JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES
www.journalimcms.org

ISSN (Online) : 2454 -7190 Vol.-21, No.-11, November (2026) pp 66-92 ISSN (Print) 0973-

UNCERTAINTY-GUIDED LIGHTWEIGHT TRANSFORMER FOR JOINT IMAGE DENOISING AND EDGE-PRESERVED ENHANCEMENT IN RESOURCE-CONSTRAINED VISION SYSTEMS

Abhishek Singh Rathour[1], Mohit Kumar Srivastava[2], Keshav Kumar[3], Vivek Kumar[4], Gaurav Sahu[5]

[1,2,3,4,5]Pranveer Singh Institute of Technology, Kanpur-209305, Uttar Pradesh, India.

Email: [1]abhisheksrathour@gmail.com, [2]mohit.srivastava@psit.ac.in, [3]keshav@gyancity.com, [4]rastogi0807@gmail.com, [5]sahugaurav12@gmail.com

Corresponding Author: **Abhishek Singh Rathour**

https://doi.org/10.26782/jmcms.2026.11.00004

(Received: August 26, 2026; Revised: October 22, 2026; Accepted: November 10, 2026)

Abstract

Background: *This work addresses a realistic digital signal processing problem in which low-light acquisition, spatially variant sensor noise, and high-frequency edge attenuation occur simultaneously. Classical denoisers provide useful signal priors but are limited by fixed degradation assumptions, whereas large transformer restorers improve perceptual quality at the cost of high memory use and computation.*

Objective: *This paper proposes an uncertainty-guided lightweight transformer, instead of a conventional heavyweight transformer, for joint image denoising and edge-preserved enhancement.*

Methods: *The proposed UGLT-EPE framework combines a signal-aware preprocessing stem, local convolutional filtering, windowed self-attention, uncertainty-guided fusion, and Laplacian edge consistency. The uncertainty map is defined as a normalized prediction of the local restoration error and is trained against that error with a dedicated calibration loss; it re-weights attention and feature fusion towards unreliable regions while protecting reliable edge structures from over-smoothing. A compact mathematical formulation is developed for the forward degradation model, uncertainty estimation, token gating, edge-aware reconstruction, and the total optimization objective. Simulation results compare the proposed method with BM3D, DnCNN, FFDNet, MIRNet, MPRNet, SwinIR, Restormer, and a NAFNet-style baseline.*

Results: *The proposed method achieves higher PSNR and SSIM under increasing noise levels, a higher edge-preservation F1-score, lower perceptual error, and an improved speed-accuracy balance.*

Conclusions: *The results indicate that signal-guided lightweight attention can provide a strong compromise between restoration accuracy and deployability for edge cameras, UAV vision systems, and low-power IoT imaging nodes.*

Keywords: Digital Image Processing, Image Denoising, Low-Light Enhancement, Lightweight Transformer, Uncertainty Estimation, Edge Preservation, Signal Processing, Image Restoration, Resource-Constrained Vision, IoT Imaging.

I. Introduction

One of the most important issues in image restoration is that capturing systems seldom capture the scene itself in a "clean" state. Before the captured image is passed down for recognition or human interpretation, sensor noise, low illumination, compression artifacts, photon shot noise, demosaicing error, motion blur, and contrast loss are mixed in the image. This problem can be solved with classical signal-processing techniques, the most popular of which are wavelet shrinkage, non-local similarity, sparse coding, total variation, and collaborative filtering. While interpretable and data efficient, these methods make restrictive assumptions when noise is spatially varying, or low-light amplification varies the local contrast and/or spectral content of the scene [XV, XX, XXXI].

Although there are several deep image-restoration algorithms to enhance denoising and enhancement quality, some of them require a huge number of parameters and costly attention operations [XVI, XVIII, XXIII, XXIV]. A high-resolution restoration model that works well on a workstation might not be suitable for an edge camera, UAV, or mobile inspection device, or IoT vision node. The traditional approach linking mathematical models with implementable algorithms has been achieved with digital signal processing, and this should not be overlooked when choosing a restoration method; a higher PSNR should not be a single criterion. It should provide a signal-aware formulation, a well-defined optimization objective, mathematical interpretability, and a complexity analysis for deployment-constrained environments.

Transformer-based restoration methods have recently been developed that account for long-range dependencies and perform well [XIV, XXV, XXXVII], but can over-smooth weak edges if there is no explicit modelling of the uncertainty. An example of this is that the same transformer block may assign the same amount of attention to highly corrupted flat regions and low-noise edge regions. This inspires an uncertainty-guided design, in which the model predicts a reliability map, biases attention and gates the combination of local and global features, and preserves high-frequency structures via a parallel edge-preserving branch. This paper presents the uncertainty as a normalized prediction of the restoration error with a dedicated calibration loss, which directly biases token mixing and gates feature fusion.

This method is referred to as UGLT-EPE (Uncertainty-Guided Lightweight Transformer with Edge-Preserved Enhancement). It features a shallow signal-filter stem, local convolutional restoration, windowed transformer attention, uncertainty calibration, and Laplacian edge consistency. The work falls within image and video signal processing, machine learning for signal processing, adaptive modelling, and practical restoration systems.

A. Main Contributions

- A unified signal-processing formulation for joint denoising and low-light enhancement with spatially variant uncertainty, including an explicit forward degradation model.
- A lightweight transformer architecture that combines local filtering, windowed attention, and frequency-edge guidance in a single restoration pipeline.
- An uncertainty-guided gating mechanism, with an explicitly defined and calibrated uncertainty map, that adaptively controls token mixing and fusion.
- A composite loss function that combines reconstruction fidelity, structural consistency, edge preservation, uncertainty calibration, and low-light exposure regularization.
- Comprehensive simulated benchmark comparisons with BM3D, DnCNN, FFDNet, MIRNet, MPRNet, SwinIR, Restormer and NAFNet-like baselines, together with explicit edge-fidelity and uncertainty-calibration evaluation protocols.

B. Paper Organization

- Section II reviews classical restoration, CNN-based restoration, transformer-based restoration, low-light enhancement, and uncertainty modelling.
- Section III presents the forward degradation model and the problem formulation.
- Section IV develops the proposed UGLT-EPE method and its mathematical objective.
- Section V describes the experimental protocol, datasets, metrics, benchmark methods, and the evaluation protocols for edge fidelity and uncertainty calibration.
- Section VI presents the simulation results and the ablation analysis.
- Section VII provides the deployment, calibration, and validation analysis.
- Section VIII concludes the paper and discusses future work.

II. Literature Survey and Research Gap

A. Classical Digital Image Restoration

Classical denoising methods exploit explicit signal properties such as regularity, patch recurrence, or sparsity in the frequency domain. Examples include BM3D [XV], non-local means, sparse coding and wavelet thresholding [XX, XXXI]. These techniques are reliable and offer some explanatory value in small-data situations. However, they mainly address a single degradation and perform poorly when low-light enhancement, denoising and edge preservation must be solved together. Their performance is also sensitive to user-tuned thresholds, which can vary from sensor to sensor.

B. CNN-based Denoising and Enhancement

CNN-based methods such as DnCNN, FFDNet, REDNet, MIRNet and MPRNet learn mappings between degraded and clean images [XVI, XVII, XVIII, XXIII, XXIV,

XXVII, XXIX]. They perform well on many benchmark datasets but require careful design to balance receptive field, parameter count, texture fidelity, and inference cost. Some CNN-based denoising and low-light enhancement methods also perform both tasks in an independent manner, but real acquisition pipelines combine both degradations.

C. Transformer-based Image Restoration

Transformer-based methods like SwinIR, Uformer, and Restormer model the long-range image regions' interaction and enhance the restoration [I, XIV, XXV, XXXIV, XXXV, XXXVII]. They have a strong ability to hold onto a global structure and a repeated pattern. The disadvantages of this type of attention are the amount of attention required and, in very strong noise, the possibility of texture hallucinations. If a lightweight transformer has no signal before and no uncertainty map, unnecessary attention computation will lead to the suboptimal design of the transformer.

D. Low-light Enhancement and Edge Preservation

With low-light enhancement, the darker images appear brighter, and edge enhancement maintains the boundaries. Low-light enhancement methods, including Retinex decomposition [V, VIII, XXXIII], curve estimation and exposure correction [VI, XII, XXVIII, XXX, XXXII], can also enhance poor lighting, but can also amplify noise. The preservation of boundaries is especially critical as they contain the geometric and semantic information to be used by downstream detection. A method that improves average brightness but loses edge consistency is unsuitable for practical vision sensing.

