Low-Light Image Enhancement for Edge-Based Security Surveillance in 6G-IoT Visual Systems

Vishal Krishna Singh, Niharika Anand, Krishna Sharma S, Anjali, *Member, IEEE*, Mahendra Kumar Shukla , *Member, IEEE*, Rajkumar Singh Rathore, *Senior Member, IEEE*, Weiwei Jiang, *Senior Member, IEEE*.

***Abstract*—Application areas such as real-time visual analytics over high-bandwidth 6G networks, low-power camera networks in remote or low-light environments and surveillance drones, usually operate under insufficient lighting conditions. The captured images are often of low quality, poor resolution and poor visual clarity, leading to reduced visibility, color distortion, and amplified noise. Existing methods of Low-light image enhancement suffer from low accuracy with compromised reliability, trust and fairness. Inspired by the zero-reference learning paradigm of Zero-DCE++, this work aims to investigate the impact of data pre-processing and augmentation strategies for improving the performance of real-time, mission-critical security systems where low-light surveillance images are used for critical decision making. The proposed method uses FFDNet for denoising, exposure fusion for illumination improvement and data augmentation for bias mitigation and performance optimization through diverse training samples. The method is curated for edge deployment on constrained IoT hardware, with low latency and energy efficient usage in 6G-IoT visual systems. The proposed model is aimed at performance improvement on trust driven visual improvements, reduced distributional bias, and deployment fairness across diverse lighting conditions and scenarios. Comparative analysis demonstrates that with the help of zero-reference deep curve estimation, the proposed, *DA-Zero-DCE++*, pipeline achieves improved performance as compared to state-of-the-art low-light image enhancement methods. Our best configuration, which combines exposure fusion-based augmentation and mild denoising using FFDNet, achieves an average PSNR of $15.34\ dB$, SSIM of 0.4869, and MAE of 40.87 on the *SICE dataset* at 1200×900 resolution. For high-level vision applications such as real-time visual analytics over high-bandwidth 6G networks, low-power camera networks in remote or low-light environments, the performance is further validated on DarkFace dataset where high average precision at intersection over union of 0.5 is achieved.**

***Index Terms*—6G Internet of Things, Low-Light Image Enhancement, Security Surveillance, Trustworthy AI, Bias Mitigation, Deep Curve Estimation, Zero-Reference Learning.**

I. Introduction

VISUAL data plays a key role in smart cities with varying applications such as autonomous vehicles, surveillance systems, autonomous security drones, real-time visual analytics over high-bandwidth 6G networks, low-power camera networks in remote or low-light environments, disaster response and emergency monitoring. In real world scenarios such as real-time visual analytics over high-bandwidth 6G networks and low-power camera networks in remote or low-light environments, the captured images are often of low quality, poor resolution and poor visual clarity, leading to reduced visibility, color distortion, and amplified noise. Research in the domain of Low-light Image Enhancement (LLIE) aims to address these issues with the aim of achieving improved accuracy with reliable and trustworthy data-driven decision making.

LLIE has evolved significantly in terms of optimized performance and improved trust with advanced DL methods. Although early LLIE methods relied on hand-crafted techniques such as *histogram equalization* and *retinex-based* algorithms [8]. While these methods improved global contrast, they often amplified noise and failed to preserve fine details in extreme low-light conditions. Modern methods like FFENet [14] work similar to the traditional methods, processing the information at high and low information separately. Yet, these methods suffer from limitations like color distortions. Recent works leverage deep networks to learn end-to-end mappings between low-light and normal-light images. For example the RetinexNet [8] uses decomposed images into reflectance and illumination maps but suffers from noise amplification. Subsequent works like MBPNet [15] used multi-branch and progressive networks to handle diverse degradation patterns. Similarly, the broader field of image enhancement has leveraged fusion techniques to tackle other challenges; for instance, multi-focus image fusion is used in surveillance to combine images with different focal depths into a single, sharp picture [16] and multi-modality medical image fusion merges data from different sources (e.g., CT, MRI) for improved diagnostics [1]. Other notable CNN-based approaches like MBLLEN [2] have also been developed for enhancing low-light images and videos, while others like HFMNet [3] focus on hierarchically mining features, using different network layers to specifically handle illumination and edge details. Similarly, MLLEN-IC [4] employs multiscale feature extraction and an illumination constraint to prevent overexposure and color distortion. Many GAN based methods have been developed over the years, such as the EnlightenGAN [11] which uses global and local discriminators to enhance realism. More recently, works like the CycleGAN proposed in [17] have emerged as seminal contributions addressing the issue of low accuracy and degraded trust and poor

Corresponding author: Vishal Krishna Singh (email: v.k.singh@essex.ac.uk)
Vishal Krishna Singh is with School of Computer Science and Electronics Engineering, University of Essex, Colchester Campus, United Kingdom.
Niharika Anand and Krishna Sharma S are with the Department of Computer Science, Indian Institute of Information Technology Lucknow, India.
Anjali and Mahendra Kumar Shukla are with the Department of Information Technology, Indian Institute of Information Technology Gwalior, India.
Rajkumar Singh Rathore is with the Department of Computer Science, School of Technologies, Cardiff Metropolitan University, United Kingdom
Weiwei Jiang is with School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing, China

TABLE I
COMPARATIVE ANALYSIS OF THE PROPOSED METHOD WITH STATE-OF-THE-ART METHODS

Models	Complexity	Performance	Suited For Dataset		Sensitivity		Advantages of Proposed Approach over Existing Methods
			Large & Annotated	Scalable, Generalized & Diverse	Data Augmentation	Multi-level Denoising	
MMIF [1]	Medium	Fair	✓	✓	–	–	1- Significantly improved performance in terms of visual quality and complexity, through data augmentation, illumination correction, and noise suppression.
MBLLEN [2]	Medium	Fair	✓	✓	✓	–	
HFMNet [3]	Medium	Good	✓	–	✓	–	
MLLEN-IC [4]	High	Fair	✓	✓	✓	–	2 - Scalable, robust and generalized architecture across heterogeneous lighting scenarios through exposure fusion.
SRIE [5]	High	Fair	✓	–	✓	–	
LIME [6]	High	Good	✓	–	✓	–	3- Light-head architecture validated on diverse training samples and multiple exposure levels.
Li *et al*. [7]	Medium	Good	–	✓	✓	–	
RetinexNet [8]	Medium	Good	–	✓	–	–	4- Pixel-wise enhancement curve estimation to reduce paired reference image dependency.
LightenNet [9]	Medium	Good	✓	✓	–	–	
Wang *et al*. [10]	High	Good	–	✓	–	–	6- Multilevel denoising.
EnlightenGAN [11]	Medium	Good	✓	–	✓	–	5- Lightweight and computationally efficient architecture; well suited for real-time deployment on resource-constrained edge devices.
Zero-DCE [12]	High	Good	–	✓	✓	–	
Zero-DCE++ [13]	Medium	Good	✓	✓	✓	–	
Proposed *DA-Zero-DCE++*	**Medium**	**Excellent**	✓	✓	✓	✓	7- Practically beneficial for high-level vision applications.

image resolution for improving the performance of real-time, mission-critical security systems where low-light surveillance images are used for critical decision making. In one of the important contributions, Zero-DCE [12] introduced a zero-reference framework using deep curve estimation, eliminating the need for paired data. It's lightweight design (DCE-Net) inspired efficient variants like Zero-DCE++ [13] and more recent frameworks like Zero-DCE xt [18] which are aimed to be computationally cheap. Other important methods include CPGA-Net [19], which uses gamma correction in a lightweight architecture and IRNet [20], which improved zero-shot Retinex decomposition using dual networks (Decom-Net and Enhance-Net), achieving state-of-the-art metrics on benchmark datasets. The Table I presents a comprehensive comparison of select state-of-the-art methods to highlight the major challenges in existing scenario.

However, it is empirical to consider that existing LLIE methods, including histogram equalization and Retinex-based algorithms [21], have shown limited ability to handle the complex degradations present in real-world low-light scenarios. In recent years, deep learning (DL) based approaches have achieved remarkable progress. Notable methods such as RetinexNet [8], EnlightenGAN [11], and Zero-DCE++ [13] have demonstrated superior performance by learning powerful mappings from low-light to well-exposed images. Zero-DCE++ is one of the most efficient zero-reference frameworks that gives competitive results on LLIE tasks while also being computationally efficient. Despite these advances, many DL-based LLIE methods require large-scale, well-annotated datasets. Additionally, the performance of these methods are too sensitive to the choice of training data and pre-processing steps. To be specific, the roles of data augmentation, and denoising in the context of LLIE pipelines are not yet fully understood. Furthermore, the generalization ability of these models to diverse real-world scenarios and their impact on high-level vision tasks, such as face detection in extremely dark images, warrant further investigation.

