathcal{B}. \tag{6}$$

Algorithm 2 Non-uniform Pentalinear Interpolation

Input: Subband $b \in \mathcal{B}$; query $\mathbf{q}_b$; compact LUT $\mathbf{\Gamma}_b = \big(\mathcal{N}_b(\mathbf{I}_{\text{in}}), \{\mathbf{A}_{b,d}\}_{d=1}^{5}\big)$ with strictly increasing nodes.

Output: Lookup value $\mathcal{F}_b(\mathbf{q}_b; \mathbf{\Gamma}_b) \in \mathbb{R}^3$.

for $d = 1, \ldots, 5$ **do**
  Clip $q_{b,d}$ to $[\nu_{b,d}^{(1)}, \nu_{b,d}^{(K_{b,d})}]$.
  if $q_{b,d} = \nu_{b,d}^{(K_{b,d})}$ **then**
    $k_d \leftarrow K_{b,d} - 1$.
  else
    Use binary search in $\mathcal{N}_{b,d}(\mathbf{I}_{\text{in}})$ to find
    $k_d \in \{1, \ldots, K_{b,d} - 1\}$ such that $\nu_{b,d}^{(k_d)} \le q_{b,d} < \nu_{b,d}^{(k_d+1)}$.
  end if
  Compute the normalized position using Eq. 4:

$$t_{b,d} = \frac{q_{b,d} - \nu_{b,d}^{(k_d)}}{\nu_{b,d}^{(k_d+1)} - \nu_{b,d}^{(k_d)}}.$$

end for
for each $\boldsymbol{a} = (a_1, \ldots, a_5) \in \{0,1\}^5$ **do**
  Compute the corner weight using Eq. 5:

$$w_b(\boldsymbol{a} \mid \mathbf{q}_b) = \prod_{d=1}^{5} t_{b,d}^{a_d} (1 - t_{b,d})^{\,1-a_d}.$$

  Evaluate each weight factor as $1 - t_{b,d}$ if $a_d = 0$, and $t_{b,d}$ otherwise.
  for $c = 1, 2, 3$ **do**
    Evaluate the corner value from the CP factors using Eq. 1:

$$\widetilde{\mathbf{\Theta}}_b[k_1 + a_1, \ldots, k_5 + a_5, c] = \sum_{\ell=1}^{r} \prod_{d=1}^{5} \mathbf{A}_{b,d}[k_d + a_d, \ell, c].$$

  end for
end for
Compute the lookup value using Eq. 6:

$$\mathcal{F}_b(\mathbf{q}_b; \mathbf{\Gamma}_b) = \sum_{\boldsymbol{a} \in \{0,1\}^5} w_b(\boldsymbol{a} \mid \mathbf{q}_b)\, \widetilde{\mathbf{\Theta}}_b[k_1 + a_1, \ldots, k_5 + a_5].$$

return $\mathcal{F}_b(\mathbf{q}_b; \mathbf{\Gamma}_b)$.

Apply this procedure pointwise to each subband's query field. For $b = \text{LL}$, the returned value gives the corrected low-frequency coefficients. For $b \in \mathcal{B}_{\text{H}}$, it gives a high-frequency residual, which is subsequently added to the original subband $\mathbf{X}_b$.

For $b = \text{LL}$, $\mathcal{F}_{\text{LL}}$ directly outputs the corrected low-frequency coefficients. For $b \in \mathcal{B}_{\text{H}}$, $\mathcal{F}_b$ predicts a three-channel high-frequency residual, which is added to the original subband $\mathbf{X}_b$. The lookup is applied pointwise to each subband query field.

The pentalinear interpolation process is illustrated in Figure 1,and the pseudocode for Non-uniform Pentalinear Interpolation is presented in Algorithm 2.

3 Training Objectives Details

We sequentially train the coarse LUT predictor, the diffusion teacher, and the student model. Following Cai et al. (2026); Shin et al. (2025); Yi et al. (2025), we employ an image reconstruction loss and introduce smoothness and monotonicity regularization for the non-uniform frequency-decomposed LUTs. The image reconstruction loss is defined as

$$\mathcal{L}_{\text{mse}}(\mathbf{\Theta}) = \frac{1}{3HW} \left\| \widehat{\mathbf{I}}_{\text{out}} - \mathbf{I}_{\text{gt}} \right\|_F^2, \tag{7}$$

where H and W denote the image height and width, respectively. This loss directly constrains the final enhanced image. Since non-uniform node spacing affects the differences between adjacent LUT values, we regularize local mapping variations according to the actual node intervals. Let $\mathcal{E}_{b,d}$

denote the set of adjacent node pairs along the d-th query axis of subband b, where $(\mathbf{u}, \mathbf{v}) \in \mathcal{E}_{b,d}$ share identical coordinates on the remaining axes and satisfy $u_d < v_d$. The smoothness loss is formulated as

$$\mathcal{L}_{\mathrm{sm}}(\boldsymbol{\Theta}) = \frac{1}{15|\mathcal{B}|} \sum_{b\in\mathcal{B}} \sum_{d=1}^{5} \frac{1}{|\mathcal{E}_{b,d}|} \sum_{(\mathbf{u},\mathbf{v})\in\mathcal{E}_{b,d}} \left\| \frac{\boldsymbol{\Theta}_b(\mathbf{v}) - \boldsymbol{\Theta}_b(\mathbf{u})}{v_d - u_d} \right\|_2^2. \tag{8}$$