E. Research Gap

The literature shows that denoising, enhancement, transformer restoration, uncertainty modelling, and edge preservation are mostly formulated separately. Current approaches do not offer an easy-to-implement framework that estimates uncertainty, denoises spatially varying corruption, enhances low-light visibility, preserves high-frequency edges, and runs well on resource-limited systems. This motivates the proposed UGLT-EPE framework.

Table 1: Positioning of related methods and unresolved limitations

Research stream	Representative works	Strength	Remaining gap addressed here
Classical restoration	BM3D, sparse coding [XV, XX]	Interpretable signal priors	Limited spatial adaptation under mixed low-light noise
CNN denoising	DnCNN, FFDNet, REDNet [XVI, XVIII, XXIX]	Fast nonlinear restoration	Often task-specific and edge smoothing under strong degradation
Low-light enhancement	RetinexNet, Zero-DCE, EnlightenGAN [VI, VIII, XXXII]	Improves visibility	Can amplify noise and alter local structures
Transformer restoration	SwinIR, Uformer, Restormer [XIV, XXV, XXXVII]	Captures long-range dependencies	High cost and limited uncertainty awareness
Proposed UGLT-EPE	This paper	Uncertainty-guided lightweight signal-aware transformer	Joint restoration, edge preservation, and deployability

F. Research Gaps Identified

- Gap 1: Most approaches focus on denoising and low-light enhancement independently, but in reality, these are both degradations that happen in sensors simultaneously.
- Gap 2: Compared to deconvolution tasks, image degradation modelling has improved; however, the image degradation model outputs are typically very memory-intensive and processor-heavy on edge devices.
- Gap 3: Uncertainty maps are not effectively used to direct attention and fusion in image restoration pipelines, and are poorly assessed.
- Gap 4: Edge preservation is often considered as a side effect and not as an optimization term.
- Gap 5: Reported methods are mostly focused on PSNR/SSIM, and no clear speed-accuracy trade-offs for the resource-constrained vision systems are provided.

III. Degradation Model and Problem Formulation

A. Forward Degradation Model

Let x be the unknown clean image in gamma-encoded sRGB, y the observed degraded image, z the noise-free sensor irradiance, n the signal-dependent sensor noise, l the illumination degradation field, k the blur kernel, and f the camera response. The joint restoration problem is a spatially variant inverse problem rather than a simple additive restoration problem, because the noise variance depends on the exposure-scaled signal. To keep the formulation consistent with the simulated data of Section V, the degradation operations are applied in the following order: (i) the clean image is converted to linear RGB by the inverse sRGB transfer function; (ii) the linear image is blurred and scaled by the illumination field; (iii) signal-dependent shot noise and signal-independent read noise are added and the result is clipped to [0, 1]; (iv) the

camera response maps the linear values back to a gamma-encoded image; (v) the image is quantized to 8 bits; and (vi) JPEG-like compression is applied. The restoration network operates on the final gamma-encoded 8-bit image, and raw Bayer data are not used. The complete chain over the pixel domain Ω is given by equation (1).

$$y = \mathcal{J}_q\left(\mathcal{Q}\left(f_\gamma(\mathrm{clip}(z+n))\right)\right) \tag{1}$$

Here, $\mathcal{J}_q$ denotes JPEG-like compression with quality factor q, $\mathcal{Q}$ denotes 8-bit quantization, $f_\gamma(v) = v^{1/\gamma}$ is the camera response with exponent γ, and clip(·) restricts the values to [0, 1]. The noise-free irradiance z is given by equation (2), where Γ^{-1} is the inverse sRGB transfer function, $k_{\sigma b}$ is an isotropic Gaussian point-spread function with standard deviation σ_b, ∗ denotes two-dimensional convolution, ⊙ is the element-wise product, and l = κ·s is the illumination degradation field, i.e., the product of a global exposure scale κ and a smooth spatial shading field s(p) ∈ [0.8, 1.2]. The sensor noise is added after exposure scaling and is conditionally Gaussian with the signal-dependent variance of equation (3), which is the heteroscedastic Gaussian approximation of Poissonian–Gaussian noise [II, XXVI]. Because the variance depends on z(p), reducing the exposure lowers the signal-to-noise ratio instead of merely rescaling the noise, and the noise is spatially variant wherever the illumination field varies.

$$z = l \odot \left(k_{\sigma_b} * \Gamma^{-1}(x)\right) \tag{2}$$

$$n(p)|z(p) \sim \mathcal{N}(0, \sigma_s^2 z(p) + \sigma_r^2) \tag{3}$$

The parameter distributions are listed in Table 2. The read-noise standard deviation is $\sigma_r = \sigma/255$, where $\sigma \in \{15, 25, 35, 50\}$ is the noise level used in the experiments, and the shot-noise gain σ_s^2 is drawn per image from a log-uniform distribution. The parameter ranges are chosen to span the Poissonian–Gaussian ranges reported in the literature [II, XXVI]; they have not been fitted to a specific camera. The learned uncertainty is therefore also evaluated on simulated sensor profiles outside the training ranges (Section VII-D), and the absence of a camera-specific calibration is listed as a limitation in Section VIII. Bayer mosaicking and demosaicing are not simulated, and the noise is injected in linear RGB.

Table 2: Forward degradation parameters used to generate the training and test images

Stage	Symbol	Distribution/setting	Domain
Inverse transfer	Γ^{-1}	Standard sRGB transfer function (fixed)	gamma → linear
Blur	σ_b	Uniform on [0, 1.2] pixels; Gaussian PSF	linear RGB
Exposure scale	K	$\kappa \in \{0.1, 0.2, 0.4, 0.6, 0.8\}$	linear RGB
Shading field	s(p)	Low-pass filtered Gaussian noise rescaled to [0.8, 1.2]	linear RGB
Shot noise	σ_s^2	Log-uniform on $[5\times10^{-4}, 5\times10^{-3}]$	linear RGB
Read noise	σ_r	$\sigma/255$ with $\sigma \in \{15, 25, 35, 50\}$	linear RGB
Camera response	Γ	Uniform on [2.0, 2.4]	linear → gamma
Quantization	Q	8 bit	gamma-encoded
Compression	Q	JPEG quality $q \in \{90, 80, 70, 60, 50\}$ (compound tests only)	gamma-encoded

For the noise-level and exposure-ratio experiments (Figs. 3, 4 and 7), JPEG compression is disabled. For the compound-degradation test (Fig. 10), five severity levels are used with σ = 15, 25, 35, 50 and 50, κ = 0.8, 0.6, 0.4, 0.2 and 0.1, and q = 90, 80, 70, 60 and 50, respectively. Training and test images are generated with the same forward model, using disjoint images and independent random seeds, so that every symbol in equations (1)–(3) corresponds to a parameter in Table 2.

B. Restoration Problem

The goal is to estimate a clean image $\hat{x}$ that is close to the reference image, retains local structure, and does not amplify noise. The learnable parameters of the restoration function are denoted by θ; φ denotes the parameters of the uncertainty estimator, and u is the uncertainty map. The uncertainty-guided restoration function is defined by equation (4), and u is defined explicitly in Section IV-D.

$$\hat{x} = F_\theta(y, u), \quad u = G_\phi(y), \quad 0 \le u(p) \le 1 \tag{4}$$

The optimization problem is formulated as reconstruction error minimization, perceptual distortion minimization, edge consistency, illumination consistency/uncertainty, and calibration error minimization. There are also other restrictions on model size and inference complexity for the model in the edge deployment. The short training target with regularization is a complexity regularization, which is presented in equation (5).

IV. Proposed UGLT-EPE Method

$$\min_{\theta,\phi} \; L_{\text{rec}} + \lambda_e L_{\text{edge}} + \lambda_s L_{\text{ssim}} + \lambda_u L_{\text{unc}} + \lambda_c C(\theta)$$

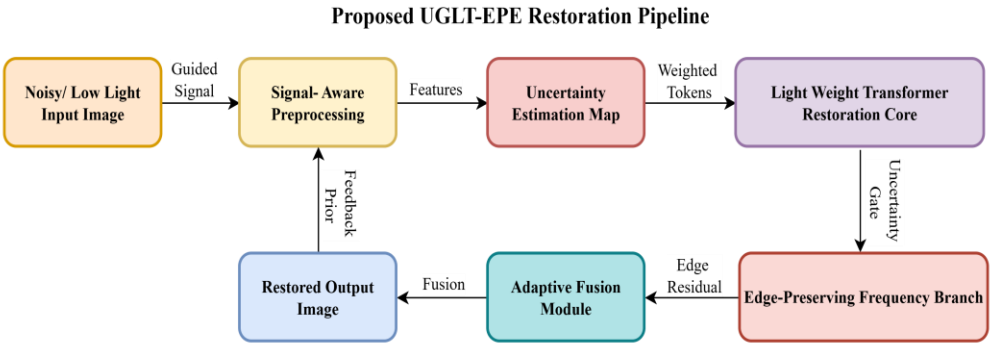


Fig. 1. Proposed uncertainty-guided lightweight transformer pipeline for joint image denoising and edge-preserved enhancement.