To address these issues, this work proposes a data augmentation driven framework, *DA-Zero-DCE++*, which builds upon the *Zero-DCE++* framework. The idea is to improve visual quality through a combination of data augmentation, illumination correction, and noise suppression techniques, all within a zero-reference learning framework designed for edge-based surveillance systems in 6G-IoT environments. The proposed *DA-Zero-DCE++* integrates exposure fusion as a data augmentation strategy, generating diverse training samples by fusing multiple exposure levels from the same scene to simulate varied low-light conditions. This helps improve the model's robustness and generalization across heterogeneous lighting scenarios. In parallel, FFDNet is incorporated as a pre-processing step for image denoising, with varying levels of noise applied during training to study its impact on the stability and reliability of the enhancement process. These components are seamlessly combined with the Zero-DCE++ network, which estimates pixel-wise enhancement curves without requiring paired reference images. The proposed method is designed to be lightweight and computationally efficient, enabling real-time deployment on resource-constrained edge devices in a smart city. The major contributions of this work are as below:

1) A novel approach to use exposure fusion to generate diverse training samples from multi-exposure sequences from the SICE dataset [22]. This improves the model's ability to generalize across various low-light scenarios.
2) Using FFDNet, a comprehensive framework is proposed for denoising at varying noise levels to analyze its impact on both image quality and downstream vision tasks.
3) Combining exposure fusion-based augmentation and mild denoising, the proposed *DA-Zero-DCE++* achieves

significant improvement over state-of-the-art methods especially with respect to perceptual quality, while remaining lightweight.
4) We quantitatively assess how our enhancement pipeline improves face detection performance on the DarkFace dataset [23].
5) We present a use case model to highlight the practical benefits of our proposed pipeline for high-level vision applications such as real-time visual analytics over high-bandwidth 6G networks, low-power camera networks in remote or low-light environments.

The organization of the paper is as follows: Section II presents the details of the system model followed by the Section III where the proposed *DA-Zero-DCE++* is explained in detail. This is followed by the experimental setup, data set description, scenarios and outputs in Section IV, the benchmark evaluation parameters on a second dataset in Section V and concluding comments in Section VI.

II. System Model

A model scenario of a fully connected *6G-enabled smart city* is considered to demonstrate the operational flow of the proposed *DA-Zero-DCE++* framework. The system model, illustrated in Fig. 1, is structured into three interconnected layers: (i) the Perception Layer, (ii) the Edge Processing Layer, and (iii) the Cloud and Control Layer. Each layer is designed to ensure low-latency data transfer, adaptive enhancement, and intelligent decision-making within a high-bandwidth 6G-IoT visual ecosystem. The model is aimed at establishing an end-to-end intelligent surveillance pipeline. The Perception Layer performs sensing, the Edge Processing Layer executes enhancement and initial inference using the proposed *DA-Zero-DCE++* framework, and the Cloud Layer coordinates large-scale decision support. Through the combination of denoising, exposure fusion, and zero-reference enhancement, the system ensures trust-driven, bias-mitigated, and energy-efficient visual enhancement suitable for deployment in mission-critical 6G-IoT environments.

The model city has surveillance cameras to monitor public spaces like parks, streets, parking lots, banks etc. and often operate under varying lighting conditions (Fig. 1). The data acquisition considers adverse scenarios where low quality images are captured due to poor lighting conditions, adverse weather conditions and possible obstructions in line of sight. The following assumptions are considered for the captured images:

1) Footage Acquisition: The surveillance cameras continuously record and store video streams in varying lighting conditions, including low-light conditions.
2) Incident Detection and Footage Extraction: In case of an event being reported, relevant video segments are extracted from storage for further analysis in real-time or almost real-time.
3) Frame Enhancement: The extracted low-light frames are processed using the proposed *DA-Zero-DCE++* for inference.

A. Perception Layer

The Perception Layer consists of distributed low-power imaging sensors and surveillance cameras deployed across urban and critical environments such as streets, parking areas, parks, and transport hubs. These cameras continuously capture visual data under diverse illumination and environmental conditions. In many real-world cases, the captured frames suffer from degradation due to poor lighting, motion blur, or atmospheric distortions such as fog and rain. The data acquisition process therefore considers a set of low-visibility images I_{low} that require enhancement prior to higher-level analytics.

B. Edge Processing Layer

Each edge node is equipped with constrained IoT hardware capable of executing lightweight AI inference with limited computational and energy resources. The proposed *DA-Zero-DCE++* model is deployed at this layer to enhance low-light images in real time. The main components being; (i) Denoising: The raw image is first processed using FFDNet to suppress noise while retaining structural details, (ii) Exposure Fusion: An illumination correction module based on exposure fusion (iii)Zero-Reference Enhancement: The refined image is passed through a zero-reference curve estimation network that adaptively adjusts illumination and color consistency without relying on paired ground-truth data.

C. Cloud and Control Layer

The Cloud and Control Layer integrates the outputs from multiple edge nodes across the 6G-IoT infrastructure. Enhanced images and extracted features are transmitted via ultra-reliable low-latency communication channels to centralized servers for large-scale analytics, event correlation, and long-term storage. The cloud layer supports global decision-making, and adaptive model updates.

III. Proposed Methodology

The proposed *DA-Zero-DCE++* builds upon the *Zero-DCE++* framework with the aim to improve visual quality through a combination of data augmentation, illumination correction, and noise suppression techniques, all within a zero-reference learning framework designed for edge-based surveillance systems in 6G-IoT environments. The proposed *DA-Zero-DCE++* integrates exposure fusion as a data augmentation strategy, generating diverse training samples by fusing multiple exposure levels from the same scene to simulate varied low-light conditions. This helps improve the model's robustness and generalization across heterogeneous lighting scenarios. In parallel, FFDNet is incorporated as a pre-processing step for image denoising, with varying levels of noise applied during training to study its impact on the stability and reliability of the enhancement process.

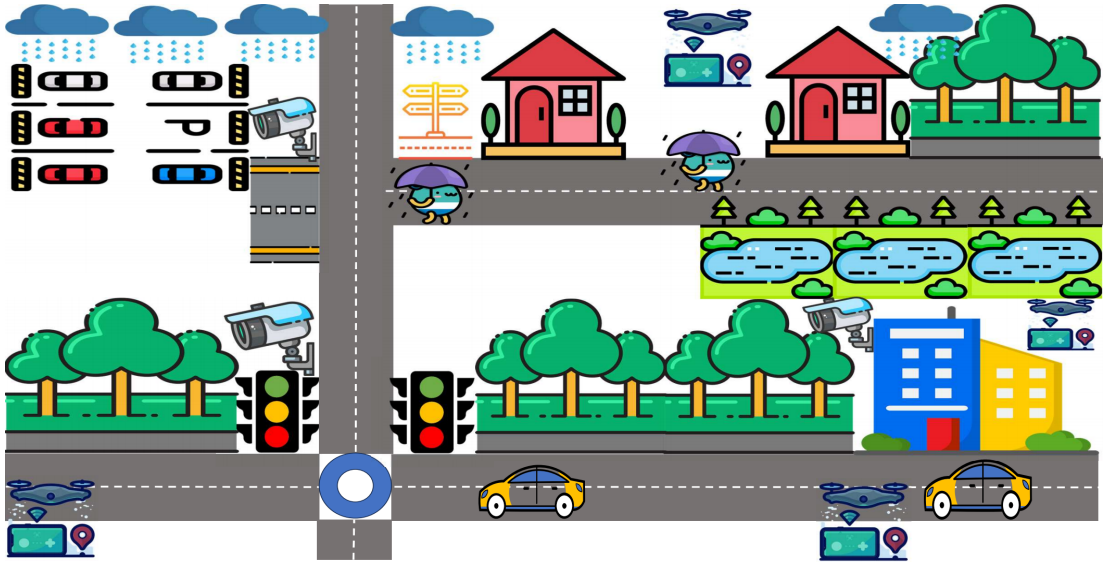


Fig. 1. System Model for the Proposed *DA-Zero-DCE++*

A. Exposure Fusion

Exposure fusion is achieved by combining multiple images of the same scene taken at different exposure levels. We use exposure fusion to create images for training and testing in approach 1 and to augment the dataset in approaches 2 and 3 and is described in Section IV. The idea is to take advantage of well-exposed regions from each image to produce a final output with improved detail and balanced illumination, making it particularly suitable for enhancing low-light images. The fusion allows bias mitigation and optimizes the learning process by providing a balanced set of training images with varying features and properties. The Fig. 4 shows the step wise execution of the exposure fusion process over the selected images.

The exposure fusion approach used in our methodology is grounded in the technique proposed in [24]. Building on the existing work, we consider three measures of quality:

1) **Contrast:** For each input image, the absolute value of the Laplacian filter applied to the grayscale version provides a per-pixel measure of local contrast. High-frequency regions like edges receive higher weights.
2) **Saturation:** The saturation measure is computed as a standard deviation of R, G and B channel, at each pixel.
3) **Well-exposedness:** This measure prefers pixels that are neither under-exposed nor over-exposed. It is quantified by applying a Gaussian function peaking at the ideal exposure value where the pixel intensities are near 128 for 8-bit images.

These three measures are combined into a scalar weight map using multiplication and is normalized across all images so the sum of weights at each pixel location is one. The weighted images are blended using a Laplacian pyramidal decomposition to avoid seam artifacts and ensure smooth blending.