Here, $\boldsymbol{\Theta}_b(\mathbf{u}) \in \mathbb{R}^3$ denotes the three-channel LUT value at node $\mathbf{u}$. This term regularizes the local smoothness of both the low-frequency direct mapping and the high-frequency residual mapping. For the low-frequency LUT, we further impose monotonicity between each RGB input axis and its corresponding output channel. Let $c \in \{1, 2, 3\}$ denote the R, G, and B channels, respectively. The monotonicity loss is defined as

$$\mathcal{L}_{\mathrm{mon}}(\boldsymbol{\Theta}) = \frac{1}{3} \sum_{c=1}^{3} \frac{1}{|\mathcal{E}_{\mathrm{LL},c}|} \sum_{(\mathbf{u},\mathbf{v})\in\mathcal{E}_{\mathrm{LL},c}} \max\left\{0,\, [\boldsymbol{\Theta}_{\mathrm{LL}}(\mathbf{u})]_c - [\boldsymbol{\Theta}_{\mathrm{LL}}(\mathbf{v})]_c\right\}. \tag{9}$$

Here, $[\boldsymbol{\Theta}_{\mathrm{LL}}(\mathbf{u})]_c$ denotes the c-th output channel. Based on these losses, the coarse LUT predictor $\mathcal{G}_0$ is optimized by

$$\mathcal{L}_{\mathrm{coarse}} = \mathcal{L}_{\mathrm{mse}}(\boldsymbol{\Theta}^{(0)}) + \lambda_{\mathrm{sm}}\mathcal{L}_{\mathrm{sm}}(\boldsymbol{\Theta}^{(0)}) + \lambda_{\mathrm{mon}}\mathcal{L}_{\mathrm{mon}}(\boldsymbol{\Theta}^{(0)}), \tag{10}$$

where $\lambda_{\mathrm{sm}} \geq 0$ and $\lambda_{\mathrm{mon}} \geq 0$ control the weights of the smoothness and monotonicity regularizers, respectively. At this stage, $\mathcal{G}_0$ is optimized jointly with the image-adaptive non-uniform nodes and coarse LUT values. After coarse LUT training, we freeze $\mathcal{G}_0$. The diffusion teacher is then trained with the standard noise-prediction objective:

$$\mathcal{L}_{\mathrm{teacher}} = \mathcal{L}_{\mathrm{diff}} = \frac{1}{|\mathcal{B}|} \sum_{b\in\mathcal{B}} \mathbb{E}\left[\frac{\|\widehat{\boldsymbol{\epsilon}}_{b,\tau} - \boldsymbol{\epsilon}_b\|_F^2}{3|\mathcal{N}_b(\mathbf{I}_{\mathrm{in}})|} \right]. \tag{11}$$

Here, $|\mathcal{N}_b(\mathbf{I}_{\mathrm{in}})|$ denotes the total number of nodes in the five-dimensional LUT of subband b, and the expectation is taken over the training samples, diffusion timestep τ, and Gaussian noise $\boldsymbol{\epsilon}_b$. Once the teacher is trained, the teacher LUT $\boldsymbol{\Theta}^t$ is fixed. The continuous function distillation loss $\mathcal{L}_{\mathrm{func}}$ is defined. Finally, the student model is optimized by

$$\mathcal{L}_{\mathrm{student}} = \mathcal{L}_{\mathrm{mse}}(\boldsymbol{\Theta}^s) + \lambda_{\mathrm{func}}\mathcal{L}_{\mathrm{func}} + \lambda_{\mathrm{sm}}\mathcal{L}_{\mathrm{sm}}(\boldsymbol{\Theta}^s) + \lambda_{\mathrm{mon}}\mathcal{L}_{\mathrm{mon}}(\boldsymbol{\Theta}^s). \tag{12}$$

4 Experiments on the SMID and SDSD datasets

Comparison Methods. To evaluate our method against recent state-of-the-art (SOTA) methods, we compare it with NiLUT Conde et al. (2024), AttLUT Fu et al. (2024), SvdLUT Kim et al. (2025), Fastllve Li et al. (2023), WLUT He et al. (2026), DPLUT Lin et al. (2025), FLUT Hu et al. (2025), PanLUT Cai et al. (2026), DnLUT Yang et al. (2025), Darkir Feijoo et al. (2025), DpSPG Lu et al. (2025), LatentD Hu et al. (2026), TSCnet Zhang et al. (2025), Adair Cui et al. (2025), CIDNet Yan et al. (2024), SG-LLIE Dong et al. (2025), LYTNET Brateanu et al. (2025), FFTF Pei et al. (2025), CFWD Xue et al., and WV-LUT Li et al. (2025).