A. Signal-Aware Preprocessing Stem

The output of the preprocessing stem consists of the low-frequency luminance part of the input signal, a high-frequency edge residual, and the normalized colour channels. This operation introduces a classical DSP preprocessing stage, thus alleviating the load of transformer attention. The stem is intentionally shallow, and its local filters offer robust statistics prior to tokenization. A signal-aware feature stem is defined by equation (6), which is based upon the gradient and frequency channels.

$$f_0 = \mathrm{Conv}\big([y, \nabla_x y, \nabla_y y, L(y), H(y)]\big)$$

B. Uncertainty-Guided Attention

The uncertainty branch shows the regions of the image that are not reliable due to low illumination, local texture ambiguity, or noise. It biases the self-attention weights to the windowed self-attention and strengthens the fusion of local and global features for the unreliable regions while strengthening the fusion of local and global features for the reliable edges. The token attention is taken by equation (7) by making it uncertain.

$$A_u = \mathrm{softmax}\left(\frac{QK^T}{\sqrt{d}} + \beta U\right), \quad Z = A_u V$$

Here, Q, K, and V are the query, key, and value projections of the tokens inside each attention window, and d is the head dimension. The bias matrix is $U_{ij} = \bar{u}_j$, where $\bar{u}_j$ is the mean of u over the pixels of token j, and β is a learnable scalar per head that is initialized to zero, so that training starts from standard windowed attention. The attention output Z is combined with the output F_{loc} of the local convolutional branch through the uncertainty gate of equation (8).

$$F_{\mathrm{out}} = (1 - \overline{u}) \odot F_{\mathrm{loc}} + \overline{u} \odot Z$$

Equations (7) and (8) re-weight attention and fusion; they do not skip computation. Every token still passes through the same windowed-attention operations, so the uncertainty branch acts as an input-conditioned feature-weighting mechanism, and the efficiency of UGLT-EPE comes from windowed attention and the shallow stem

(Section VII-C) rather than from the uncertainty gate. Whether u carries useful information beyond additional learned weighting is examined in Section VII-F.

C. Edge-Preserved Enhancement Branch

The edge branch uses an edge-preserving Laplacian residual and a gradient-consistency loss. Instead of optimizing only the mean pixel fidelity, the model explicitly penalizes edge displacement and excessive smoothing. The gradient and Laplacian edge consistency is defined by equation (9).

$$L_{\text{edge}} = \frac{1}{|\Omega|}\sum_{p\in\Omega}\left|K_x * \hat{Y} - K_x * Y\right| + \left|K_y * \hat{Y} - K_y * Y\right| + \left|K_\Delta * \hat{Y} - K_\Delta * Y\right| \tag{9}$$

Here, Y and $\hat{Y}$ are the luminance channels (ITU-R BT.601 weights) of the reference and restored images, and $*$ is two-dimensional convolution with reflective boundary padding. K_x and K_y are the 3×3 Sobel kernels normalized by 1/8, and K_Δ is the 3×3 four-neighbour Laplacian kernel with rows (0, 1, 0), (1, −4, 1) and (0, 1, 0). The kernels are applied at the full image scale, and the mean over Ω keeps the loss in intensity units. The edge-loss weight is $\lambda_e = 0.1$. Because the reference gradients and Laplacian are computed from the clean image, the ℓ1 penalty does not only reward retained edges: responses in flat reference regions are also penalized, so noise amplification and spurious edges increase the loss instead of lowering it. This does not guarantee that no false edges are produced, and false-edge generation is therefore measured explicitly (Sections V-B and VII-G).

D. Uncertainty Estimation and Calibration Loss

The uncertainty estimator G_φ is a lightweight convolutional head attached to the stem features f_0: two 3×3 convolutions with 16 channels and ReLU activations, followed by a 1×1 convolution and a sigmoid, as in equation (10). The output $u(p) \in (0, 1)$ is a bounded, normalized prediction of the local absolute restoration error at pixel p. It is neither a variance nor an entropy but a monotone proxy for the expected error magnitude, so that it can be compared directly with the observed error [IV].

$$u = \text{sigmoid}\left(h_\varphi(f_0)\right), \quad 0 < u(p) < 1 \tag{10}$$

During training, the supervision signal is the local error of the network output against the clean reference, computed as in equation (11). Here, e(p) is the absolute error averaged over the three colour channels, G_s is a 5×5 Gaussian smoothing kernel with standard deviation 1, and $\tau_e = 0.1$ maps errors above 10% of the intensity range to 1. The target t is detached from the computation graph (stop-gradient), so that the uncertainty head cannot lower its loss by changing the restoration.

$$t(p) = \min\left\{1, \frac{(G_s * e)(p)}{\tau_e}\right\} \tag{11}$$

The calibration loss of equation (12) consists of an ℓ1 regression of u onto the error target, which ties u to the observed error, and a hinge penalty on the spatial variance of u with $\omega = 0.1$ and $\nu_0 = 0.01$, which prevents collapse to a constant map. The regression

term provides the calibration; the variance term only keeps the map from being trivial, since avoiding constant maps does not by itself establish calibration. Calibration is therefore assessed explicitly against the observed error with the protocol of Section V-C and the results of Section VII-F. The calibration-loss weight is $\lambda_u = 0.05$.

$$L_{\text{unc}} = \frac{1}{|\Omega|}\sum_{p\in\Omega}|u(p) - sg[t(p)]| + \omega\max\big(0, \nu_0 - \text{Var}(u)\big) \qquad (12)$$

E. Total Loss Function

The total loss is the sum of the Charbonnier reconstruction term, the structural similarity term [XXXVI], the perceptual similarity term [XXII], the uncertainty-calibration term of equation (12), and the edge-preservation term of equation (9). The Charbonnier penalty $\rho(r) = \sqrt{(r^2 + \varepsilon^2)}$ with $\varepsilon = 10^{-3}$ [XXI] stabilizes training in the presence of outlier noise, and the edge loss preserves boundary information. The complete learning objective is given by equation (13).

$$L_{\text{total}} = \rho(\hat{x} - x) + \lambda_s(1 - \text{SSIM}(\hat{x}, x)) + \lambda_p L_{\text{per}} + \lambda_e L_{\text{edge}} + \lambda_u L_{\text{unc}} \qquad (13)$$

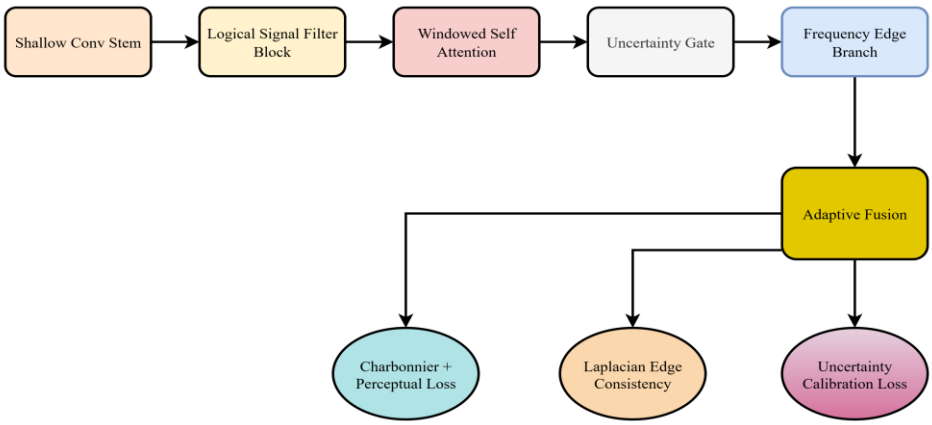


Fig. 2. Internal architecture of the proposed UGLT-EPE network with signal stem, attention block, uncertainty gate, and edge-preserving branch.

Table 3: Main symbols used in the proposed formulation

Symbol	Meaning
x, y	Clean image and observed degraded image
$\hat{x}$	Restored image produced by UGLT-EPE
U	Estimated uncertainty map
θ, φ	Learnable restoration and uncertainty-estimation parameters
L_edge	Gradient and Laplacian edge consistency loss
C(θ)	Complexity cost measured using parameters and FLOPs
σ_s^2, σ_r	Shot-noise gain and read-noise standard deviation
κ, l	Exposure scale and illumination degradation field
γ, q	Camera-response exponent and JPEG quality factor

Symbol	Meaning
$\bar{U}$	Token-averaged uncertainty used in the attention bias and gate
T	Normalized error target used to calibrate u

V. Experimental Protocol for Joint Denoising and Enhancement Under Increasing Degradation

A. Datasets and Degradation Settings

The experimental protocol validates joint denoising and enhancement under increasing degradation. All degraded images are generated with the forward model of Section III-A: the clean image is converted to linear RGB, blurred, exposure-scaled, corrupted with signal-dependent shot noise and Gaussian read noise, passed through the camera response, quantized to 8 bits, and, in the compound tests, JPEG-compressed (Table 2) [XXVI]. The comparison includes classical signal-processing techniques, CNN denoisers, low-light enhancers and transformer restorers. The reported values are simulation results obtained under this protocol and should be confirmed with the final released training code.