For approach 1, we consider only the *darkest* to the *mid-exposure* images in an exposure sequence to create a new image. For example, if a multi-exposure sequence consists of 5 images, we take the first 3 images to create our training image. We do this for all sequences in the *SICE* dataset to create a new training set. Exposure fusion is used to augment the Part 1 of SICE dataset for training in approaches 2 and 3. As illustrated in Fig 4, the three darkest images from a given sequence are fused. This fused image is then added to the original set, thereby augmenting the training data with a high-quality composite.

B. Denoising

Low-light images typically suffer from significant noise due to high sensor gain and low photon counts under poor illumination. To mitigate this, the proposed framework integrates FFDNet [25] as a front-end denoising module within the Zero-DCE++ pipeline. FFDNet provides a flexible and efficient CNN-based denoising mechanism capable of handling a broad range of spatially invariant or variant noise levels through the use of a tunable noise-level map. In the integration, the input low-light image is first subjected to reversible downsampling into sub-images to improve computational efficiency and enlarge the receptive field. A corresponding noise-level map, denoted as M, is concatenated with the downsampled tensor to guide adaptive noise suppression while maintaining structural detail. The denoised output is then forwarded to the Zero-DCE++ enhancement module, ensuring that the illumination correction operates on cleaner image data with reduced noise amplification. During training, Additive White Gaussian noise (AWGN) with varying noise levels (σ) is applied to the input images to simulate various degradation levels, following the FFDNet training strategy. The network learns to map (I_{noisy}, M) pairs to clean images by minimizing the reconstruction loss. Orthogonal initialization of convolutional filters is adopted to stabilize training and preserve the role of M in balancing noise reduction and detail retention. The noise map M acts as a dynamic control parameter, allowing the denoising

strength to adapt across regions with different illumination conditions.

C. DA-Zero-DCE++

The proposed *DA-Zero-DCE++* is a fast and effective framework which can enhance low-light image without the need for any reference image. It works by estimating a set of Light-Enhancement (LE) curves and progressively enhancing the image. The existing model of Zero-DCE++ [13] enhances the DCE-Net which learns the mapping between the input image and the best-fitting curve parameter maps. The input is a low-light image and the output is a set of pixel-wise curve parameters. The curve parameters are generated solely based on the input image, allowing for fully automated adjustment. The design of this curve follows three main objectives:

1) All enhanced pixel values remain within the normalized range $[0, 1]$, preventing information loss due to overflow or truncation.
2) The curve is monotonically increasing in order to preserve the contrast between neighboring pixels.
3) The curve is simple and differentiable, enabling efficient optimization during neural network training.

To achieve these criteria, the objective function is designed as a quadratic curve, which is expressed as:

$$LE(I(\mathbf{x}); \alpha) = I(\mathbf{x}) + \alpha I(\mathbf{x})(1 - I(\mathbf{x})) \tag{1}$$

Here $\mathbf{x}$ denotes pixel coordinates, $LE(I(x); \alpha)$ is the enhanced version of the given input $I(x)$ and the trainable parameter $\alpha \in [0, 1]$ adjusts the magnitude of LE-curve while controlling the exposure level as well. The curve defined in Equation 1 is more effective in adjustment when used iteratively. For n iterations, the equation 1 is rewritten as:

$$LE_n(\mathbf{x}) = LE_{n-1}(\mathbf{x}) + \alpha_n LE_{n-1}(\mathbf{x})(1 - LE_{n-1}(\mathbf{x})) \tag{2}$$

To prevent over or under enhancement of local regions, we formulate α as a pixel-wise parameter by using a parameter map $\mathcal{A}$ which is of the same dimensions as the image. The equation 2 is therefore updated to:

$$LE_n(\mathbf{x}) = LE_{n-1}(\mathbf{x}) + \mathcal{A}_n LE_{n-1}(\mathbf{x})(1 - LE_{n-1}(\mathbf{x})) \tag{3}$$

In addition to the

IV. Experimental Setup and Model Scenarios

A. Datasets

Our experiments utilize several public datasets for training, evaluation, and performance validation on downstream vision tasks.

SICE Dataset [22]: The primary dataset for training and evaluating our enhancement pipeline is the Single Image Contrast Enhancement (SICE) dataset. It contains 589 high-resolution, multi-exposure sequences, totaling 4413 images. The dataset covers a diverse range of lighting conditions, with sequences captured in both indoor and outdoor environments. Each sequence consists of between 3 to 18 images with varying exposure levels. For outdoor scenes in particular, the sequences were collected with exposure values (EV) shifted by $\pm\{0.5, 0.7, 1.0, 2.0, 3.0\}$, ensuring a broad and well-defined spectrum of under- and over-exposed conditions. The ground truth for each sequence is a single high-quality reference image, generated by fusing the multi-exposure images and selecting the best result through subjective human evaluation. For our experiments, we primarily use Part 1 of the dataset for training and Part 2 for testing. To ensure the robustness and generalization of the proposed framework, we further employ a 5-fold cross-validation strategy across the entire SICE dataset. The sequences are randomly partitioned into 5 subsets, where 4 folds are used for training and 1 fold for testing in each iteration. The reported results represent the average performance across all folds, reducing bias due to dataset partitioning. During preprocessing, all images are resized to 512×512 pixels for training, and, in specific experiments, are denoised using FFDNet to suppress sensor noise while retaining structural details. Critically, we leverage the multi-exposure sequences to create augmented training data via exposure fusion. We use both Part 1 and Part 2 to generate the training images in the experiment described in Approach 1. For Approach 2 and Approach 3 we follow the established protocol of using sequences from Part 1 of the dataset for training and Part 2 for testing.

DarkFace Dataset [26]: To validate the practical utility of our method in real-world security applications, we use the challenging DarkFace dataset. This real-world dataset consists of $6,000$ low-light images captured at night in various unconstrained scenes, including streets, parks, and overpasses. All images in this set are annotated with bounding boxes for human faces. The dataset also provides thousands of unlabeled images and a smaller set of paired low-light/normal-light images, though these were not used in our evaluation. We benchmark our method's impact on the downstream task of face detection by evaluating it against the hold-out test set, quantitatively measuring how our enhancement pipeline improves the Average Precision (AP) of a face detector.

Custom Dataset from Reddit Images: To validate our model's generalization on challenging, real-world data, we curate a custom 'in-the-wild' dataset. It consists of image from the Reddit. We perform targeted searches for keywords indicative of difficult lighting conditions, such as "night mode," "candlelit," and "no flash". We also filter out any images with titles containing terms associated with high-quality night photography, like "long exposure" or "tripod" to ensure that the images are similar to what we would find in real-world scenarios. Following this, we manually review the images to discard any remaining well-lit images, ensuring the final set of 125 images consisted exclusively of low-light scenarios. As our model handles variable input sizes, each image is preprocessed by resizing it only if its longest dimension exceeded 1024 pixels, preserving its original aspect ratio. Furthermore, to ensure bias mitigation and accurate results, we use cross-validation similar to the SICE dataset.

Other Benchmark Datasets: For a comprehensive perceptual quality comparison against other state-of-the-art methods, we also report performance on the DICM [27], LIME [6], and MEF [28] datasets. These are commonly used benchmarks in the low-light enhancement literature for no-reference quality assessment.

B. Experiments using Various Approaches

The proposed *DA-Zero-DCE++* is considered and validated with 3 model scenarios:

1) ***Approach 1:*** *Fused Images Only:* For this scenario, images from the SICE dataset [22] (Part 1 and Part 2) are considered by combining them as one. The SICE dataset consists of images from various exposure levels. There are 589 multi-exposure sequences from Part 1 and Part 2 of the SICE dataset. We use only the *darkest* to the *mid exposure* images, to create a distinguished subset of images for training. For example, if a multi-exposure sequence consists of 5 images, we take the first 3, if there are 7 images, we take the first 4 and so on. The resulting image is resized to 512×512 which is the recommended dimension used to train Zero-DCE++. These images are then denoised using FFDNet and are then split to train and test, obtaining 471 training images and 118 test images. Apart from the 118 test images generated after the split, we also take the lowest exposure image corresponding to the test images to create a generalized test set. This set of 118 images are not denoised and are closer to the kind of images that we expect to find in real world scenarios. While recording metrics for inference, we do not resize the image. Instead, the resolution remains at 1200×900 just like how it was originally present in the SICE dataset.

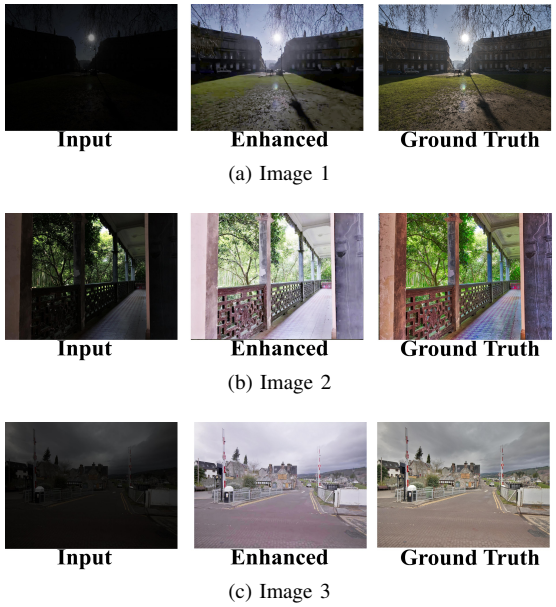
The method is implemented by replicating the parameters and values from the original work by the authors in [13] and was accessed using the version publicly provided *. Fig. 2 shows a comprehensive comparison between the *input* image, *enhanced* image and *ground truth* of the experiments. The performance is validated on several metrics such as Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Mean Absolute Error (MAE), Learned Perceptual Image Patch Similarity (LPIPS) and perceptual quality by calculating the Perceptual Index (PI) [29], [30]. While LPIPS evaluates perceptual similarity by comparing deep feature representations between the enhanced image and the ground truth, PI measures the perceptual quality of the enhanced image without requiring a reference, focusing on how realistic the image appears to human observers. The results on 1200×900 images from the generalized test set are presented in Table II.

