Dark Light Enhancement For Dark Scenes Urban Object Recognition

Keshava Kumar Boddepalli
Department of CSE,
VIT-AP University, India
keshavakumar.21bce9407@vitapstudent.ac.in

Krishnaja Venkata Naga Sri Lasya P
Department of CSE,
VIT-AP University, India
srilasya.21bce7477@vitapstudent.ac.in

Munisaiteja Sharan Tupakula
Department of CSE,
VIT-AP University, India
munisaiteja.21bce9423@vitapstudent.ac.in

Nirmal Viswas Jami
Department of CSE,
VIT-AP University, India
nirmalviswas.21bce9285@vitapstudent.ac.in

Koduru Hajarathaiah
Department of CSE,
VIT-AP University, India
hajarathaiah.k@vitap.ac.in

***Abstract*—The performance of four advanced deep learning architectures, UNet, Autoencoder, Deep Convolutional Neural Network (DCNN), and ResNet will be discussed, focusing in particular on the problem of low-light urban object recognition. So far, other works focused on a single model. In this work, a comparative evaluation is conducted across these frameworks with regard to their ability in terms of image de-noising, contrast enhancement, and especially to object recognition accuracy. Synthetic dataset was created using replicated noise patterns that were derived from low illumination urban environment enhanced with data augmentation and RGB-to-RAW unprocessing for generalization of the model. The results show that there are huge variations in the strengths and weaknesses exhibited by the models and thus provide some meaningful insights related to application areas such as autonomous driving, urban surveillance, and pedestrian detection in low-light environments. The comparative study will help in the identification of best frameworks that enhance the optimization of images captured under low-light conditions and object recognition while advancing methodologies in deep learning based on complex environments within urban settings.**

***Index Terms*—Low-light image enhancement, urban object recognition, deep learning models, image de-noising, synthetic noise generation.**

I. Introduction

Enhancing dark-light images has become a fundamentally critical input considering recent advancements in autonomous technologies such as self-driving cars and robotic navigation systems. In metropolitan cities, fluctuating conditions at nighttime greatly degrade the effectiveness of camera-based perception systems. Because of this, improving low-light images itself turns to be an essential undertaking. Inadequate illumination reduces the number of photon signals detected by camera sensors, and images generated therefore contain high noise levels and poor visibility. This degradation impacts object detection and scene segmentation as well as many other tasks in vision that are required to support the autonomous operation of aircraft. Moreover, a failure in appropriate object identification under low light conditions will lead to higher safety issues, thus making it critical for autonomous systems to embrace advanced image enhancement techniques. Such systems can thereby receive improved situational awareness through enhanced image quality captured in low-light environments, thus enhancing their operational efficiency and dependability across diverse environments. Many solutions have been proposed to solve the problem of image degradation in low illumination levels. Traditional techniques like median filtering and NLM methods attempt to remove noise but often at the cost of image detail. Deep architectures, such as CNNs, have opened new avenues to low-light image enhancement. Models like UNet are highly effective in applications of image de-noising and object recognition to improve the clarity and resolution of images captured in low-light environments. A single architecture may not have the potency to withstand all types of noises and distortion. Therefore, there is a need to discuss different architectures for handling variable levels of degradation in the image.

This paper therefore explores and compares four state-of-the-art deep learning models—namely, UNet, Autoencoder, Deep Convolutional Neural Networks (DCNN), and ResNet—on their effectiveness for low-light urban object recognition. Although the prior studies, including that of Zhou et al., focused solely on the implementation of UNet for enhancing low-light images, our goal is to provide an in-depth comparison among these four models in terms of their performance for image de-noising, contrast enhancement, and accuracy in object recognition.This work tests every model with a synthetic dark-light dataset whose statistical properties mimic those of real noise. Here, the aim is to bring forth the good and less desirable sides of each model in challenging urban scenarios. This comparative analysis is intended to study the low-light image enhancement capability on UNet, Autoencoder, DCNN, and ResNet and how these contributions may affect applications that could demand autonomous navigation as well as object detection capabilities. Key contributions include a noise generation process that simulates low-light scenarios and evaluates implementation through qualitative and quantitative analyses to enhance dark images.