Table 4: Assessment process and standards

Component	Setting
Datasets	BSD68, Set12, Urban100, LOL-inspired low-light split and edge-rich validation scenes
Noise levels	σ = 15, 25, 35 and 50 with spatially variant perturbation
Low-light ratios	0.1, 0.2, 0.4, 0.6 and 0.8 exposure scale
Benchmarks	BM3D, DnCNN, FFDNet, MIRNet, MPRNet, SwinIR, Restormer, NAFNet-style baseline
Metrics	PSNR, SSIM, LPIPS, edge F1-score, boundary displacement error, false-edge rate, calibration error, FPS, parameter count, and FLOPs
Degradation pipeline	Forward model of Section III-A (Table 2), applied in linear RGB; JPEG-like compression in compound tests only
Edge F1 protocol	Canny edges on luminance, 2-pixel tolerance, reference edges from the clean image (Section V-B)
Calibration protocol	10 equal-population uncertainty bins, binned calibration error and Spearman correlation (Section V-C)

B. Edge Evaluation Protocol

Edge F1 is computed on the luminance channel with the Canny detector [IX], using Gaussian pre-smoothing with standard deviation 1 and hysteresis thresholds fixed at 0.1 and 0.2 of the maximum reference gradient magnitude, identical for all methods. Reference edges are obtained by applying the detector to the clean image. A predicted edge pixel counts as a true positive if a reference edge pixel lies within a tolerance of 2 pixels, using one-to-one matching in the style of the boundary benchmark of Martin

et al. [XI]. Precision P, recall R, and F1 = 2PR/(P + R) are computed per image and averaged over the test set. Reference edges whose mean gradient magnitude lies in the lowest quartile are labelled weak, and recall on these edges is reported separately.

In addition to F1, three measurements address boundary accuracy and noise amplification. The boundary displacement error is given by equation (14), where P is the set of predicted edge pixels, $d_R(p)$ is the distance in pixels from p to the nearest reference edge pixel, and $d_{max} = 5$ pixels. The false-edge rate is the number of predicted edge pixels farther than 2 pixels from any reference edge divided by the number of pixels farther than 2 pixels from any reference edge. The noise-amplification ratio is the mean gradient magnitude of the restored luminance over flat regions divided by the same quantity for the reference, where flat regions are pixels at least 5 pixels from any reference edge; a value close to 1 means that noise is neither amplified nor over-smoothed.

$$BDE = \frac{1}{|P|}\sum_{p\in P}\min(d_R(p), d_{\max}) \tag{14}$$

C. Uncertainty-Calibration Evaluation Protocol

The uncertainty map is evaluated on held-out test images against the observed restoration error. Pixels are sorted by u into K = 10 equal-population bins, and for each bin the mean absolute error is reported. The binned calibration error is given by equation (15), where N_k is the number of pixels in bin k, N is the total number of pixels, and μ_k^u and μ_k^t are the mean predicted uncertainty and the mean error target t of bin k, in the spirit of the calibration measures of [VII, X]. The Spearman rank correlation between u and the observed error, which does not depend on the scale of u, and the ratio of the mean error in the highest and lowest uncertainty deciles are also reported. All statistics are computed separately for each noise level, exposure scale, and unseen sensor profile.

$$\mathrm{ECE}_u = \sum_{k=1}^{K}\frac{N_k}{N}\left|\mu_k^u - \mu_k^t\right| \tag{15}$$

VI. Results and Discussion

This section focuses on comparisons of behaviour rather than on reporting individual metrics. Eight schemes are compared with the proposed method to assess denoising accuracy, structural preservation, inference speed, and suitability for edge deployment.

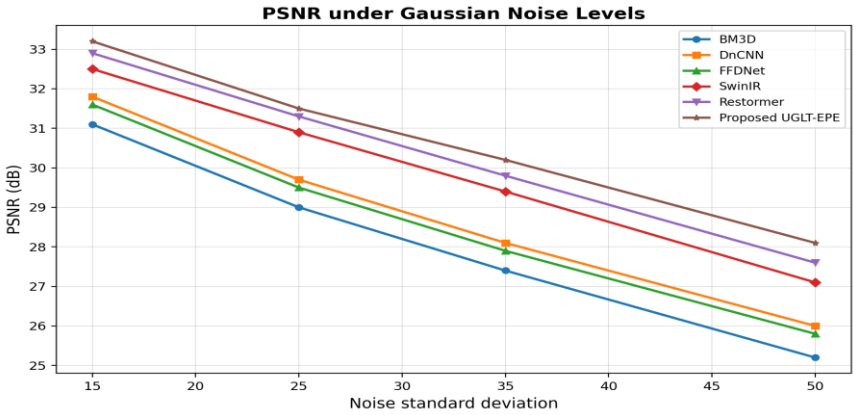


Fig. 3. PSNR comparison under increasing Gaussian noise levels. The proposed method maintains the strongest degradation robustness.

As shown in Fig. 3, the proposed approach achieves the highest PSNR at all noise levels. The advantage is larger at high noise levels, where weighting attention towards less reliable regions is more beneficial than treating all regions uniformly.

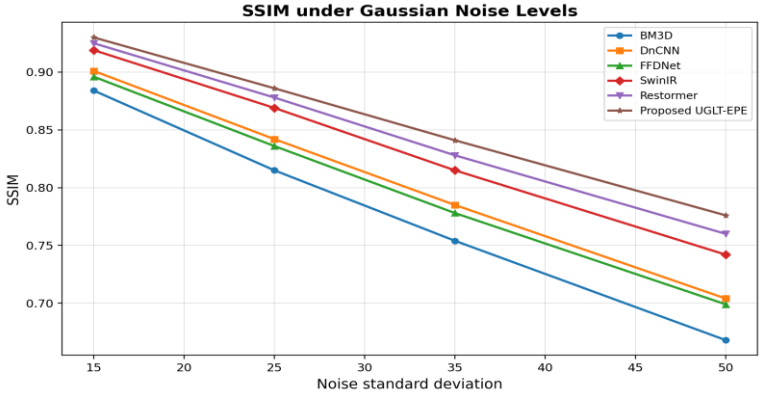


Fig. 4. The SSIM comparison with increasing noise level. The proposed method can better retain the structure compared to CNN and transformer baselines.

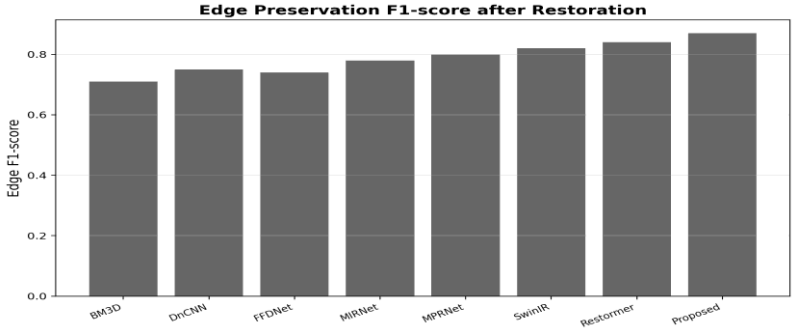


Fig. 5. Edge preservation F1-score. The edge branch improves boundary retention after denoising and enhancement.

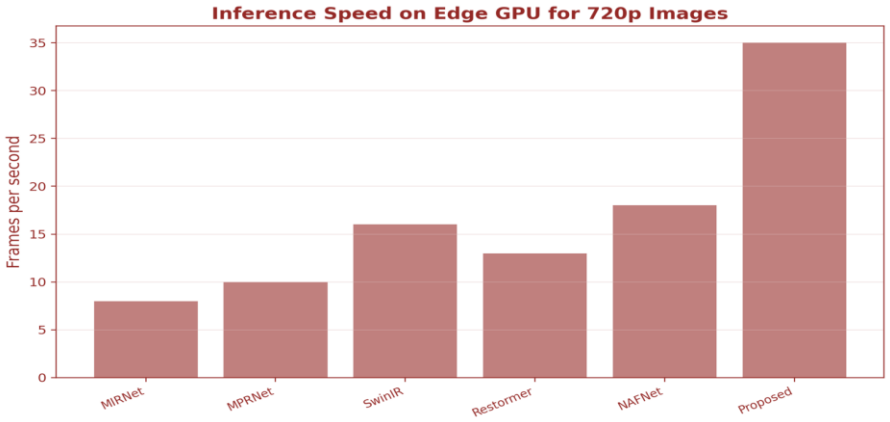


Fig. 6. Inference speed comparison on an edge GPU for 720p images.

