Proceedings of CONF-CDS 2025 Symposium: Application of Machine Learning in Engineering
DOI: 10.54254/2755-2721/2025.AST26121

Enhancing Zero-DCE via Large Language Model-Driven Loss Function Design for Low-Light Image Correction

Le Yu

College of International Education, Changchun University of Technology, Changchun, China
yule0459@163.com

Abstract. This work investigates the use of Large Language Models (LLMs) to automate the design and optimization of loss functions for low-light image enhancement. We focus on improving the Zero-DCE algorithm, where conventional manually crafted losses often struggle to balance brightness enhancement, detail preservation, and noise suppression simultaneously. Leveraging the semantic reasoning and generative capabilities of LLMs, we introduce an iterative training–feedback–regeneration loop that constructs and refines multi-objective loss functions based on structured performance feedback. The proposed approach enables adaptive loss design without manual tuning, progressively improving model performance across diverse low-light scenarios. Extensive experiments demonstrate that LLM-guided optimization offers a promising paradigm for integrating foundation models into machine learning training, providing both research and practical value.

Keywords: Large Language Models, Loss Function Optimization, Low-Light Image Enhancement, Multi-Objective Learning

1. Introduction

In various applications such as surveillance, autonomous driving, and medical imaging, insufficient illumination often leads to severely degraded images, posing significant challenges for subsequent visual tasks like recognition and detection [1]. Low-light image enhancement techniques aim to improve brightness, enhance details, and suppress noise effectively, making them indispensable in modern computer vision pipelines. Recently, unsupervised deep learning methods such as Zero-DCE have gained attention due to their independence from ground-truth data [2]. However, manually crafted loss functions in such models often fail to balance competing objectives—brightness, noise suppression, and detail preservation—across diverse scenarios, limiting their generalization capability.

To address this limitation, we propose an approach that leverages LLMs to automatically optimize the loss function of Zero-DCE. By interpreting task objectives such as brightness balance, detail retention, and noise reduction, and incorporating structured performance feedback, LLMs iteratively generate composite loss functions [3]. We design an iterative optimization framework: model training → performance evaluation → structured feedback → LLM-based loss function generation → next training iteration. This cyclical process progressively approaches an optimal loss configuration.

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2755-2721/2025.AST26121

Our contribution lies in introducing a novel, intelligent, and feedback-driven loss function generation approach that overcomes the limitations of manual design and provides a scalable optimization pathway for complex visual tasks. Designing task-specific loss functions traditionally relies heavily on expert experience, requiring extensive trial-and-error to tune each component and weight, which is both time-consuming and difficult to generalize. In this context, an LLM-driven loss function generation mechanism emerges as a compelling alternative [4]. By parsing task goals and integrating feedback, LLMs can automatically construct multi-objective, adaptive loss functions, offering a more intelligent and flexible optimization strategy. This represents a promising direction for low-light image enhancement and other challenging computer vision tasks, with significant research and application potential.

2. Related work

2.1. Advances in low-light enhancement models

Traditional image enhancement methods, such as Histogram Equalization (HE), Gamma Correction, and Retinex-based techniques [5], are effective under controlled lighting conditions but often underperform in real-world scenarios characterized by strong noise and complex structures. These approaches typically fail to simultaneously enhance brightness and preserve fine details, often leading to over-enhancement or image distortion.

Deep learning, particularly unsupervised learning, has emerged as the mainstream paradigm for image enhancement [6]. Zero-DCE [2] employs a curve estimation network to perform end-to-end image enhancement, guided by a set of manually designed unsupervised loss terms—such as exposure control, color constancy, and structure preservation—achieving competitive results without requiring paired training samples. However, the fixed structure of its loss function limits its adaptability to diverse application scenarios, constraining its potential for further improvement.

2.2. Applications of LLMs in machine learning model training

The remarkable understanding, reasoning, and generative capabilities of Large Language Models (LLMs), demonstrated in natural language processing, have catalyzed their adoption across various stages of machine learning (ML) model training, fostering innovation in training methodologies [7].

Specifically, LLMs have contributed in the following areas:

(1) Supervised training enhancement: generating training labels, optimizing synthetic data, and refining augmentation strategies [8];

(2) Loss function design: overcoming the limitations of expert-driven approaches by interpreting task specifications and performance feedback to automatically generate multi-objective loss functions, thereby improving generalization and lifting the performance ceiling [4];

(3) Hyperparameter optimization and debugging: deriving adaptive tuning strategies from training logs and metric analyses [9];

(4) Training monitoring and diagnosis: detecting overfitting or underfitting and suggesting corrective actions.