TABLE II
APPROACH 1: IMAGE QUALITY WITH VARYING NOISE LEVELS OF FFDNET

Method	PSNR↑	SSIM↑	MAE↓	PI↓	LPIPS↓
FFDNet ($\sigma = 25$)	14.56	0.5677	45.7593	3.9562	0.3698
FFDNet ($\sigma = 10$)	14.16	0.5553	47.7899	3.9695	0.3749
FFDNet ($\sigma = 5$)	14.10	0.5536	48.3695	3.9790	0.3795

Replicating the parameters, the images are scaled to a factor of 12, so even though the test images were 512×512 originally, the used images were cropped to 504×504. We updated the original code to retain the original dimensions and use it for all our approaches. The PSNR and SSIM for the images from the generalized test set of dimension 504×504 is reported at 12.95 and 0.3599 respectively, when the FFDNet had $\sigma = 25$.

*https://github.com/Li-Chongyi/Zero-DCE_extension



(a) Image 1

(b) Image 2

(c) Image 3

Fig. 2. Output of Approach 1: Input, Enhanced and Ground Truth

We recorded a PSNR of 13.06 and SSIM of 0.3357 when FFDNet had $\sigma = 10$. While calculating the image quality metrics for 512×512 test images, we obtained PSNR: 15.35, SSIM: 0.6583 and MAE: 39.9654 and using the same model for 512×512 images from the generalized test set, we obtain PSNR: 14.27, SSIM: 0.5664 and MAE: 47.8010.

2) ***Approach 2:*** *Augmented Dataset and Denoising with FFDNet:* The original Zero-DCE++ used images of all exposures from Part 1 of SICE dataset for training. We use exposure fusion to augment Part 1 of the SICE dataset. To augment the training data, we create a new image for each multi-exposure sequence by fusing the first 3 exposures. The Fig. 4 shows the step wise execution of the exposure fusion process over the selected images. The images were resized to 512×512 and denoised with FFDNet and then were used to train Zero-DCE++. The training pipeline is described in Fig. 3. The test set was created from SICE Part 2. We selected the *darkest* exposures per sequence and used the first 3 or 4 images depending on sequence length. For example, if a sequence had 7 images, we took the first 3 exposures. If it had 9 images, we took the first 4 images. In total, we get 3381 train images and 784 test images. We evaluate the image quality metrics using the test set and obtain results as shown in Table III. The Fig. 5 shows the performance comparison and highlights the *input* image, *enhanced* image and the *ground truth*. Lower values indicate better performance for MAE, PI, and LPIPS. We note that the Approach 2 is able to outperform the Approach 1 in terms of PSNR, MAE, PI metric and LPIPS, showing the importance of using images of all exposures during training.

3) Approach 3: Augmented Dataset Only: The third approach requires augmentation of the training data as the Approach 2, but without any denoising. The images are resized from Part 1 of the SICE dataset to 512×512 and the model is trained on it. Just like before, we use Part 2 of the dataset for testing and obtain image quality as the evaluation metric, which is reported in Table IV. The output, as obtained with

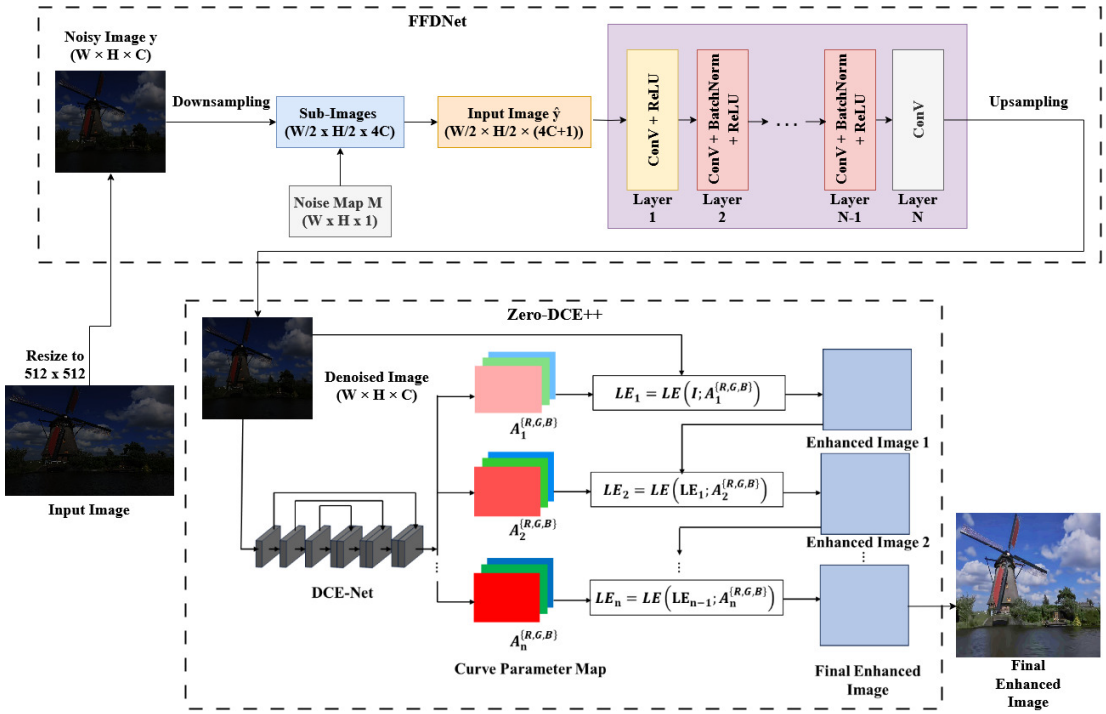


Fig. 3. Model Training Workflow, Data Flow, Processing Stages of Approach 2: Step Wise Execution

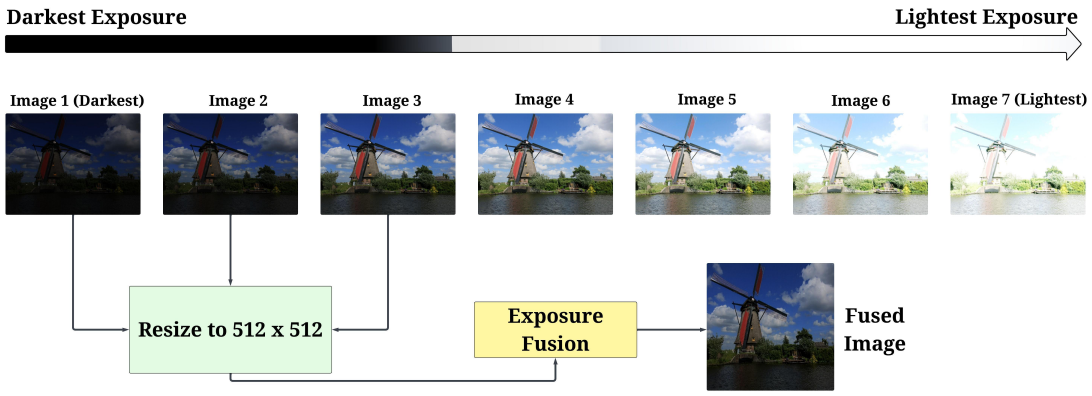


Fig. 4. Exposure Fusion on Sample Images from SICE Part 1 Dataset: Step Wise Execution

TABLE III
APPROACH 2: AVERAGE IMAGE QUALITY WITH VARYING RESOLUTIONS AND VALUES OF σ.

σ	Resolution	PSNR↑	SSIM↑	MAE↓	PI↓	LPIPS↓
5	512 × 512	15.29	0.5024	40.4594	3.4559	0.3171
5	1200 × 900	15.34	0.4869	40.4028	3.3280	0.3581
25	512 × 512	15.62	0.5049	38.3700	3.4396	0.3123
25	1200 × 900	15.67	0.4890	38.2831	3.3195	0.3536

TABLE IV
APPROACH 3: AVERAGE IMAGE QUALITY WITH VARYING RESOLUTIONS

Resolution	PSNR↑	SSIM↑	MAE↓	PI↓	LPIPS↓
512 × 512	15.22	0.5030	40.8739	3.4588	0.3217
1200 × 900	15.27	0.4874	40.8178	3.3347	0.3627

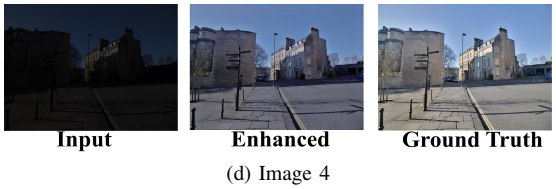
this approach for identified 6 images, is presented in Fig. 6.