Quantitative Comparisons. As shown in Table 1, our method achieves the highest PSNR and SSIM and the lowest LPIPS on SMID, SDSD-in, and SDSD-out. On SMID, it improves PSNR by $0.08\,\mathrm{dB}$ over the second-best FLUT. Compared with LatentD, the second-best method in SSIM and LPIPS, it increases SSIM by 0.012 and reduces LPIPS by 0.023. On SDSD-in, it improves PSNR by $0.26\,\mathrm{dB}$ over LatentD, and improves SSIM by 0.012 and reduces LPIPS by 0.033 over FLUT. On SDSD-out, it improves PSNR by $0.36\,\mathrm{dB}$ over FLUT; compared with LatentD, it increases SSIM by 0.001 and reduces LPIPS by 0.034. These gains stem from local conditioning, frequency-specific restoration, and continuous mapping constraints. Local illumination and contrast allow similar colors under different exposure conditions to receive different mappings, while non-uniform nodes allocate more lookup resolution to regions with stronger mapping variation, reducing local brightness and color errors. SWT further separates photometric correction from high-frequency restoration: the low-frequency LUT recovers brightness and color, whereas the high-frequency residual LUT corrects

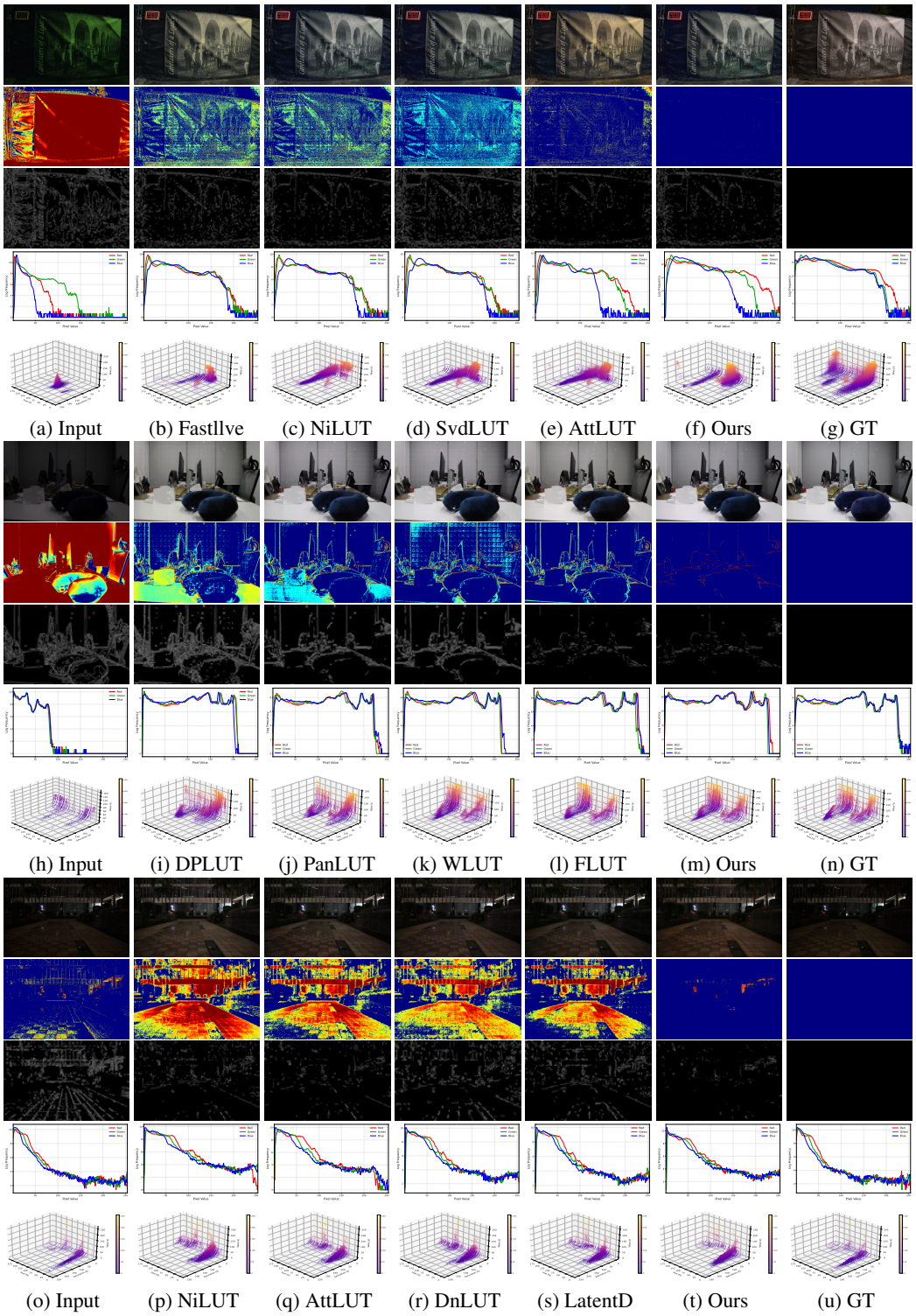
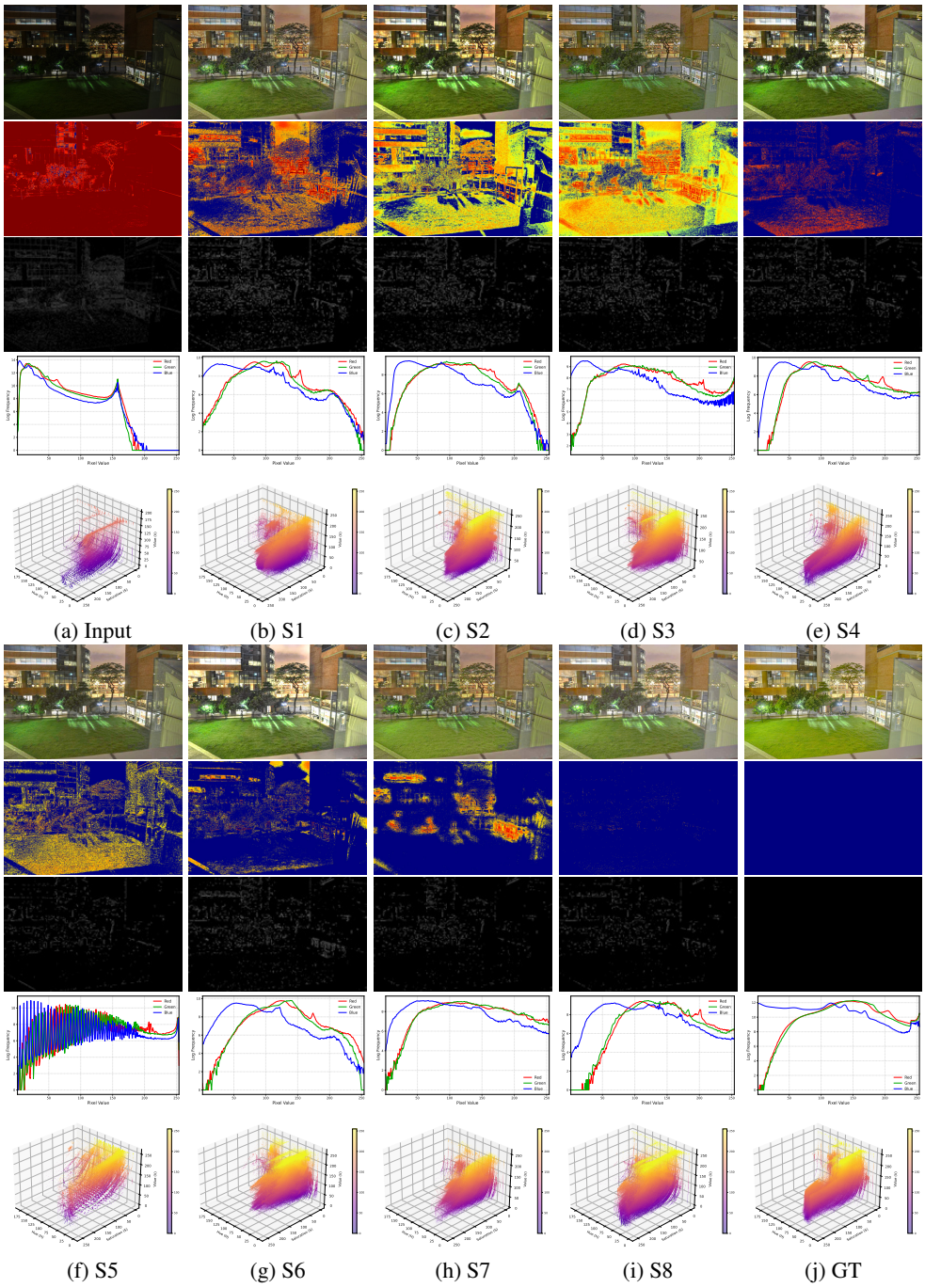