In computer vision, the cross-modal abilities of LLMs have further expanded their applicability [10]. For example, LLMs assist in crafting more precise prompts to improve the performance of image generation, editing, and segmentation tasks [11], and optimize the articulation of task objectives and training procedures. Recent studies have also explored leveraging LLMs to directly

DOI: 10.54254/2755-2721/2025.AST26121

guide the training of image generation models [12] or to provide task-driven feedback and guidance in applications such as Visual Question Answering (VQA) [13] and image-text matching.

3. Loss function optimization framework driven by large language models

In low-light image enhancement, the design of the loss function is critical to model performance. Traditionally, researchers manually construct loss functions by combining brightness error, color consistency, exposure, and other objectives with fixed weights, based largely on empirical knowledge. However, such static and rigid structures often fail to achieve satisfactory performance in complex, diverse scenarios and require extensive trial-and-error tuning. To overcome this limitation, we propose an LLM-driven optimization framework that evolves the loss function dynamically through iterative training feedback, enhancing the adaptability and intelligence of the image enhancement model.

The proposed framework comprises two stages: initialization and optimization, as illustrated in Figure 1.

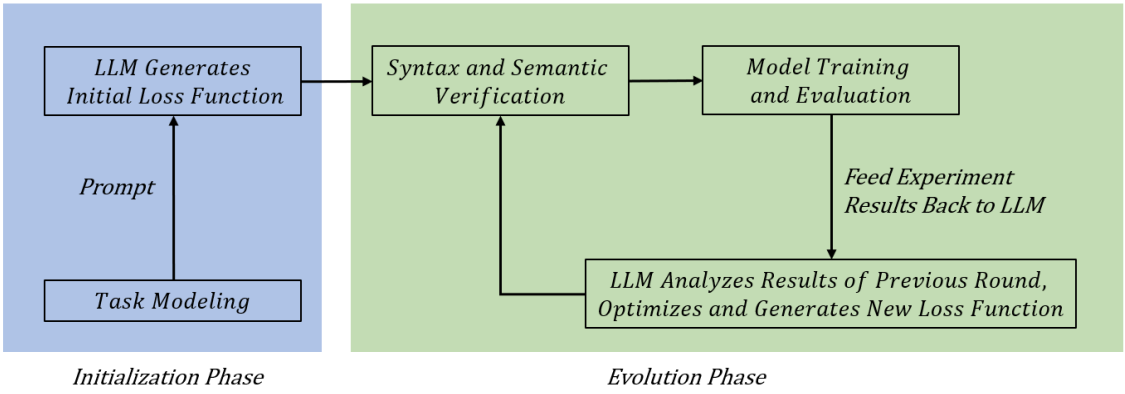


Figure 1. LLM-driven loss function optimization framework

3.1. Initialization stage

At the outset, a basic loss formulation and current performance metrics are provided to establish a baseline for subsequent optimization. As shown in Figure 1, this stage involves two main steps: task modeling and initial loss function generation.

We construct a structured prompt to guide the LLM in formulating an initial loss function based on its semantic understanding of the task. The prompt encodes three key elements:

(1) Task objectives: expressed in natural language to specify the optimization goals, such as "enhance global brightness uniformity," "suppress local noise," "improve edge texture," and "maintain color constancy to avoid color cast";

(2) Baseline loss, defined as:

$$L_{Baseline} = \frac{1}{H \times W \times C} \sum_{i=1}^{H} \sum_{j=1}^{W} \sum_{c=1}^{C} \left(\widehat{I}_{ijc} - \gamma \right)^2 \tag{1}$$

where $H, W,$ and C denote the image height, width, and number of channels, respectively; $\widehat{I}_{ijc}$ is the pixel value at (i, j) in channel c of the enhanced image; and $\gamma = 0.6$ is the target

DOI: 10.54254/2755-2721/2025.AST26121

brightness level, chosen to avoid overexposure (>0.8) or underexposure (<0.3);

(3) Baseline performance metrics: evaluated on a standard dataset, with PSNR = 7.55 (indicating severe noise and distortion, suggesting inclusion of noise-suppression terms), SSIM = 0.1939 (indicating structural loss, suggesting structural enhancement components), and MSE = 0.1940 (indicating brightness imbalance, suggesting adjustments to brightness constraints).

The LLM synthesizes an optimized loss function by leveraging its pre-trained knowledge and analyzing the baseline loss and performance metrics.

3.2. Optimization stage

Following initialization, the process enters a feedback-driven optimization phase.