Table V presents a comprehensive comparison of the performance achieved by the proposed *DA-Zero-DCE++* (all three approaches) along with several state-of-the-art methods. A careful analysis of the output of these methods reveal that while the PSNR and SSIM of the proposed method is at par with the existing methods (better in some cases), the MAE achieved by the proposed method is significantly reduced. The impact of reduced MAE is reflected in terms of algorithm accuracy and trust development. Clearly, the data pre-processing and augmentation strategies have worked in the favor of the proposed method, where the performance improvement is significant. The reduced MAE is crucial for real-time, mission-critical security systems where low-light surveillance images are used for critical decision making. The use of FFDNet for denoising, exposure fusion for illumination improvement and

(a) Image 1

(b) Image 2

(c) Image 3

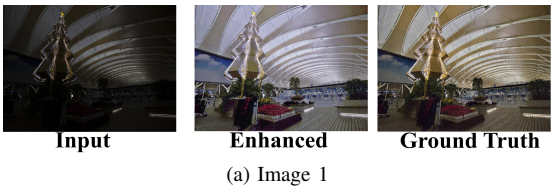


(d) Image 4

(e) Image 5

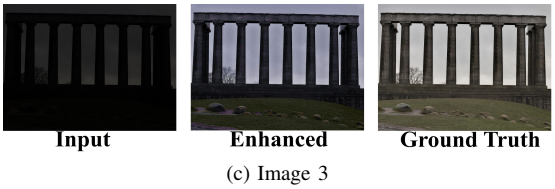
(f) Image 6

Fig. 5. Output of Approach 2: Input, Enhanced and Ground Truth



(a) Image 1

(b) Image 2



(c) Image 3

(d) Image 4

(e) Image 5

(f) Image 6

Fig. 6. Output of Approach 3: Input, Enhanced and Ground Truth

data augmentation for bias mitigation and performance optimization through diverse training samples, in amalgamation are responsible for the improvement in the performance of the proposed method. The proposed method is able to achieve performance improvement on trust driven visual improvements, reduced distributional bias, and deployment fairness across diverse lighting conditions and scenarios. Comparative analysis demonstrates that with the help of zero-reference deep curve estimation, the proposed, *DA-Zero-DCE++*, pipeline achieves improved performance as compared to state-of-the-art low-light image enhancement methods. As shown in Fig. 7, our best configuration, which combines exposure fusion-based augmentation and mild denoising using FFDNet, achieves an average PSNR of $15.34\ dB$, SSIM of 0.4869, and MAE of 40.87 on the *SICE dataset* at 1200×900 resolution.

TABLE V
COMPARISON OF LLIE METHODS ON SICE DATASET 1200×900

Method	**PSNR↑**	**SSIM↑**	**MAE↓**
SRIE [5]	14.41	0.54	127.08
LIME [6]	16.17	0.57	108.12
Li *et al.* [7]	15.19	0.54	114.21
RetinexNet [8]	15.99	0.53	104.81
LightenNet [9]	13.17	0.55	140.92
MBLLEN [2]	15.02	0.52	119.14
Wang *et al.* [10]	13.52	0.49	142.01
EnlightenGAN [11]	16.21	0.59	102.78
Zero-DCE [12]	16.57	0.59	98.78
Zero-DCE++ [13]	16.42	0.58	102.87
Approach 1 ($\sigma = 25$)	14.56	0.57	45.76
Approach 2 ($\sigma = 25$)	15.67	0.49	38.28
Approach 3 (No denoising)	15.27	0.49	40.82

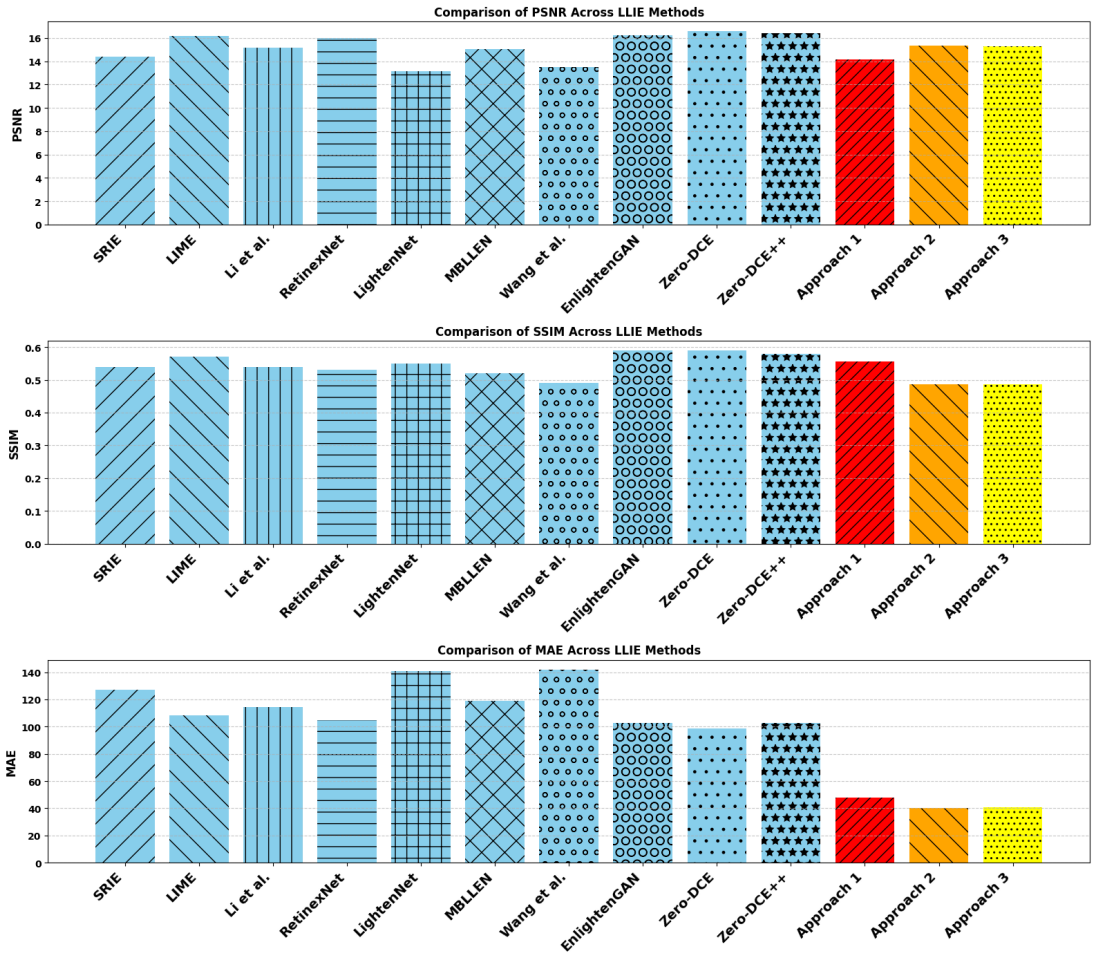


Fig. 7. Performance Comparison of the proposed LLIE Methods Across Evaluation Metrics

C. Naturalness Image Quality Evaluator

The perceptual quality of the enhanced image is measured using Naturalness Image Quality Evaluator (NIQE) for all the three approaches. Mathematically, NIQE is computed as:

$$\text{NIQE}(I) = \sqrt{(f(I) - \mu)^T \Sigma^{-1} (f(I) - \mu)} \tag{4}$$

where $f(I)$ represents the feature vector extracted from image I, and μ and Σ are the mean and covariance of the natural scene statistics model. The NIQE scores of the three approaches are summarized in Table VI. The calculations reflect the model performance and impact of varying parameters. Evidently, the Approach 2 has the minimum score owing to the augmentation and pre-processing. The Approach 1, has the highest score as the set of 118 images are not denoised and are closer to the kind of images that we expect to find in real world scenarios.

TABLE VI
NIQE SCORES ON SICE DATASET

Method	NIQE↓
Approach 1 ($\sigma = 25$)	4.5678
Approach 2 ($\sigma = 25$)	3.7248
Approach 3 (No denoising)	3.7321

V. BENCHMARK EVALUATIONS

For high-level vision applications such as real-time visual analytics over high-bandwidth 6G networks, low-power camera networks in remote or low-light environments, the performance is further validated on DarkFace dataset [23], DICM [27], LIME [6] and MEF [28]. We quantitatively assess how our enhancement pipeline improves face detection performance on the identified dataset. The purpose of this analysis is to highlight the practical benefits of the proposed approach for high-level vision applications such as real-time visual analytics over high-bandwidth 6G networks, low-power camera networks in remote or low-light environments.