II. Literature Review

Xiao et al. (2023) [1] discusses the dark light enhancement techniques designed for low-light urban object recognition. Techniques they present enhance clarity and details in images, thereby improved object identification that usually takes place in conditions poorly lit. Their results have brought significant breakthroughs in precision algorithms detecting objects and have pointed to the need for enhancement of low-light images in applications for urban usage. Panagiotou & Bosman (2024) [2] proposed a novel denoising diffusion post-processing technique to improve the quality of low-light images. The method effectively removed noise without affecting most features of the image, thus leading to enhancement. Experimentation proved that such an approach could be taken to make it possible to generate more effective images, thus allowing performance excellence in subsequent computer vision applications. Zhang et al. (2019) [3] proposed an effective low-light image enhancement model that utilizes state-of-the-art algorithms for improving the clarity of the images. They overcome general low-light challenges and developed an approach that enhances not only aesthetic quality but also enhances the overall performance of object recognition systems. It offers insights into how to balance the quality of enhancement with computational efficiency. Li et al. (2023) [4] proposes an approach for lowlight image enhancement under nonuniform dark conditions. Their adaptation of illumination level variations across the image and tailored enhancements for localized lighting effects makes the proposed technique suitable for real scenarios. This work points towards the necessity of context-aware enhancement techniques to enhance object visibility in varied lighting conditions. Lore et al. (2017) [5] designed LLNet, a deep autoencoder for improving low-light images of nature. The paper is based on the principle that deep learning surpasses the traditional state of affairs in capturing small details existing in lowlight images. The LLNet model shows that neural networks are fit to improve image quality with a lower computational cost. Ye et al. (n.d.) [6] proposed an external memory-based low-light image enhancement method that enhances images effectively. The contextual information remembered by the model through the use of the external memory structures improves the quality of the image. This work indicates the requirement of memory-based methods in the handling of complexities in low-light images. Tao et al. put forth a Convolutional Neural Network model, LLCNN for low-light image enhancement in 2017 [7]. The Conclusion gives an overview that this model, LLCNN improves the observance of dark images with natural representations of colors as was true before. The work clarifies the usability of Convolutional Neural networks about such low light challenges to implement into various applications. Wang et al. (2022) [8] proposed the normalizing flow-based technique for low-light image enhancement. Their technique is built on learning data distributions, which can enhance the quality of images, providing new insights into low-light enhancement. The results show normalizing flow to be quite a powerful tool for generating highly realistic and enhanced low-light images. Jeong & Lee (2021) [9] details the optimization-based gamma correction method for lowlight image enhancement. This approach provides a systematic way of estimating the parameters of gamma correction, thus providing the better clarity of the image. Therefore, it is further contribution toward understanding in how the gamma correction has to be used under low lighting conditions. Wang et al. (2023) [10] present an illumination-aware gamma correction method that is embedded in a full image modeling network for enhancing low-light images. It clearly explains how the work focuses on the modeling of illumination to get good enhancement results, demonstrating how a fusion of various techniques enhances the quality of an image significantly.

The reviewed literature thus encompasses the range from typical low-light image enhancement algorithms to more sophisticated deep learning methods. Each of the related works adds to the cumulative knowledge regarding how enhancements relate to improving the quality and detection of objects in poor light. Evolution has naturally led to a trend on context-awareness and complexity of the algorithms used on these techniques, which hopefully can be a stepping point toward even more robust applicability.

III. Methodology

We explored and applied four advanced deep learning models, including UNet, Autoencoder, Deep Convolutional Neural Networks (DCNN), and ResNet, for the enhancement of low-light urban images and object recognition accuracy. The proposed methodology handles noise, contrast, and low-light conditions in a structured approach with synthetic dataset generation, preprocessing, augmentation, training, and evaluation. Mathematical formulations for preprocessing and model performance evaluation are incorporated throughout.

A. Datasets:

The datasets for this paper were carefully collected to include a wide variety of lighting conditions and object types to overcome the difficulties of low-light environments. Focus was given to urban images that reflect diverse settings where different levels of illumination affect object visibility. Notably, we utilized the ExDark dataset, which features paired low-light and enhanced images for training models under realistic conditions. Besides, custom datasets were developed targeting some object classes, including cars, pedestrians, and traffic signs, in order to demonstrate realistic scenario conditions.

B. Model Selection:

U-Net: The architecture of the U-Net is pretty effective for any tasks demanding detailed image restoration and enhancement. Its multi-scale feature extraction capabilities and skip connections are efficient in capturing the features which are low level and high level to bring out the lost and detect objects in images shot with low light. The Skip connections allow the network to retain the essentials at every step of the processing chain to yield robust performance for enhancing and detecting objects.

Deep Convolutional Neural Network (DCNN): DCNNs are powerful in complex feature extraction and more so handle deeper layers of more abstract representations of images. They are best suited in finding intricate patterns, which usually is the case for object detection which has poor visibility. In our system, DCNN helps in learning detailed features that distinguish between objects even under tough lighting conditions.