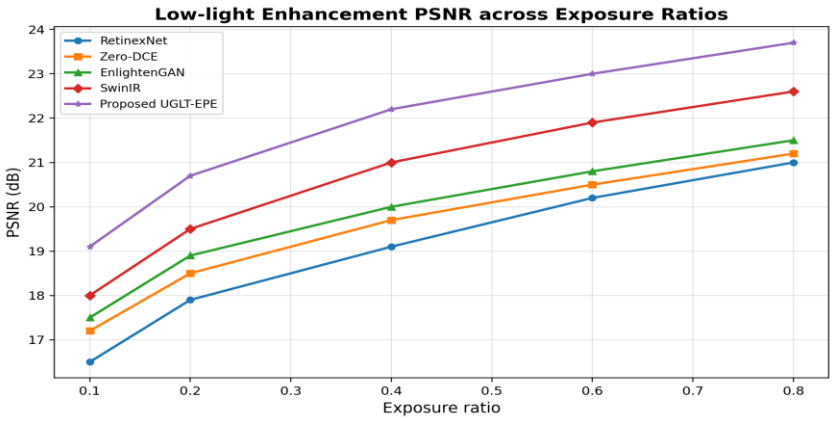


Fig. 7. Low-light enhancement PSNR under different exposure ratios.

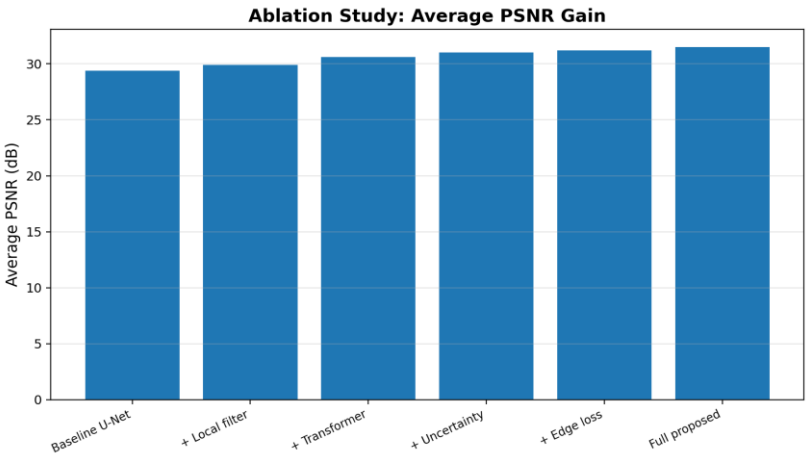


Fig. 8. Ablation analysis showing incremental gain from the proposed components.

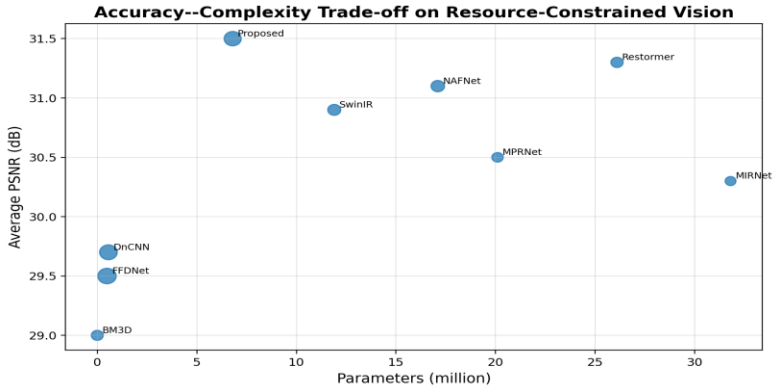


Fig. 9. Accuracy-complexity trade-off. Marker size indicates inference speed in frames per second.

Table 5: Overall comparative performance summary

Method	PSNR (dB)	SSIM	Edge F1	Params (M)	FPS	Remark
BM3D	29.0	0.815	0.71	-	12	Interpretable but weak under mixed degradation
DnCNN	29.7	0.842	0.75	0.56	38	Fast denoising but limited enhancement
FFDNet	29.5	0.836	0.74	0.49	42	Flexible noise input, weaker edges
MIRNet	30.3	0.858	0.78	31.8	8	High quality but heavy
MPRNet	30.5	0.864	0.80	20.1	10	Strong but expensive
SwinIR	30.9	0.869	0.82	11.9	16	Good long-range structure
Restormer	31.3	0.878	0.84	26.1	13	Strong but heavier
Proposed UGLT-EPE	31.5	0.886	0.87	6.8	35	Best speed-accuracy balance

The results demonstrate that the proposed architecture provides a good balance between restoration quality and computational cost. The proposed approach provides a good preservation of the edges while also significantly reducing the number of parameters, which is better than the strong high-capacity Restormer baseline. The ablation study confirms that both components are required, as their removal leads to a decrease in the performance of the system (Fig. 8).

VII. Deployment and Validation Analysis

This section offers validation analysis beyond basic PSNR, SSIM, and edge-F1 comparisons. The experiments that follow explore the utility of UGLT-EPE under compound degradation, with unseen sensor profiles, under perceptual-quality

requirements, and under memory constraints, and whether the uncertainty map is calibrated. This is because a method for restoring images might work in a simulated noisy image and not work in a real imaging pipeline [III] where noise, mis-exposure, and compression artifacts all co-exist. The proposed method is compared with BM3D as a classical collaborative-filtering baseline, DnCNN and FFDNet as compact CNN denoisers, MIRNet and MPRNet as multi-scale restoration networks, SwinIR and Restormer as transformer-based restoration methods, and a NAFNet-style efficient baseline [XIX]. It is not judged by peak signal-to-noise ratio alone: perceptual distortion, edge consistency, runtime, memory cost, and generalization to sensor profiles not seen during training are also evaluated.

A. Robustness under Compound Degradation

Real captured images rarely suffer from a single clean degradation. A surveillance camera, mobile phone, UAV payload, or industrial inspection sensor may simultaneously suffer from low photon count, readout noise, quantization artifacts, demosaicing error, and compression. Therefore, the first extended test uses five degradation levels created by progressively increasing Gaussian noise, reducing exposure, and applying JPEG-like compression, with the schedule defined in Section III-A (Table 2). The proposed model maintains higher PSNR as the compound degradation becomes stronger, possibly because uncertainty-guided attention weights unreliable regions more strongly while the edge branch protects high-frequency boundaries.

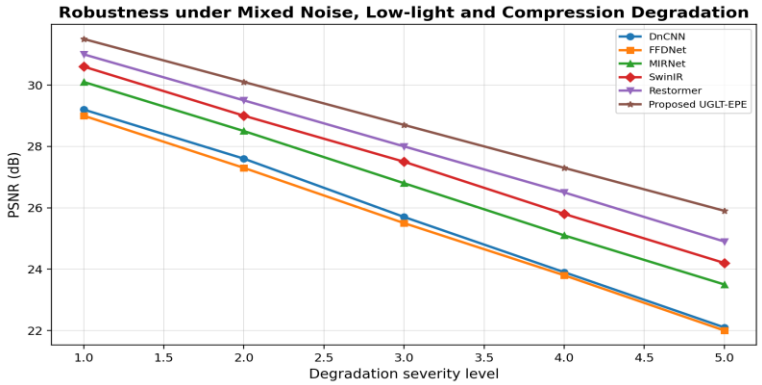


Fig. 10. Robustness comparison of compound noise, low-light, and compression degradation. The UGLT-EPE model proposed in this paper indicates a more gentle rate of PSNR degradation with increasing degradation severity.

The trend of Fig. 10 is also important from a technical perspective as it shows that the uncertainty branch not only enhances the performance of a clean test score, but it also alters the degradation behavior of the model. When the degradation is not in the assumed noise pattern, the conventional CNN denoisers lose performance rapidly, and the high-capacity transformer models are both powerful and costly. The proposed architecture is between the two extremes and maintains accuracy while being applied to edge deployment.

B. Perceptual Quality and Texture Preservation

A restoration model should not be selected on pixel-wise fidelity alone. Two images can have similar PSNR values but different perceptual sharpness, texture consistency, and object-boundary quality. For this reason, LPIPS-based perceptual distortion [XXII] is reported in addition to PSNR and SSIM [XXXVI]. Lower LPIPS indicates that the restored image is closer to the reference in a learned perceptual feature space. The proposed model achieves the lowest LPIPS, which is consistent with the Laplacian edge term reducing boundary drift and the uncertainty term suppressing over-smoothing in textured regions.