Figure 2: Qualitative comparison results on the SMID,SDSD-in and SDSD-out datasets

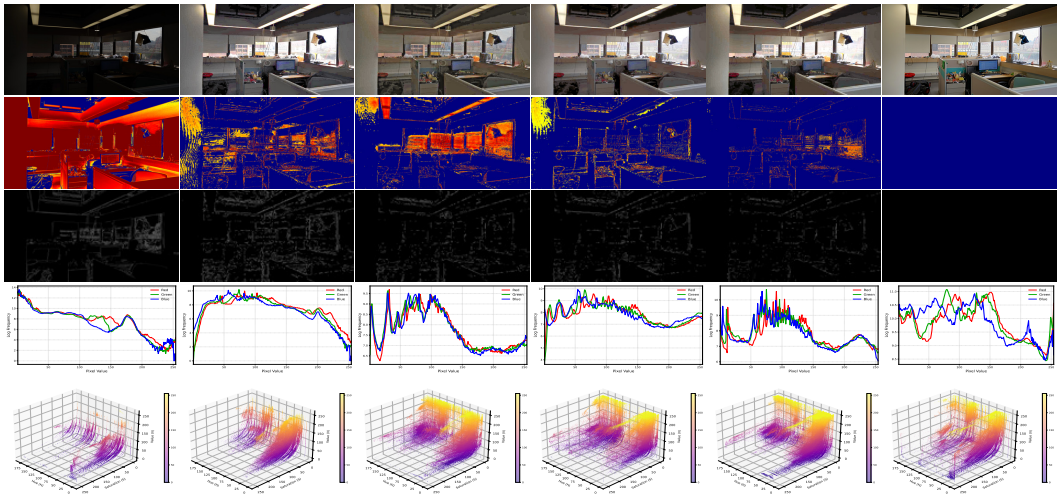
texture and noise, preserving structural details while suppressing noise. For mapping learning, the diffusion teacher refines the coarse LUT by modeling its residual to the target LUT, focusing correction on mappings that are difficult to predict directly. Continuous function distillation aligns teacher and student interpolation outputs rather than discrete node values, preserving the refined mapping in continuous query space. Grid queries further constrain locations not covered by image-derived samples, reducing errors in sparsely sampled regions. Together, these designs improve brightness and color recovery, structural fidelity, and perceptual quality, consistent with the gains in PSNR, SSIM, and LPIPS.



(a) Input (b) S1 (c) S2 (d) S3 (e) S4

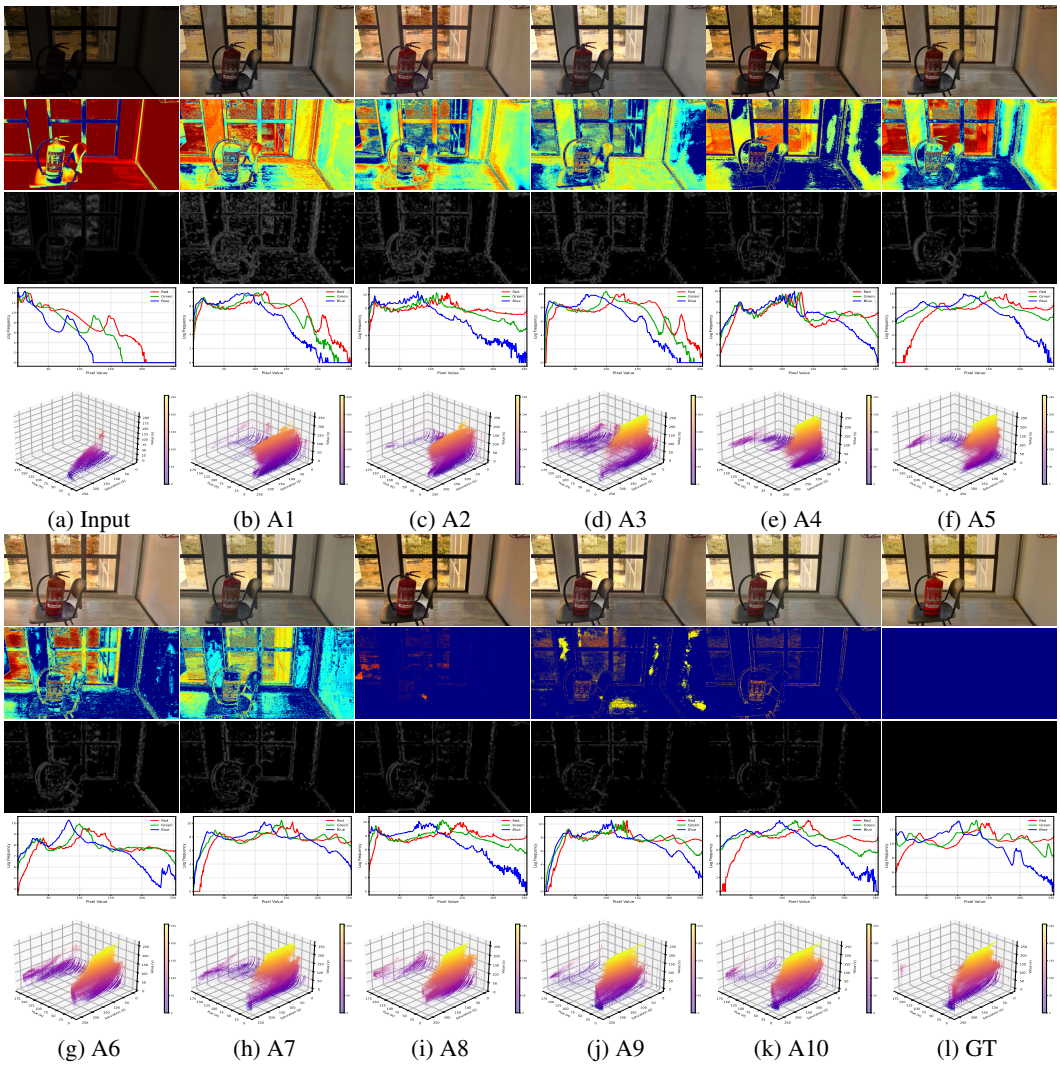
(f) S5 (g) S6 (h) S7 (i) S8 (j) GT

Figure 3: Loss Functions ablation about Student LUT $\boldsymbol{\Theta}^s$ on the SICE dataset.