A. Perceptual Quality Comparison

We measure the perceptual quality of the image enhanced by the proposed method across various datasets like DICM [27], LIME [6], MEF [28] and VV [†]. To measure the perceptual quality, we use the PI as a non-reference metric such that lower PI values indicate better perceptual quality. It combines the MA [30] and NIQE scores as follows:

[†]https://sites.google.com/site/vonikakis/datasets

TABLE VII
PI SCORE COMPARISON ACROSS VARIOUS DATASETS

Method	DICM	LIME	MEF	VV	Average
SRIE [5]	3.17	**2.76**	2.61	3.37	2.98
LIME [6]	3.35	3.00	2.78	3.03	3.04
Li *et al.* [7]	3.43	3.02	3.61	3.37	3.36
LightNet [9]	3.13	2.84	2.51	3.29	2.94
MBLLEN [2]	3.19	3.18	3.04	3.63	3.26
RetinexNet [8]	3.24	3.08	2.86	**2.95**	3.03
Wang *et al.* [10]	3.20	2.90	2.72	3.42	3.06
EnlightenGAN [11]	3.13	2.83	2.45	4.71	3.28
Zero-DCE [12]	3.04	**2.76**	2.43	3.33	2.89
Zero-DCE++ [13]	3.21	2.97	2.50	3.31	2.99
Proposed	**2.33**	2.94	**2.47**	3.34	**2.77**

$$\mathrm{PI} = \frac{1}{2}\left[(10 - \mathrm{Ma}) + \mathrm{NIQE}\right] \tag{5}$$

where *Ma* is the quality score proposed the authors in [30], and NIQE is the Natural Image Quality Evaluator. The Table VII shows the comparison of PI scores across various methods and datasets such that the bold values denote the best scores in each column. The analysis reveals that while the proposed method achieves at par performance for majority of the the compared methods, at several instances, it is able to even outperforms them with a significant margin.

B. Face Detection in Low-Light Conditions

To evaluate the effectiveness of the proposed method, specifically Approach 2, on a practical downstream task, we apply it to the problem of face detection in low-light conditions. We use our pipeline to enhance low-light images prior to detection, and then perform face detection using the Dual Shot Face Detector (DSFD) [31]. For this evaluation, we employ the challenging DarkFace dataset [26], which contains a wide range of low-light facial images. We benchmark the performance of our method using the official evaluation tools[‡] provided by the DarkFace dataset [26]. Our evaluation focuses on the Average Precision (AP) metric computed at an Intersection-over-Union (IoU) threshold of 0.5. We obtained an AP score of 0.2174 and an AP score of 0.1559 without any enhancement. A set of random output samples are shown in the Fig. 9. The *red* bounding boxes are the faces labeled as the ground truth in the *Darkface dataset* and the *green* boxes are the faces detected by DSFD. We see that LLIE greatly enhances the performance of face detection algorithms like DSFD by enabling it to detect more faces accurately. Further comparative analysis of the performance of the identified methods is presented in Table VIII, where the detection performance (AP, IoU=0.5) is presented for DarkFace Dataset. The graph in Fig. 8 shows that the proposed method is able to achieve significant improvement in the face detection performance when compared to the raw image.

C. Runtime Performance Analysis & Suitability for Real World Deployment

The runtime complexity of the proposed method is compared against other state-of-the-art methods. With only $10,561$ parameters (shown in Table IX), our model demonstrates high efficiency as the modular integration allows the proposed DA-Zero-DCE++ to process denoised, illumination-corrected images at the edge efficiently, enhancing both perceptual quality and computational robustness under diverse low-light and noise conditions. This lightweight design makes it highly suitable for deployment on resource-constrained edge devices as intended.

To evaluate the model's practical performance and scalability, we measure the total inference time on a NVIDIA T4 GPU. The experiment is conducted on progressively larger subsets of the SICE test set for 2 fixed resolutions: 512×512 and 1200×900. Fig 10 shows a linear and predictable increase in computation times for both resolutions. The graph clearly shows the computational trade-off associated with image resolution, with the higher-resolution images requiring significantly more processing time.

Furthermore, to obtain a more granular view of the enhance-

[‡]https://github.com/Ir1d/DARKFACE_eval_tools

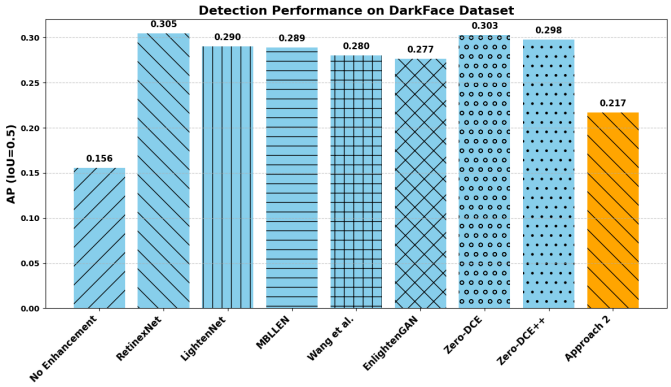


Fig. 8. Comparative Analysis of the AP for Various Approaches

TABLE VIII
DETECTION PERFORMANCE (AP, IOU=0.5) ON DARKFACE DATASET

Method	AP (IoU=0.5)
No enhancement	0.1559
RetinexNet [8]	0.3049
LightenNet [9]	0.2901
MBLLEN [2]	0.2892
Wang *et al.* [10]	0.2800
EnlightenGAN [11]	0.2765
Zero-DCE [12]	0.3031
Zero-DCE++ [13]	0.2979
Proposed (Approach 2, $\sigma = 25$)	**0.2334**

TABLE IX
TRAINABLE PARAMETERS(#P) COMPARISON

Method	#P
LightenNet [9]	29,532
MBLLEN [2]	450,171
RetinexNet [8]	555,205
Wang *et al.* [10]	998,816
EnlightenGAN [11]	8,636,675
Zero-DCE [12]	79,416
Zero-DCE++ [13]	10,561
Proposed (Approach 2, $\sigma = 25$)	**10,561**

(a) Image 1

(b) Image 2

(c) Image 3

(d) Image 4

(e) Image 5

(f) Image 6

(g) Image 7

(h) Image 8

Fig. 9. Comparison of DSFD's Face Detection on Original and Enhanced Images (Part 1). Red Boxes: Ground Truth. Green Boxes: DSFD Detections. Enhanced Images Show Improved Face Detection

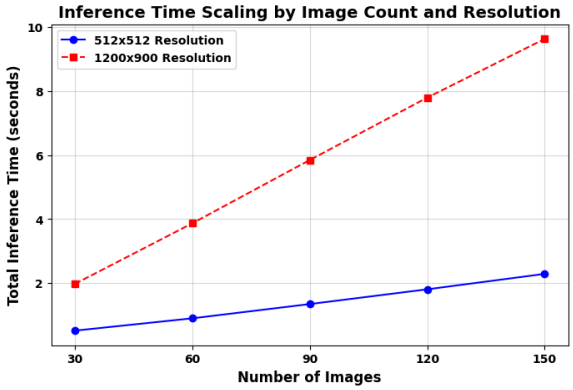


Fig. 10. Inference Time Scaling on the SICE Test Set for 2 Identified Resolutions.

ment quality, we performed histogram and intensity profile analysis on a sample image from the SICE dataset. The outputs are presented in the Fig 11. A careful analysis of the output reveals; (i) a direct visual comparison of the input, our enhanced output, and the ground truth (Fig 11 - a), (ii) the corresponding pixel intensity histograms (Fig 11 - b), and (iii) the intensity profiles along a central row (Fig 11 - c). The histogram profile illustrates the successful redistribution of pixel values across the full dynamic range, while the intensity profile provides clear evidence of fine-detail recovery in underexposed regions.

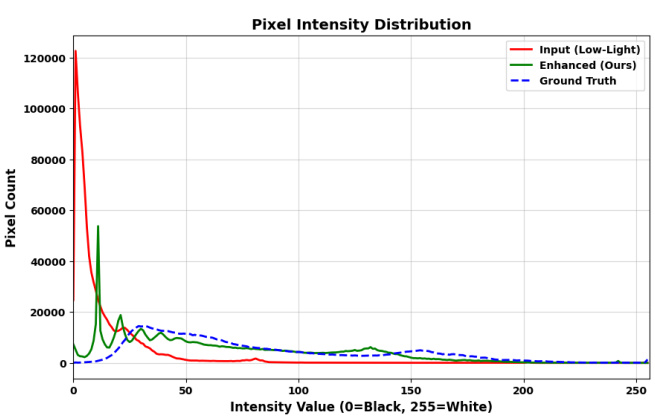
D. Evaluation on Self Curated Real-World Data

To validate the model's performance in a realistic scenario and to test our model's generalization capabilities, we create a new 'in-the-wild' dataset and is described in the Section IV. This dataset is specifically designed to be representative of challenging, user-generated content found in real-world scenarios. We scrape public media forums like Reddit by performing targeted searches for keywords indicative of difficult lighting conditions, such as "night mode," "candlelit," and "no flash". Considering that the proposed model can handle variable input sizes, we pre-process each image by resizing it only if its longest dimension exceeds 1024 pixels, preserving its original aspect ratio. Following this automated process, we manually review the images to discard any remaining well-lit images, ensuring the dataset consisted exclusively of challenging low-light scenarios resulting in a set of 125 low-light images. Some of the images enhanced by the model on the self-curated dataset, are shown in Fig. 12. The images showcase the model's ability to handle diverse and challenging real-world low-light conditions.