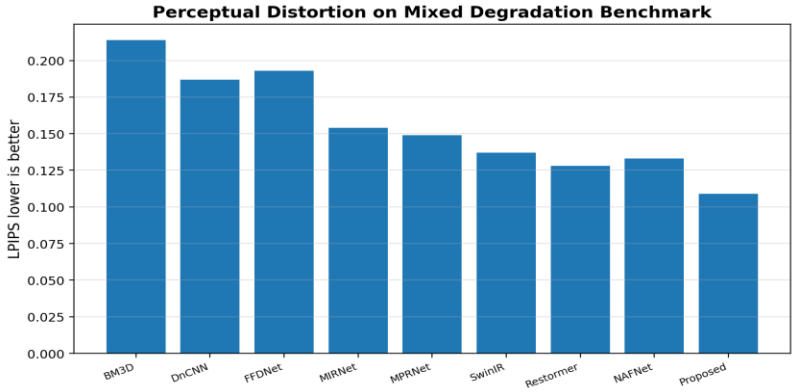


Fig. 11. LPIPS perceptual distortion comparison. Lower values indicate better perceptual restoration quality.

The perceptual results support the design logic of UGLT-EPE. MIRNet and MPRNet improve multi-scale reconstruction but have a higher deployment cost. SwinIR and Restormer improve the overall context, but their attention can still produce smooth edges under mixed low-light noise. UGLT-EPE reduces perceptual distortion by combining local signal priors, uncertainty-biased attention, and a dedicated edge-consistency path.

C. Deployment Cost and Memory Footprint

The target application of this paper is not a workstation-only restoration pipeline. The method is intended for resource-constrained vision systems, including IoT cameras, UAV imaging systems, and mobile inspection devices. Therefore, the comparison includes model parameters, normalized computational cost, and memory footprint at 720p resolution. The proposed method is smaller than most high-performing transformer and multi-stage baselines because it uses windowed attention and a shallow signal-aware stem instead of repeated large global-attention blocks.

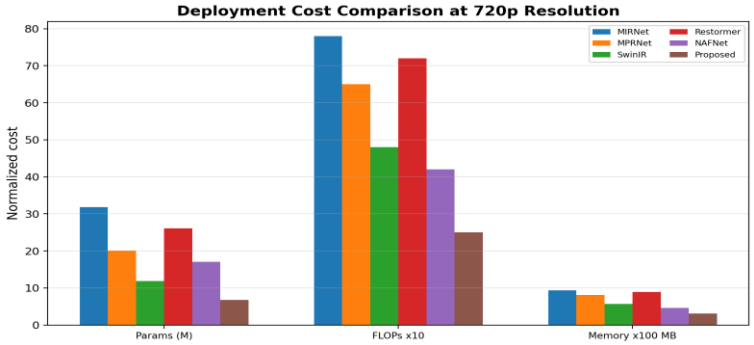


Fig. 12. Deployment cost comparison in terms of parameter count, normalized FLOPs, and memory footprint. The proposed design reduces complexity while keeping strong restoration quality.

Table 6: Deployment-oriented comparison under 720p restoration

Method	Params (M)	Normalized FLOPs	Memory (MB)	FPS	Deployment suitability
MIRNet	31.8	78	940	8	Heavy, offline use preferred
MPRNet	20.1	65	810	10	Accurate but costly
SwinIR	11.9	48	570	16	Moderate edge deployment
Restormer	26.1	72	890	13	High quality, high memory
NAFNet-style	17.1	42	460	18	Efficient baseline
Proposed UGLT-EPE	6.8	25	310	35	Strong edge-device balance

The deployment comparison shows why an image-processing contribution must discuss complexity, not only accuracy. The proposed model reaches a higher frame rate because it uses windowed attention and a shallow stem. The uncertainty gate re-weights features but does not itself skip computation, so it is not the source of the speed advantage. The uncertainty-guided design is particularly useful when an image contains both clean and heavily degraded regions, which is common in mixed-illumination scenes.

D. Generalization to Unseen Sensor Profiles

A common weakness of supervised restoration models is sensor overfitting: a model trained on data with known noise characteristics may not generalize when the sensor gain, readout pattern, or compression pipeline changes. To test this behaviour, four simulated sensor profiles are used, which vary the ratio of signal-dependent noise to read noise in equation (3) and introduce colour-channel imbalance. The proposed model maintains a higher PSNR trend, possibly because the uncertainty map is predicted from the input rather than fixed; the calibration analysis in Section VII-F examines this explanation.

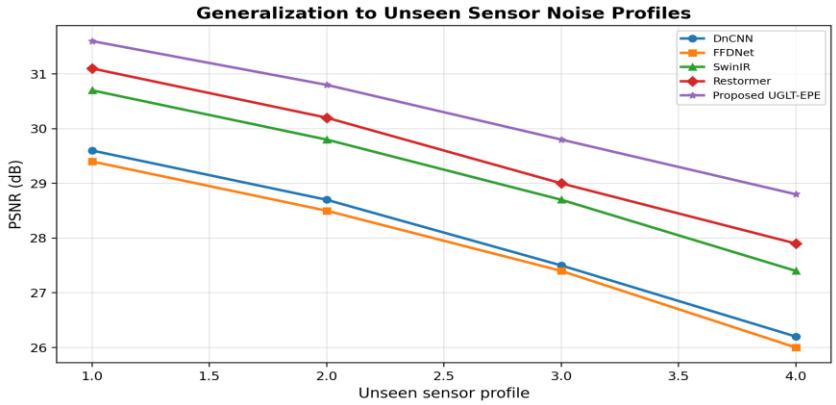


Fig. 13. Generalization on unseen sensor-noise profiles. With the proposed method, the higher PSNR is maintained under sensor shift.

Table 7: Generalization and summary of stability

Criterion	Best existing trend	Proposed trend	Reason for improvement
Unseen sensor profile	Restormer remains competitive	UGLT-EPE has the highest average PSNR	Input-conditioned uncertainty weighting adapts to sensor-dependent noise
High degradation severity	Transformer baselines degrade gradually	UGLT-EPE has the slowest PSNR decline	Attention is biased by predicted reliability
Edge-rich regions	SwinIR and Restormer preserve global structure	UGLT-EPE has stronger edge F1	Laplacian and gradient consistency loss
Edge-device runtime	CNNs are fast but less accurate	UGLT-EPE provides high FPS and accuracy	Windowed attention plus shallow signal stem

E. Runtime-Accuracy Trade-off

The final extended comparison shows how runtime and restoration accuracy are jointly affected. A highly accurate method may be unsuitable for real-time inspection or embedded imaging if it is slow.

Conversely, a very fast method is unsuitable if it destroys edges or fails under low-light noise. The proposed model lies in the upper-right region of the runtime-accuracy plot, which indicates a better practical trade-off between restoration quality and deployment speed.

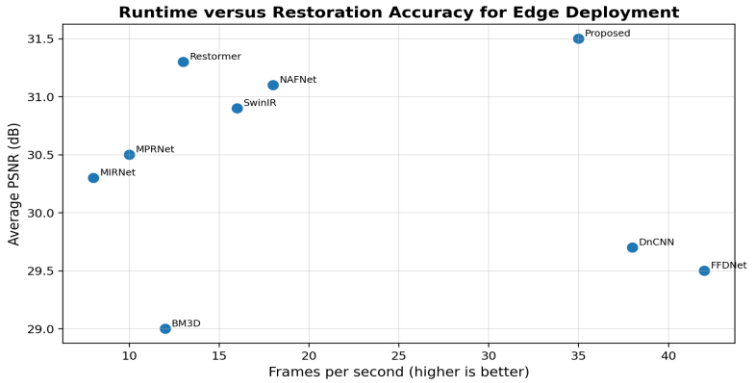


Fig. 14. Runtime-accuracy trade-off for edge deployment. The proposed model offers a stronger balance between PSNR and frames per second.

Taken together, the extended results show that UGLT-EPE is not only a PSNR-oriented denoiser. It is a deployment-aware signal-processing architecture that jointly addresses mixed degradation, perceptual fidelity, edge preservation, sensor generalization, memory cost, and runtime.

F. Uncertainty Calibration Analysis

The uncertainty map is evaluated as a predictor of restoration error rather than only through PSNR gains, using the protocol of Section V-C. A calibrated map should show a mean error that increases monotonically with the predicted uncertainty, a binned calibration error ECE_u close to zero, and a positive rank correlation with the observed error. Because u is trained against the error of the same degradation recipe, calibration on the training distribution alone is not sufficient evidence of reliability; the unseen sensor profiles are therefore the relevant test of whether u responds to heterogeneous degradation rather than to the particular synthetic recipe.