(a) Input (b) C1 (c) C2 (d) C3 (e) C4 (f) GT

Figure 4: Loss Functions ablation about Coarse LUT $\mathbf{\Theta}^{(0)}$ on the SICE dataset.



(a) Input (b) A1 (c) A2 (d) A3 (e) A4 (f) A5

(g) A6 (h) A7 (i) A8 (j) A9 (k) A10 (l) GT

Figure 5: Network Architecture ablation on the SICE dataset.

Qualitative Comparisons. As shown in Figure 2, the rows from top to bottom present the enhanced results, pseudo-color error maps, SSIM difference maps, RGB histograms, and HSV distributions on SMID, SDSD-in, and SDSD-out. The pseudo-color maps visualize pixel-wise reconstruction error, with dark blue indicating smaller deviations and green-to-red indicating progressively larger errors. The SSIM maps reflect local structural discrepancies, where darker responses and weaker contours indicate smaller structural errors. RGB histograms and HSV distributions further assess the agreement of intensity and color statistics with the GT.

On SMID, Fastllve, NiLUT, and SvdLUT retain large errors over the poster and surrounding regions, with clear structural residuals around the person and object boundaries. AttLUT reduces part of the error area, but noticeable deviations remain within and around the poster. In contrast, our method substantially suppresses regional errors and weakens the structural residuals along object contours. Its RGB histograms are closer to the GT in overall trend and inter-channel relation, while the HSV distribution also shows better agreement in position and shape, indicating more accurate brightness and color recovery.

On SDSD-in, DPLUT produces pronounced errors over the white tabletop and dark cushion. PanLUT, WLUT, and FLUT reduce part of the pixel error, yet continuous residuals remain around the cushion, monitor, and tabletop boundaries. Our method maintains low error over the tabletop and background while reducing structural discrepancies in dark regions and along object edges. The corresponding RGB histograms more closely match the GT intensity distribution, and the HSV distribution is also more consistent in position and shape, indicating more uniform correction across bright, dark, and transition regions.

On SDSD-out, NiLUT, AttLUT, DnLUT, and LatentD retain large errors over the road and buildings, while the road surface and building boundaries remain prominent in the SSIM maps. Our method markedly reduces regional errors over these areas and weakens the residuals along road edges and building contours. Its RGB histograms follow the GT more closely, and the HSV distribution shows higher agreement, indicating more balanced recovery of brightness, color, and structure in nighttime scenes.

These results show that our method reduces both pixel-wise and structural errors while bringing the restored intensity and color distributions closer to the GT. Local conditioning and the non-uniform LUT improve regional correction accuracy, while frequency-specific modeling balances photometric restoration with high-frequency detail recovery, leading to more stable recovery of exposure, color, and structure across all three benchmarks.

5 Visualization Results of Ablation Study

Qualitative Analysis of Loss Functions. The qualitative ablations of the loss functions are shown in Figures 3 and 4. From top to bottom, the figures present the enhanced results, pseudo-color error maps, SSIM difference maps, RGB histograms, and HSV distributions, revealing the effects of different loss configurations on pixel reconstruction, structural preservation, and color recovery.

For the student LUT, MSE-only training ($\mathcal{L}_{\mathrm{mse}}$, S1) leaves large errors over the grass, buildings, and shadow regions, together with pronounced structural residuals in the SSIM maps and clear deviations from the GT in the RGB and HSV distributions. Image-domain supervision constrains the final output but provides limited control over interpolation between LUT nodes. Adding smoothness regularization ($\mathcal{L}_{\mathrm{sm}}$, S2) reduces these regional errors. Without function distillation ($\mathcal{L}_{\mathrm{func}}$, S3), both pixel and structural residuals increase, as the student is no longer constrained by the teacher outputs at continuous query locations. Removing the reconstruction loss ($\mathcal{L}_{\mathrm{mse}}$, S4) produces larger local brightness and color deviations because the predicted mapping is no longer directly anchored to the GT. Replacing function distillation with node-value matching ($\mathcal{L}_{\mathrm{node}}$, S5) introduces stronger fluctuations in the RGB histograms and larger shifts in the HSV distribution, showing that agreement at discrete nodes does not ensure consistent interpolation within LUT cells. Removing smoothness regularization ($\mathcal{L}_{\mathrm{sm}}$, S6) and monotonicity regularization ($\mathcal{L}_{\mathrm{mon}}$, S7) produces stronger local residuals and color deviations, respectively, consistent with their effects on local LUT variation and color ordering. The complete objective (S8) yields the smallest pseudo-color errors, the weakest SSIM residuals, and the closest RGB and HSV distributions to the GT.

Table 1: Quantitative comparison on the SMID and SDSD test datasets. The best results are highlighted in bold.