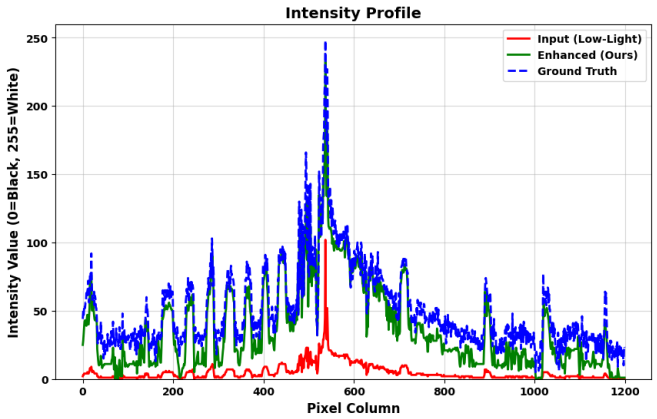
Further, to validate the performance of the proposed method on the self-curated dataset, we perform a quantitative assessment. The metrics for the evaluation was considered as before i.e the same no-reference metrics : the *Naturalness Image Quality Evaluator (NIQE)* and the *Perceptual Index (PI)*. The average scores are reported in Table X

The average NIQE score of 3.7227 is consistent with the score seen on the SICE benchmark (Table VI) from the model trained using Approach 2 which is the model we used here as

(a) Original, Enhanced and Ground Truth Images



(b) Histogram of Pixel Intensity Distribution



(c) Intensity Profile

Fig. 11. Pixel Intensity Distribution and Intensity Profile for a Single Row for a Sample Image. Our Method (Green) Effectively Redistributes the Pixel Values of the Input (Red) to Better Match the Ground Truth (Blue).

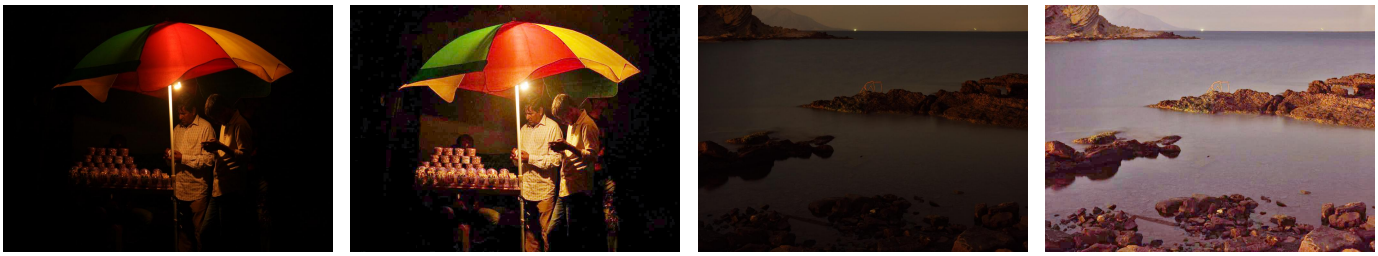
TABLE X
NO-REFERENCE QUALITY ASSESSMENT ON CUSTOM DATASET

Metric	Average Score
NIQE ↓	3.7227
PI ↓	3.31

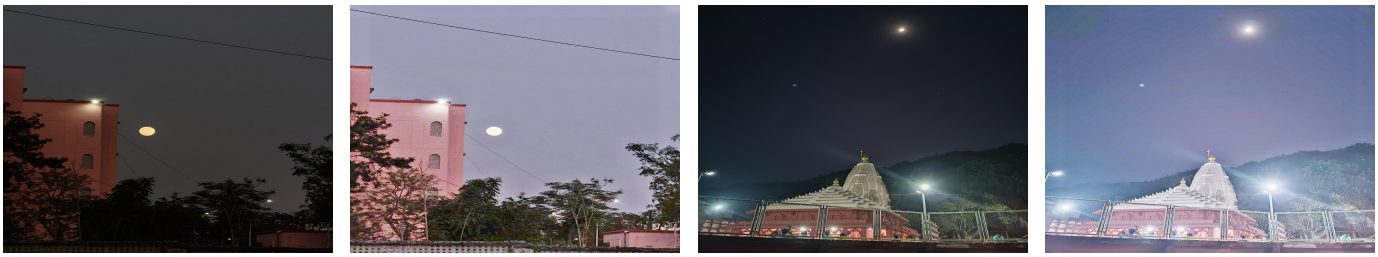
well, indicating that the model maintains its ability to produce natural-looking enhancements on completely unseen data. The PI score of 3.31 is on the higher side of the scores recorded on curated benchmark datasets in Table VII, but this is expected given the inherently more challenging and unpredictable nature of real-world images. These results collectively confirm that our model is not overfitted to a specific benchmark and generalizes effectively to diverse, real-world conditions.

VI. CONCLUSION AND FUTURE WORK

This work proposes *DA-Zero-DCE++*, a data augmentation driven framework to improve visual quality through a combination of data augmentation, illumination correction, and noise suppression techniques, all within a zero-reference learning framework designed for edge-based surveillance systems in 6G-IoT environments. The proposed *DA-Zero-DCE++* integrates exposure fusion as a data augmentation strategy, generating diverse training samples by fusing multiple exposure levels from the same scene to simulate varied low-light conditions. This helps improve the model's robustness and generalization across heterogeneous lighting scenarios. In parallel, FFDNet is incorporated as a pre-processing step for image denoising, with varying levels of noise applied during training to study it's impact on the stability and reliability of the enhancement process. These components are seamlessly combined with the Zero-DCE++ network, which estimates pixel-wise enhancement curves without requiring paired reference images. The proposed method is designed to be lightweight and computationally efficient, enabling real-time deployment on resource-constrained edge devices in a smart city. Through systematic experiments on publicly available datasets, we demonstrated that explicit denoising of training images has a minimal impact when using robust DL models such as Zero-DCE++. Our best approach achieved competitive results compared to state-of-the-art LLIE methods, in terms of PSNR, SSIM, and MAE values. It also surpassed some state-of-the-art methods with respect to perceptual quality. Furthermore, we evaluated the proposed method on the DarkFace dataset and a self curated dataset, showing that our method provides improvements for downstream tasks such as face detection in extremely low-light conditions. Building on these promising results, we continue our work on deploying and evaluating the pipeline on real-world, resource-constrained IoT hardware to benchmark its latency and energy efficiency in the future. We also plan to extend the evaluation to a broader range of downstream vision tasks, including object tracking and action recognition, to further validate its practical utility in security surveillance. We aim to explore more advanced data augmentation techniques and alternative lightweight model architectures to further enhance performance and efficiency.



(a) Input (b) Enhanced (c) Input (d) Enhanced



(e) Input (f) Enhanced (g) Input (h) Enhanced

Fig. 12. Comparison of the Proposed Model's Enhancement on the Self Curated In-The-Wild Datset Images. Each pair (Input–Enhanced) Shows Visual Improvement in Challenging Illumination Conditions.

REFERENCES

[1] M. Diwakar, P. Singh, and P. Kumar, "Chapter 14 - multimodality medical image fusion in shearlet domain," in *Digital Image Enhancement and Reconstruction*, ser. Hybrid Computational Intelligence for Pattern Analysis, S. S. Rajput, N. U. Khan, A. K. Singh, and K. V. Arya, Eds. Academic Press, 2023, pp. 317–328. [Online]. Available: https://www.sciencedirect.com/science/article/pii/B9780323983709000214

[2] F. Lv, F. Lu, J. Wu, and C. Lim, "Mbllen: Low-light image/video enhancement using cnns," in *British Machine Vision Conference (BMVC)*, 2018.

[3] K. Xu, H. Chen, X. Tan, Y. Chen, Y. Jin, Y. Kan, and C. Zhu, "Hfmnet: Hierarchical feature mining network for low-light image enhancement," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–14, 2022.

[4] G.-D. Fan, B. Fan, M. Gan, G.-Y. Chen, and C. L. P. Chen, "Multiscale low-light image enhancement network with illumination constraint," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 11, pp. 7403–7417, 2022.

[5] X. Fu, D. Zeng, Y. Huang, X.-P. Zhang, and X. Ding, "A weighted variational model for simultaneous reflectance and illumination estimation," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2782–2790.

[6] X. Guo, Y. Li, and H. Ling, "Lime: Low-light image enhancement via illumination map estimation," *IEEE Transactions on Image Processing*, vol. 26, no. 2, pp. 982–993, 2017.

[7] M. Li, J. Liu, W. Yang, X. Sun, and Z. Guo, "Structure-revealing low-light image enhancement via robust retinex model," *IEEE Transactions on Image Processing*, vol. 27, no. 6, pp. 2828–2841, 2018.

[8] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," in *British Machine Vision Conference*. British Machine Vision Association, 2018.