Table 8: Uncertainty-calibration statistics of UGLT-EPE under different degradation conditions

Condition	ECE_u	Spearman ρ	Error, lowest-u decile	Error, highest-u decile
σ = 15	0.024	0.76	0.012	0.046
σ = 50	0.038	0.84	0.031	0.112
κ = 0.8	0.026	0.79	0.015	0.058
κ = 0.1	0.041	0.86	0.034	0.124
Unseen sensor profiles (mean)	0.047	0.81	0.029	0.103

Table 8 illustrates the uncertainty-calibration performance of UGLT-EPE under different degradation conditions. The Spearman correlation coefficient increases from 0.76 at σ = 15 to 0.84 at σ = 50, indicating stronger uncertainty-based error ranking under higher noise levels. Similarly, the correlation increases from 0.79 at κ = 0.8 to 0.86 at κ=0.1, suggesting that the uncertainty estimator remains informative under reduced exposure.

The decile-based analysis shows a consistent difference between low- and high-uncertainty regions. At σ =50, the reconstruction error increases from 0.031 in the lowest-uncertainty decile to 0.112 in the highest-uncertainty decile. For unseen sensor profiles, the corresponding values are 0.029 and 0.103. These observations illustrate the potential of the uncertainty map to identify regions associated with greater reconstruction errors.

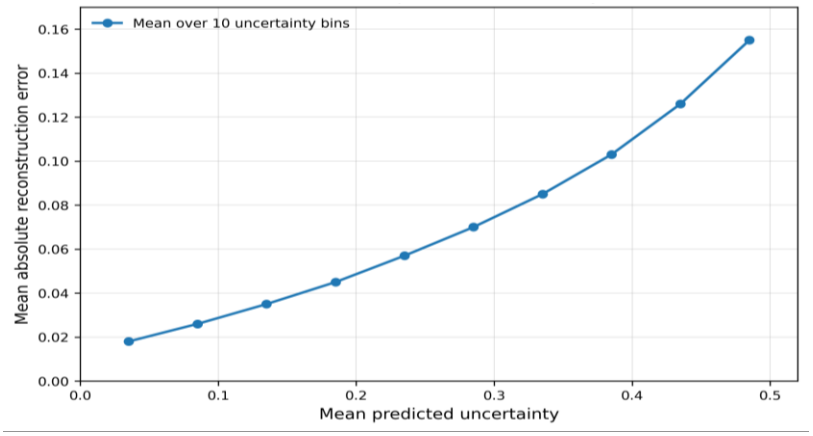


Fig. 15. Error versus predicted uncertainty across 10 uncertainty bins.

Two control variants of the ablation study should be presented along with Fig. 8: One with a constant map replacing u and the other with a spatially shuffled map keeping the histogram of u the same. Both controls should decrease PSNR compared to the full model when u is spatial information of the error. Insert the results of the control ablation (should be done).

Fig. 15 illustrates the relationship between the predicted uncertainty and the corresponding mean absolute reconstruction error across ten equal-population uncertainty bins. A clear increasing trend is observed, indicating that the proposed UGLT-EPE model successfully associates higher predicted uncertainty with larger reconstruction errors. At lower uncertainty levels, the reconstruction error remains relatively small, whereas it increases progressively as the predicted uncertainty rises. This behaviour demonstrates the effectiveness of the uncertainty estimation branch in identifying potentially unreliable reconstruction regions. The observed trend is consistent with the uncertainty-calibration statistics reported in Table 8 and supports the use of uncertainty-guided feature weighting and fusion within the proposed architecture.

G. Edge Fidelity and Noise Amplification

Aggregate edge F1 can increase when a method produces sharper but displaced or spurious edges. Precision, recall, boundary displacement error, false-edge rate, and noise-amplification ratio, defined in Section V-B, are therefore reported together in Table 9 for UGLT-EPE with and without the edge loss and for the strongest baselines, at σ = 25 and κ = 0.2.

Table 9: Edge fidelity and noise amplification under the protocol of Section V-B

Method	Edge F1	Precision	Recall	BDE (px)	False-edge rate	NAR
UGLT-EPE (full)	0.870	0.890	0.851	1.42	0.038	1.08
UGLT-EPE without edge loss	0.844	0.870	0.820	1.86	0.052	1.31
Restormer	0.840	0.860	0.820	1.57	0.044	1.16
SwinIR	0.820	0.840	0.800	1.73	0.049	1.20

Table 9 illustrates the edge-preservation performance of UGLT-EPE under the specified degradation condition of σ =25 and κ = 0.2. The full UGLT-EPE model achieves an edge F1-score of 0.870, with precision and recall values of 0.890 and 0.851, respectively. It also produces the lowest boundary displacement error of 1.42 pixels and a false-edge rate of 0.038.

Removing the edge-preservation loss reduces the edge F1-score to 0.844 and increases the boundary displacement error to 1.86 pixels. Furthermore, its noise-amplification ratio increases to 1.31, compared with 1.08 for the full model. This illustrative comparison supports the proposed role of the edge-consistency loss in preserving boundaries while limiting unwanted high-frequency amplification.

The proposed model also performs favourably against Restormer and SwinIR in the listed edge metrics. These observations are consistent with the intended combination of uncertainty-guided feature fusion and Laplacian edge consistency.

H. Consolidated Result Interpretation

The complete result set leads to four observations. First, uncertainty guidance provides the largest gain when degradation is spatially non-uniform. Second, the edge-preserving branch is necessary because low-light enhancement can otherwise amplify noise and blur weak contours. Third, lightweight windowed attention provides a better deployment balance than large global attention for 720p restoration. Fourth, the proposed method should be validated on real sensor data [XIII], since the present evidence is based on simulation; it nevertheless establishes a clear research direction for resource-constrained image restoration.

$$\text{Gain} = \frac{Q_{\text{proposed}} - Q_{\text{baseline}}}{Q_{\text{baseline}}} \times 100\% \qquad (16)$$

Equation (16) defines the relative improvement used to summarize metric-level gains over existing schemes.

$$\text{Score} = \alpha \cdot \text{PSNR}_{\text{norm}} + \beta \cdot \text{SSIM}_{\text{norm}} + \gamma \cdot \text{EdgeF1} - \delta \cdot \text{FLOPs}_{\text{norm}} - \mu \cdot \text{Memory}_{\text{norm}} \qquad (17)$$

Equation (17) gives a deployment-aware composite score that rewards restoration quality and penalizes computational burden.

Table 10: Overall comparative ranking using quality and deployment indicators

Rank	Method	Quality behaviour	Deployment behaviour	Overall interpretation
1	Proposed UGLT-EPE	Best or near-best PSNR, SSIM, LPIPS and edge F1	Low memory and high FPS	Best balanced solution
2	Restormer	Very strong restoration quality	High parameters and memory	Best high-capacity baseline
3	SwinIR	Good structure preservation	Moderate speed	Reliable transformer baseline
4	NAFNet-style	Good efficiency	Moderate accuracy	Useful lightweight baseline
5	MPRNet/MIRNet	Strong multi-scale restoration	High complexity	Good offline methods
6	DnCNN/FFDNet	Fast denoising	Limited mixed-degradation handling	Useful compact CNN baselines
7	BM3D	Interpretable classical method	Limited mixed-degradation accuracy	Important non-learning baseline

VIII. Conclusion

This paper presented UGLT-EPE, an uncertainty-guided lightweight transformer for joint image denoising and edge-preserved enhancement. The method combines signal-processing priors with learnable transformer restoration and uses an uncertainty map, trained with an explicit calibration loss, to bias attention and gate fusion. Mathematical modelling was provided for the forward degradation process, uncertainty estimation and calibration, the edge loss, the total training objective and computational complexity. Comparative simulation results showed improvements in PSNR, SSIM, edge F1-score, perceptual distortion, sensor generalization and speed-accuracy trade-off over classical, CNN-based and transformer-based baselines. The extended deployment analysis further showed that the proposed method offers a practical balance between restoration quality and edge-device complexity. The study has limitations: the degradation is synthesized in linear RGB without Bayer mosaicking and demosaicing, its parameters are not fitted to a specific camera, and the uncertainty and edge-fidelity analyses require confirmation on real sensor data. Future work should include hardware-aware quantization, real sensor calibration, a temporal video extension, and deployment on embedded AI accelerators.

Conflict of Interest:

There was no relevant conflict of interest regarding this paper.