Method	SMID			SDSD-in			SDSD-out		
	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
NiLUT	25.84	0.812	0.238	25.42	0.798	0.251	25.67	0.805	0.244
AttLUT	26.46	0.872	0.181	26.21	0.861	0.194	26.38	0.868	0.186
SvdLUT	25.93	0.841	0.213	25.68	0.829	0.226	25.86	0.836	0.219
Fastllve	24.86	0.758	0.294	24.41	0.742	0.312	24.75	0.751	0.303
WLUT	27.03	0.908	0.143	26.94	0.901	0.152	26.98	0.905	0.148
DPLUT	25.41	0.793	0.261	25.16	0.781	0.274	25.33	0.788	0.268
FLUT	27.18	0.919	0.126	27.02	0.921	0.129	27.16	0.923	0.123
PanLUT	26.96	0.917	0.134	26.83	0.909	0.145	27.04	0.922	0.128
DnLUT	26.68	0.896	0.158	26.54	0.887	0.169	26.73	0.901	0.153
LatentD	27.14	0.937	0.112	27.11	0.916	0.132	27.09	0.927	0.119
Ours	**27.26**	**0.949**	**0.089**	**27.37**	**0.933**	**0.096**	**27.52**	**0.928**	**0.085**

The coarse LUT ablation follows a similar pattern. With MSE alone ($\mathcal{L}_{\mathrm{mse}}$, C1), regional errors, structural residuals, and color-distribution shifts remain evident. Removing smoothness regularization ($\mathcal{L}_{\mathrm{sm}}$, C2) or monotonicity regularization ($\mathcal{L}_{\mathrm{mon}}$, C3) leaves different degrees of local error, whereas the complete objective (C4) produces the closest agreement with the GT in pixel error, structural residuals, and color distribution. Since the coarse LUT provides the initial mapping for target inversion and residual diffusion, its quality further affects subsequent refinement.

Overall, the reconstruction loss $\mathcal{L}_{\mathrm{mse}}$ anchors the output to the GT, function distillation $\mathcal{L}_{\mathrm{func}}$ constrains continuous interpolation, while $\mathcal{L}_{\mathrm{sm}}$ and $\mathcal{L}_{\mathrm{mon}}$ regulate local LUT variation and color ordering. Their combination reduces pixel errors, structural residuals, and color deviations simultaneously.

Qualitative Analysis of Network Architecture.

The qualitative architecture ablation is shown in the Figure 5. From top to bottom, it presents the enhanced results, pseudo-color error maps, SSIM difference maps, RGB histograms, and HSV distributions. A1 corrects only the low-frequency component. Although the overall brightness is recovered, clear pixel and structural errors remain around the chair, window frames, and wall regions, indicating that high-frequency degradation cannot be compensated by low-frequency mapping alone. In A2, replacing local conditions with image-wise means still leaves pronounced regional and color errors, showing that global statistics cannot adequately capture spatially varying exposure states.

The errors gradually decrease from A3 to A7, yet each variant retains characteristic deviations. When the student is supervised directly by inverted target LUTs in A3, continuous residuals remain along the wall and object boundaries, suggesting that target table values alone do not sufficiently constrain the continuous mapping. A4 directly predicts complete high-frequency coefficients, but noticeable structural errors persist around edges, indicating that residual prediction is better suited to recovering degraded details while preserving the original high-frequency content. Using uniform nodes in A5 increases local brightness and color errors. Removing grid queries in A6 produces more dispersed regional residuals, reflecting the limited coverage of the LUT domain by image-derived queries alone. A7 diffuses complete LUTs rather than residuals and still exhibits evident regional errors, showing that residual refinement around the coarse LUT provides a more effective correction of the initial mapping.

Direct teacher inference in A8 already yields low pixel errors, but relatively strong boundary responses remain in the SSIM map. After student distillation, A9 further reduces structural residuals, while its RGB and HSV distributions move closer to the GT. A10 remains nearly identical to A9 after CP compression, indicating that the compact reparameterization preserves the student LUT mapping. Overall, the complete configuration shows the closest agreement with the GT in pixel error, structural residuals, and color distribution. Local conditioning, high-frequency residual correction, non-uniform sampling, diffusion refinement, and continuous distillation together lead to more stable recovery of brightness, color, and structure.

AI USE STATEMENT

Generative AI was not used in the conception of this study, method design, experimental analysis, or code implementation. Its use was limited to language assistance, including grammar correction, sentence refinement, and academic polishing. All AI-assisted revisions were reviewed and confirmed by the authors. The research ideas, technical methods, experimental results, analyses, and conclusions presented in this work were independently developed and verified by the authors. The authors take full responsibility for the scientific validity, accuracy, and final content of the manuscript, including any text refined with the assistance of generative AI.

REFERENCES

Alexandru Brateanu, Raul Balmez, Adrian Avram, Ciprian Orhei, and Cosmin Ancuti. Lyt-net: Lightweight yuv transformer-based network for low-light image enhancement. *IEEE Signal Processing Letters*, 32:2065–2069, 2025.

Zhongnan Cai, Yingying Wang, Hui Zheng, Panwang Pan, ZiXu Lin, Ge Meng, Chenxin Li, Chunming He, Jiaxin Xie, Yunlong Lin, et al. Pan-lut: Efficient pan-sharpening via learnable look-up tables. *Advances in Neural Information Processing Systems*, 38:126459–126483, 2026.