[9] C. Li, J. Guo, F. Porikli, and Y. Pang, "Lightennet: A convolutional neural network for weakly illuminated image enhancement," *Pattern Recognition Letters*, vol. 104, pp. 15–22, 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167865518300163

[10] R. Wang, Q. Zhang, C.-W. Fu, X. Shen, W.-S. Zheng, and J. Jia, "Underexposed photo enhancement using deep illumination estimation," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 6842–6850.

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

[12] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, and R. Cong, "Zero-reference deep curve estimation for low-light image enhancement," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2020, pp. 1777–1786.

[13] C. Li, C. Guo, and C. C. Loy, "Learning to enhance low-light image via zero-reference deep curve estimation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 8, pp. 4225–4238, 2022.

[14] Z. Liu, J. Gong, H. Lu, X. Pan, and R. Lan, "Ffenet: Learning frequency features for low-light enhancement," *IEEE Transactions on Instrumentation and Measurement*, vol. 74, pp. 1–13, 2025.

[15] K. Zhang, C. Yuan, J. Li, X. Gao, and M. Li, "Multi-branch and progressive network for low-light image enhancement," *IEEE Transactions on Image Processing*, vol. 32, pp. 2295–2308, 2023.

[16] P. Kaur, G. Panwar, N. Uppal, P. Singh, B. D. Shivahare, and M. Diwakar, "A review on multi-focus image fusion techniques in surveillance applications for image quality enhancement," in *2022 5th International Conference on Contemporary Computing and Informatics (IC3I)*, 2022, pp. 7–11.

[17] G. Tang, J. Ni, Y. Chen, W. Cao, and S. X. Yang, "An improved cyclegan-based model for low-light image enhancement," *IEEE Sensors Journal*, vol. 24, no. 14, pp. 21 879–21 892, 2024.

[18] C. Ahmadi, J.-L. Chen, and Y.-Y. Hsiao, "Zero-dce xt: A computational approach to addressing low-light image enhancement challenges," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 71, no. 9, pp. 4401–4405, 2024.

[19] S.-E. Weng, S.-G. Miaou, and R. Christanto, "A lightweight low-light image enhancement network via channel prior and gamma correction," *arXiv preprint arXiv:2402.18147*, 2024.

[20] C. Xie, H. Tang, L. Fei, H. Zhu, and Y. Hu, "Irnet: An improved zero-shot retinex network for low-light image enhancement," *Electronics*, vol. 12, no. 14, p. 3162, 2023.

[21] E. H. Land, "The retinex theory of color vision," *Scientific american*, vol. 237, no. 6, pp. 108–129, 1977.

[22] J. Cai, S. Gu, and L. Zhang, "Learning a deep single image contrast enhancer from multi-exposure images," *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 2049–2062, 2018.

[23] P. Yang, Z. Ren, D. Zhang, Q. Yu, D. Liang, X. Bai, and Y. Li, "Face detection in the dark," in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, 2019, pp. 0–0.

[24] T. Mertens, J. Kautz, and F. Van Reeth, "Exposure fusion," in *15th Pacific Conference on Computer Graphics and Applications (PG'07)*, 2007, pp. 382–390.

[25] K. Zhang, W. Zuo, and L. Zhang, "Ffdnet: Toward a fast and flexible solution for cnn-based image denoising," *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4608–4622, 2018.

[26] W. Yang, Y. Yuan, W. Ren, J. Liu, W. J. Scheirer, Z. Wang, T. Zhang, Q. Zhong, D. Xie, S. Pu, Y. Zheng, Y. Qu, Y. Xie, L. Chen, Z. Li, C. Hong, H. Jiang, S. Yang, Y. Liu, X. Qu, P. Wan, S. Zheng, M. Zhong, T. Su, L. He, Y. Guo, Y. Zhao, Z. Zhu, J. Liang, J. Wang, T. Chen, Y. Quan, Y. Xu, B. Liu, X. Liu, Q. Sun, T. Lin, X. Li, F. Lu, L. Gu, S. Zhou, C. Cao, S. Zhang, C. Chi, C. Zhuang, Z. Lei, S. Z. Li, S. Wang, R. Liu, D. Yi, Z. Zuo, J. Chi, H. Wang, K. Wang, Y. Liu, X. Gao, Z. Chen, C. Guo, Y. Li, H. Zhong, J. Huang, H. Guo, J. Yang, W. Liao, J. Yang, L. Zhou, M. Feng, and L. Qin, "Advancing image understanding in poor visibility environments: A collective benchmark study," *IEEE Transactions on Image Processing*, vol. 29, pp. 5737–5752, 2020.

[27] C. Lee, C. Lee, and C.-S. Kim, "Contrast enhancement based on layered difference representation of 2d histograms," *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 5372–5384, 2013.

[28] K. Ma, K. Zeng, and Z. Wang, "Perceptual quality assessment for multi-exposure image fusion," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3345–3356, 2015.

[29] Y. Blau and T. Michaeli, "The perception-distortion tradeoff," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 6228–6237.

[30] C. Ma, C.-Y. Yang, X. Yang, and M.-H. Yang, "Learning a no-reference quality metric for single-image super-resolution," *Comput. Vis. Image Underst.*, vol. 158, no. C, p. 1–16, May 2017. [Online]. Available: https://doi.org/10.1016/j.cviu.2016.12.009

[31] J. Li, Y. Wang, C. Wang, Y. Tai, J. Qian, J. Yang, C. Wang, J. Li, and F. Huang, "Dsfd: Dual shot face detector," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 5055–5064.

Vishal Krishna Singh received his bachelor's degree in Information Technology, in 2010, the master's degree in Computer Technology and Application, in 2013, and PhD degree in Information Technology from Indian Institute of Information Technology, Allahabad, India in 2018. He is currently working as a Lecturer and is associated with the Networks and Communications Research Group at School of Computer Science and Electronics Engineering, University of Essex, Colchester, U.K. His research interests include Internet of Things, Wireless Sensor Networks, In-Network Inference, Machine Learning and Data Analytics.

Niharika Anand received PhD from the Indian Institute of Information Technology, Allahabad, India. She is currently working as an Assistant Professor in the Department of Information Technology at the Indian Institute of Information Technology, Lucknow, India. Her research areas include Machine Learning, Deep Learning, Federated Learning, Cloud Computing, Internet of Things, Cyber Forensics, 3-D Wireless Sensor Networks, Wireless Sensor Network Localization, and WSN Topology Control and Maintenance.

Krishna Sharma S received the B.Tech. degree in Computer Science and Engineering from Amrita Vishwa Vidyapeetham, Coimbatore, India, in 2022. He is currently pursuing the M.Tech. degree in Computer Science with a specialization in AI at the Indian Institute of Information Technology, Lucknow, India. His current research focuses on low-light image enhancement, and his broader interests include Machine Learning, Image Processing, and Large Language Models.

Anjali received the Ph.D. degree in Information Technology from the Indian Institute of Information Technology Allahabad, Prayagraj, India, in 2020. Dr. Anjali is currently working as an Assistant Professor with the Department of Information Technology, ABV-Indian Institute of Information Technology, Gwalior. Her research interests include time series analysis, time series data mining, machine learning, deep learning and multimedia analysis.

Mahendra K. Shukla (Senior Member, IEEE) received the Ph.D. degree in electronics and communication engineering from the Indian Institute of Information Technology Allahabad, Prayagraj, India, in 2018. Dr. Shukla is currently working as an Assistant Professor with the Department of Information Technology, Indian Institute of Information Technology, Gwalior (an Institute of National Importance under Ministry of Education, Govt. of India). Prior to that, he was working with the Department of Electronics and Communication Engineering, The LNM Institute of Information Technology, Jaipur, India as an Assistant Professor. He worked as a Postdoctoral Fellow in the Department of Electrical and Computer Engineering, University of Saskatchewan, Canada, in 2019-2020. His research interests include intelligent wireless networks, physical layer security, cyber-physical systems, IoT, and low-power networking.

Dr. Rajkumar Singh Rathore (Senior Member IEEE) is working as Head of Cyber Security of Connected and Autonomous Systems, CINC, Head of Cyber Physical and Networks Systems, CeRISS & Programme Director for MSc Computing and IT in Department of Computer Science at Cardiff Metropolitan University's School of Technologies, United Kingdom. He has gained doctorate degree, dual master's degrees, and bachelor's degree all in Computer Science and Engineering discipline. He is the Fellow of HEA UK. His research expertise are Wireless Communications, Internet of Things/Cyber Physical Systems, Cyber Security and Privacy, Connected and Autonomous Vehicles, EV Charging Infrastructure Management, Intelligent Networking of Drones and also use cases of AI/ML.

Prof. Weiwei Jiang (Senior Member IEEE) received the B.Sc. Degree of Electronic Engineering and Ph.D. Degree of Information and Communication Engineering from the Department of Electronic Engineering, Tsinghua University, Beijing, China, in 2013 and 2018, respectively. He is currently an assistant professor with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, and Key Laboratory of Universal Wireless Communications, Ministry of Education. His current research interests include artificial intelligence, machine learning, big data, wireless communication and edge computing. He has published more than 60 academic papers in IEEE Trans and other journals, with more than 3900 citations in Google Scholar. He is one of 2023 and 2024 Stanford's List of World's Top 2% Scientists.