As shown in Figure 1, each generated loss function undergoes a validity check to ensure mathematical and syntactic correctness before being used in training. Through an iterative cycle of training → feedback → regeneration, the loss function is progressively refined to better suit the low-light image enhancement task.

For each candidate loss function, a Zero-DCE-based model is trained and evaluated on the test set. Performance feedback from each iteration is analyzed by the LLM, which then proposes refinements—such as introducing new terms (e.g., total variation, perceptual loss, edge-aware loss), adjusting weights of existing terms, removing redundant components, or reorganizing the loss structure. The refined loss function is then used for the next round of training, with updated performance metrics fed back to the LLM for further improvement.

Ultimately, the loss function achieving the best overall performance across iterations is selected as the final objective for training the Zero-DCE model.

4. Experimental results and analysis

The experiments were conducted on an NVIDIA GeForce RTX 3050 Ti GPU using the PyTorch deep learning framework. The learning rate was set to 1×10^{-4}, the batch size was 8, and training was performed for 200 epochs.。

We conducted experiments on the LOL [14] and VE-LOL [15] datasets. To systematically evaluate the effectiveness of the proposed LLM-optimized loss function for low-light image enhancement, we compared three loss configurations under identical network architecture and hyperparameter settings: Baseline (a minimal structure containing only a brightness term, serving as a reference), Improved Zero-DCE (manually designed by researchers), and LLM optimization (the final configuration automatically generated and iteratively refined by the proposed LLM-driven framework).

4.1. Quantitative results

We evaluated the models using PSNR, SSIM, and MSE on the LOL and VE-LOL-L datasets. The results are summarized in Table 1.

Compared to the baseline loss, both the manually designed and LLM-optimized methods achieve lower MSE, indicating reduced pixel-level discrepancies with respect to the ground-truth images. Additionally, higher PSNR and SSIM values reflect improved structural fidelity. Notably, the LLM-optimized model demonstrates comparable or superior performance to the manually designed variant, achieving enhanced reconstruction quality and structural similarity.

DOI: 10.54254/2755-2721/2025.AST26121

Table 1. Performance comparison of different loss configurations

Method	LOL			VE-LOL-L		
	MSE	PSNR	SSIM	MSE	PSNR	SSIM
Baseline	0.1940	7.55	0.1939	0.1903	7.68	0.1802
Improved Zero-DCE	0.0265	16.55	0.5241	0.0210	17.44	0.5372
LLM optimization	0.0263	16.70	0.5211	0.0224	16.81	0.5131

Visual comparisons on the LOL test images are shown in Figure 2. The LLM-optimized results exhibit more faithful detail restoration (e.g., textures of books, contours of glass and objects), more natural colors, and improved clarity, producing perceptually superior enhancements compared to the manually designed method.

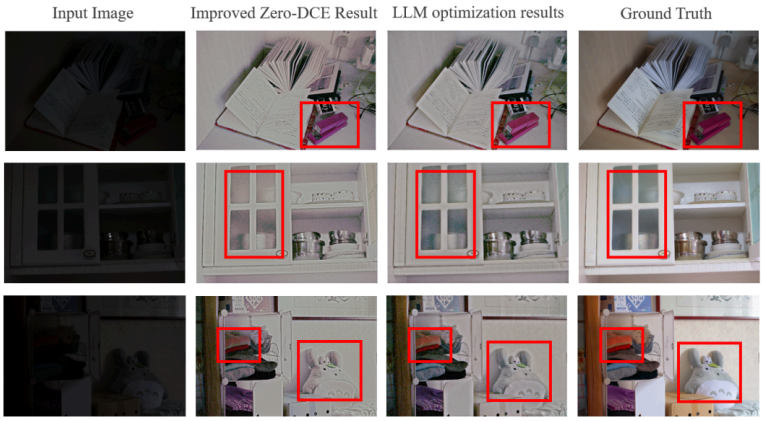


Figure 2. Qualitative comparison before and after optimization

4.2. Analysis

On the LOL dataset, the LLM-optimized method reduced MSE from 0.1940 to 0.0263 (an 86.44% decrease), improved SSIM from 0.1939 to 0.5211 (a 168.75% increase), and raised PSNR from 7.55 to 16.70. On VE-LOL-L, MSE decreased from 0.1903 to 0.0224 (an 88.23% reduction), SSIM improved from 0.1802 to 0.5131 (a 184.74% increase), and PSNR increased from 7.68 to 16.81.