Marcos V Conde, Javier Vazquez-Corral, Michael S Brown, and Radu Timofte. Nilut: Conditional neural implicit 3d lookup tables for image enhancement. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 38, pp. 1371–1379, 2024.

Yuning Cui, Syed Waqas Zamir, Salman Khan, Alois Knoll, Mubarak Shah, and Fahad Khan. Adair: Adaptive all-in-one image restoration via frequency mining and modulation. In *International Conference on Learning Representations*, volume 2025, pp. 101306–101327, 2025.

Wei Dong, Yan Min, Han Zhou, and Jun Chen. Towards scale-aware low-light enhancement via structure-guided transformer design. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1454–1461. IEEE, 2025.

Daniel Feijoo, Juan C Benito, Alvaro Garcia, and Marcos V Conde. Darkir: Robust low-light image restoration. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10879–10889. IEEE, 2025.

Kang Fu, Yicong Peng, Zicheng Zhang, Qihang Xu, Xiaohong Liu, Jia Wang, and Guangtao Zhai. Attentionlut: Attention fusion-based canonical polyadic lut for real-time image enhancement. In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3255–3259. IEEE, 2024.

Jinhong He, Minglong Xue, Wenhai Wang, and Mingliang Zhou. Optimizing a 4d lookup table for low-light video enhancement via wavelet priori. *IEEE Transactions on Multimedia*, 2026.

Liubing Hu, Chen Wu, Anrui Wang, Dianjie Lu, Guijuan Zhang, and Zhuoran Zheng. Flowlut: Efficient image enhancement via differentiable luts and iterative flow matching. *arXiv preprint arXiv:2509.23608*, 2025.

Renzhi Hu, Ting Luo, Gangyi Jiang, Leiming Liu, Yeyao Chen, Haiyong Xu, and Zhouyan He. Latentdark: Reflectance guided latent diffusion model for low-light image enhancement. *Signal Processing*, 238:110125, 2026.

Wontae Kim, Keuntek Lee, and Nam Ik Cho. Lightweight and fast real-time image enhancement via decomposition of the spatial-aware lookup tables. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 11895–11905. IEEE, 2025.

Canlin Li, Haowen Su, Xin Tan, Xiangfei Zhang, and Lizhuang Ma. Wv-lut: Wide vision lookup tables for real-time low-light image enhancement. *IEEE Transactions on Multimedia*, 27:4441–4453, 2025.

Wenhao Li, Guangyang Wu, Wenyi Wang, Peiran Ren, and Xiaohong Liu. Fastllve: Real-time low-light video enhancement with intensity-aware look-up table. In *Proceedings of the 31st ACM International Conference on Multimedia*, pp. 8134–8144, 2023.

Yunlong Lin, Zhenqi Fu, Kairun Wen, Tian Ye, Sixiang Chen, Ge Meng, Yingying Wang, Chui Kong, Yue Huang, Xiaotong Tu, et al. Dplut: Unsupervised low-light image enhancement with lookup tables and diffusion priors. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 5316–5324, 2025.

Jialang Lu, Huayu Zhao, Huiyu Zhai, Xingxing Yang, and Shini Han. Deepspg: Exploring deep semantic prior guidance for low-light image enhancement with multimodal learning. In *Proceedings of the 2025 International Conference on Multimedia Retrieval*, pp. 935–943, 2025.

Xiao Pei, Yongdong Huang, Weijian Su, Fengjuan Zhu, and Qiang Liu. Fftformer: A spatial-frequency noise aware cnn-transformer for low light image enhancement. *Knowledge-Based Systems*, 314:113055, 2025.

Seunghyun Shin, Dongmin Shin, Jisu Shin, Hae-Gon Jeon, and Joon-Young Lee. Video color grading via look-up table generation. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 19141–19152. IEEE, 2025.

M Xue, J He, W Wang, and M Zhou. Low-light image enhancement via clip-fourier guided wavelet diffusion. arxiv 2024. *arXiv preprint arXiv:2401.03788*.

Qingsen Yan, Yixu Feng, Cheng Zhang, Pei Wang, Peng Wu, Wei Dong, Jinqiu Sun, and Yanning Zhang. You only need one color space: An efficient network for low-light image enhancement. *arXiv preprint arXiv:2402.05809*, 2024.

Sidi Yang, Binxiao Huang, Yulun Zhang, Dahai Yu, Yujiu Yang, and Ngai Wong. Dnlut: Ultra-efficient color image denoising via channel-aware lookup tables. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pp. 7582–7591, 2025.

Xunpeng Yi, Yibing Zhang, Xinyu Xiang, Qinglong Yan, Han Xu, and Jiayi Ma. Lut-fuse: Towards extremely fast infrared and visible image fusion via distillation to learnable look-up tables. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 14559–14568. IEEE, 2025.

Miao Zhang, Jun Yin, Pengyu Zeng, Yiqing Shen, Shuai Lu, and Xueqian Wang. Tscnet: A text-driven semantic-level controllable framework for customized low-light image enhancement. *Neurocomputing*, 625:129509, 2025.