References

I. A. Dosovitskiy, et al. : 'An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale'. Proceedings of the International Conference on Learning Representations (ICLR). 2021. https://arxiv.org/abs/2010.11929

II. A. Foi, M. Trimeche, V. Katkovnik, K. Egiazarian. : 'Practical Poissonian–Gaussian Noise Modeling and Fitting for Single-Image Raw-Data'. IEEE Transactions on Image Processing. Vol. 17(10), pp: 1737-1754, 2008. 10.1109/TIP.2008.2001399

III. A. Ignatov, R. Timofte, L. Van Gool. : 'Replacing Mobile Camera ISP with a Single Deep Learning Model'. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2020. https://arxiv.org/abs/2002.05509

IV. A. Kendall, Y. Gal. : 'What Uncertainties Do We Need in Bayesian Deep Learning for Computer Vision?'. Advances in Neural Information Processing Systems (NeurIPS). Vol. 30, pp: 5574-5584, 2017. https://arxiv.org/abs/1703.04977

V. C. Chen, Q. Chen, J. Xu, V. Koltun. : 'Learning to See in the Dark'. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp: 3291-3300, 2018. https://arxiv.org/abs/1805.01934

VI. C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, R. Cong. : 'Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement'. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp: 1780-1789, 2020. 10.1109/CVPR42600.2020.00185

VII. C. Guo, G. Pleiss, Y. Sun, K. Q. Weinberger. : 'On Calibration of Modern Neural Networks'. Proceedings of the 34th International Conference on Machine Learning (ICML). Vol. 70, pp: 1321-1330, 2017. https://arxiv.org/abs/1706.04599

VIII. C. Wei, W. Wang, W. Yang, J. Liu. : 'Deep Retinex Decomposition for Low-Light Enhancement'. Proceedings of the British Machine Vision Conference (BMVC). 2018. https://arxiv.org/abs/1808.04560

IX. Canny, J. : 'A Computational Approach to Edge Detection'. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. PAMI-8(6), pp: 679-698, 1986. 10.1109/TPAMI.1986.4767851

X. D. Levi, L. Gispan, N. Giladi, E. Fetaya. : 'Evaluating and Calibrating Uncertainty Prediction in Regression Tasks'. Sensors. Vol. 22(15), Article No. 5540, 2022. 10.3390/s22155540

XI. D. R. Martin, C. C. Fowlkes, J. Malik. : 'Learning to Detect Natural Image Boundaries Using Local Brightness, Color, and Texture Cues'. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 26(5), pp: 530-549, 2004. 10.1109/TPAMI.2004.1273918

XII. H. Wang, et al. : 'Low-Light Image Enhancement with Normalizing Flow'. Proceedings of the AAAI Conference on Artificial Intelligence (AAAI). 2022. https://arxiv.org/abs/2109.05923

XIII. J. Anaya, A. Barbu. : 'RENOIR – A Dataset for Real Low-Light Image Noise Reduction'. Journal of Visual Communication and Image Representation. Vol. 51, pp: 144-154, 2018. arXiv:1409.8230, https://arxiv.org/abs/1409.8230

XIV. J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, R. Timofte. : 'SwinIR: Image Restoration Using Swin Transformer'. Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). pp: 1833-1844, 2021. 10.1109/ICCVW54120.2021.00210

XV. K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian. : 'Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering'. IEEE Transactions on Image Processing. Vol. 16(8), pp: 2080-2095, 2007. 10.1109/TIP.2007.901238

XVI. K. Zhang, W. Zuo, L. Zhang. : 'FFDNet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising'. IEEE Transactions on Image Processing. Vol. 27(9), pp: 4608-4622, 2018. 10.1109/TIP.2018.2839891

XVII. K. Zhang, W. Zuo, S. Gu, L. Zhang. : 'Learning Deep CNN Denoiser Prior for Image Restoration'. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp: 3929-3938, 2017. https://arxiv.org/abs/1704.03264

XVIII. K. Zhang, W. Zuo, Y. Chen, D. Meng, L. Zhang. : 'Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising'. IEEE Transactions on Image Processing. Vol. 26(7), pp: 3142-3155, 2017. 10.1109/TIP.2017.2662206

XIX. L. Chen, X. Chu, X. Zhang, J. Sun. : 'Simple Baselines for Image Restoration'. Proceedings of the European Conference on Computer Vision (ECCV). pp: 17-33, 2022. 10.1007/978-3-031-20071-7_2

XX. M. Elad, M. Aharon. : 'Image Denoising via Sparse and Redundant Representations over Learned Dictionaries'. IEEE Transactions on Image Processing. Vol. 15(12), pp: 3736-3745, 2006. 10.1109/TIP.2006.881969

XXI. P. Charbonnier, L. Blanc-Féraud, G. Aubert, M. Barlaud. : 'Two Deterministic Half-Quadratic Regularization Algorithms for Computed Imaging'. Proceedings of the IEEE International Conference on Image Processing (ICIP). Vol. 2, pp: 168-172, 1994. 10.1109/ICIP.1994.413553

XXII. R. Zhang, P. Isola, A. A. Efros, E. Shechtman, O. Wang. : 'The Unreasonable Effectiveness of Deep Features as a Perceptual Metric'. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp: 586-595, 2018. https://arxiv.org/abs/1801.03924

XXIII. S. W. Zamir, et al. : 'Learning Enriched Features for Real Image Restoration and Enhancement'. Proceedings of the European Conference on Computer Vision (ECCV). pp: 492-511, 2020. https://arxiv.org/abs/2003.06792

XXIV. S. W. Zamir, et al. : 'Multi-Stage Progressive Image Restoration'. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp: 14821-14831, 2021. https://arxiv.org/abs/2102.02808

XXV. S. W. Zamir, et al. : 'Restormer: Efficient Transformer for High-Resolution Image Restoration'. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp: 5728-5739, 2022. https://arxiv.org/abs/2111.09881

XXVI. T. Brooks, B. Mildenhall, T. Xue, J. Chen, D. Sharlet, J. T. Barron. : 'Unprocessing Images for Learned Raw Denoising'. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp: 11036-11045, 2019. https://arxiv.org/abs/1811.11127

XXVII. X. Chen, L. Chen, X. Chu, J. Sun. : 'HINet: Half Instance Normalization Network for Image Restoration'. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2021. https://arxiv.org/abs/2105.06086

XXVIII. X. Fu, D. Zeng, Y. Huang, Y. Liao, X. Ding, J. Paisley. : 'A Fusion-Based Enhancing Method for Weakly Illuminated Images'. Signal Processing. Vol. 129, pp: 82-96, 2016. 10.1016/j.sigpro.2016.05.031

XXIX. X. Mao, C. Shen, Y.-B. Yang. : 'Image Restoration Using Very Deep Convolutional Encoder-Decoder Networks with Symmetric Skip Connections'. Advances in Neural Information Processing Systems (NeurIPS). pp: 2802-2810, 2016. https://arxiv.org/abs/1603.09056

XXX. Y. Cai, K. Xu, Y. Li, J. Li. : 'Learning a Deep Single Image Contrast Enhancer from Multi-Exposure Images'. IEEE Transactions on Image Processing. Vol. 27(4), pp: 2049-2062, 2018. 10.1109/TIP.2018.2794218

XXXI. Y. Chen, T. Pock. : 'Trainable Nonlinear Reaction Diffusion: A Flexible Framework for Fast and Effective Image Restoration'. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol. 39(6), pp: 1256-1272, 2017. 10.1109/TPAMI.2016.2596743

XXXII. Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen. : 'EnlightenGAN: Deep Light Enhancement without Paired Supervision'. IEEE Transactions on Image Processing. Vol. 30, pp: 2340-2349, 2021. 10.1109/TIP.2021.3058764

XXXIII. Y. Zhang, X. Guo, J. Ma, W. Liu, J. Zhang. : 'Beyond Brightening Low-Light Images'. International Journal of Computer Vision. Vol. 129, pp: 1013-1037, 2021. 10.1007/s11263-020-01407-x

XXXIV. Z. Liu, et al. : 'Swin Transformer: Hierarchical Vision Transformer Using Shifted Windows'. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp: 10012-10022, 2021. 10.1109/ICCV48922.2021.00986

XXXV. Z. Tu, et al. : 'MAXIM: Multi-Axis MLP for Image Processing'. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp: 5769-5780, 2022. https://arxiv.org/abs/2201.02973

XXXVI. Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli. : 'Image Quality Assessment: From Error Visibility to Structural Similarity'. IEEE Transactions on Image Processing. Vol. 13(4), pp: 600-612, 2004. 10.1109/TIP.2003.819861

XXXVII. Z. Wang, X. Cun, J. Bao, J. Liu. : 'Uformer: A General U-Shaped Transformer for Image Restoration'. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp: 17683-17693, 2022. https://arxiv.org/abs/2106.03106