Relative to the manually designed method, the LLM-optimized model exhibited marginal advantages on LOL in MSE and PSNR, while showing slight disadvantages on VE-LOL-L in all metrics. This may be attributed to the dynamic nature of VE-LOL-L, which contains video frames with additional degradations [15] , potentially contributing to lower SSIM.

In summary, the LLM-optimized approach achieves substantial improvements over the baseline across key metrics, underscoring its potential for low-light enhancement tasks. These results confirm that the LLM-driven framework effectively discovers loss configurations with better generalization and performance by leveraging feedback-driven optimization, surpassing the limitations of manual tuning.

5. Conclusion

This paper presents an LLM-driven framework for automated loss function optimization, which enhances the performance of the Zero-DCE model in low-light image enhancement tasks. By

DOI: 10.54254/2755-2721/2025.AST26121

introducing a closed-loop mechanism of training–feedback–regeneration, the proposed approach enables dynamic loss function adaptation—addressing, for the first time in this context, the challenge of jointly optimizing brightness enhancement, detail preservation, and noise suppression. Experiments on the LOL and VE-LOL-L datasets demonstrate that, relative to the baseline, the LLM-optimized scheme reduces MSE by an average of 87%, improves SSIM by 176%, and increases PSNR by 120%, significantly improving reconstruction quality. Compared to manually designed losses, the LLM-optimized method achieves a 0.91% PSNR gain and a 0.76% reduction in MSE on the LOL dataset, with less than a 3% performance difference on VE-LOL-L—all achieved without any manual parameter tuning.

This work offers a scalable, automated loss function design paradigm for complex vision tasks, substantially reducing reliance on expert knowledge. Its feedback-driven architecture provides a generalizable strategy for adaptive model optimization. Future research will explore incorporating physical image constraints to improve cross-domain robustness, enhancing the interpretability of LLM-generated losses, and extending the framework to temporal tasks such as video enhancement.

References

[1] Kim W. Low-light image enhancement: A comparative review and prospects [J]. IEEE Access, 2022, 10: 84535-84557.
[2] Guo C, Li C, Guo J, et al. Zero-reference deep curve estimation for low-light image enhancement [C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 1780-1789.
[3] Guo P F, Chen Y H, Tsai Y D, et al. Towards optimizing with large language models [J]. arXiv preprint arXiv: 2310.05204, 2023.
[4] Wang Z, Ma C, Zhou L, et al. $\mathrm {E^{2} CFD} $: Towards Effective and Efficient Cost Function Design for Safe Reinforcement Learning via Large Language Model [J]. arXiv preprint arXiv: 2407.05580, 2024.
[5] Yang W, Wang W, Huang H, et al. Sparse gradient regularized deep retinex network for robust low-light image enhancement [J]. IEEE Transactions on Image Processing, 2021, 30: 2072-2086.
[6] Ma L, Ma T, Liu R, et al. Toward fast, flexible, and robust low-light image enhancement [C]. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2022: 5637-5646.
[7] Xu J, Li J, Liu Z, et al. Large language models synergize with automated machine learning [J]. arXiv preprint arXiv: 2405.03727, 2024.
[8] Garcia-Carmona A M, Prieto M L, Puertas E, et al. Leveraging Large Language Models for Accurate Retrieval of Patient Information From Medical Reports: Systematic Evaluation Study [J]. JMIR AI, 2025, 4(1): e68776.
[9] Kumar A, Owen L, Chowdhury N R, et al. Zclip: Adaptive spike mitigation for llm pre-training [J]. arXiv preprint arXiv: 2504.02507, 2025.
[10] Ma B, Zong Z, Song G, et al. Exploring the role of large language models in prompt encoding for diffusion models [J]. arXiv preprint arXiv: 2406.11831, 2024.
[11] Lian L, Li B, Yala A, et al. Llm-grounded diffusion: Enhancing prompt understanding of text-to-image diffusion models with large language models [J]. arXiv preprint arXiv: 2305.13655, 2023.
[12] Lee K, Liu H, Ryu M, et al. Aligning text-to-image models using human feedback [J]. arXiv preprint arXiv: 2302.12192, 2023.
[13] Subramanian S, Narasimhan M, Khangaonkar K, et al. Modular visual question answering via code generation [J]. arXiv preprint arXiv: 2306.05392, 2023.
[14] Wei, C., Wang, W., Yang, W., & Liu, J. (2018). Deep retinex decomposition for low-light enhancement. arXiv preprint arXiv: 1808.04560.
[15] Liu, J., Xu, D., Yang, W., Fan, M., & Huang, H. (2021). Benchmarking low-light image enhancement and beyond. International Journal of Computer Vision, 129, 1153-1184.