ResNet: This is a residual connection model. ResNet has been proven to significantly improve the detection accuracy because deep networks are usually plagued by the vanishing gradient problem. This model does not degrade with the added layers and can be very effective for object enhancement and detection in images. Its architecture also helps preserve important features in an image during enhancement.

Autoencoder: Autoencoders are very effective in feature extraction of meaningful features from images, through learning compressed representation. In low light conditions, autoencoders are very helpful as they can emphasize enhancing and refining details in the image. In our method, the autoencoder helped in enhancing the quality of features that would allow the network to correctly detect objects.

C. Noise Generation and Synthetic Dataset Creation:

Noise in low-light imaging degrades image quality quite seriously, hence causing significant degradations in object detection and recognition. In the process of developing noise generation, we used the Poisson distribution applied on the intensity values for mimicking photon randomness with variable levels of noise to mimic the real world challenges. Such a synthetic dataset forms noisy-clean pairs and, thereby, increases the model's robustness and performance across different lighting conditions.

$$\text{photon\_shot\_noise} = \text{Poisson}(\text{intensity}) \times \text{gain}$$

D. Image Transformation (RGB to RAW Conversion):

In low-light images, RGB processing tends to throw away critical details, severely impairing the model's ability to detect objects and patterns. To counteract this, we implemented an "unprocessing" technique that reverses the conventional processing path, returning the image to its RAW state without degrading high-fidelity details captured by the sensor. This preserves the crucial information, including the darker regions. The model will therefore continue to have access to a complete dataset for correctly identifying the essential features. We enhance the model's ability to identify important elements under low illumination conditions by reversing adjustments like gamma correction and white balance.

$$\text{image\_normalized} = \frac{\text{image}}{255.0}$$

$$\text{tone\_mapped\_image} = \text{smoothstep}(\text{image\_normalized}) \times 255$$

E. Gamma Compression and Inverse Gamma Compression:

It employs gamma compression on intensity for improving detail in shadowy regions so that the objects can be identified with much clarity. Smooth gradations of shadow and illumination ensure correct object identification in such environments of varying lights. The inverse gamma compression is applied at last to recover the contrast and brightness to their original state since this process maintains the features but brings forth the feature which was there but was not being viewed.

$$\text{gamma\_compression}(x) = \max(x, \epsilon)^{\frac{1}{2.2}}$$

F. Smoothstep and Inverse Smoothstep Transformation:

The smoothstep transformation enhances edge definition and clarity of low-light images by reducing step-like variations between intensity levels, which are abrupt pixel variations. It keeps the data fidelity while restoring its original structure. Therefore, it helps to preserve information significantly at edges, making detection outcomes much more accurate.

$$\text{smoothstep}(x) = 3 \cdot x^2 - 2 \cdot x^3$$

$$\text{inverse\_smoothstep}(y) = 0.5 - \sin\left(\arcsin(1-2y)^3\right)$$

G. Tone Mapping:

Understanding that dynamic range can be handled with ease in low-light images, tone mapping compresses the extreme brightness and enhances contrast. With the smoothstep function, this, in turn, permits good contrast, but when scaled to 8-bit space, lit areas remain clear. The preprocess of applying tone mapping is crucial for the model to capture features and objects in low-contrast scenes.

$$\text{tone\_mapped\_image} = \text{smoothstep}\left(\frac{\text{image}}{255}\right) \times 255$$

H. Matrix Inversion and Color Correction:

Applying the principle of matrix inversion would neutralize the unwanted color bias during RGB pipelines, especially during low-light conditions. Using an inverse combined matrix meant that we had reversed the original state of the RGB image corrected for color, thereby maintaining the original fidelity and quality of the image as it would have been before passing through the model like it was a raw capture while at the same time ensuring no loss of critical data during detection.

$$\left(\sum_{i=1}^{n} \text{random\_weights}[i] \cdot \text{color\_correction\_matrices}[i]\right)^{-1}$$

I. Model Training and Performance Metrics:

The models are trained on the prepared datasets that are optimized for noise reduction, contrast enhancement, and object detection: UNet, Autoencoder, DCNN, and ResNet. Qualitative performance metrics are based on PSNR and SSIM:

The Peak Signal-to-Noise Ratio (PSNR) is defined as:

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{L^2}{\text{MSE}} \right)$$

where L is the maximum pixel intensity and MSE is the Mean Squared Error. The Structural Similarity Index (SSIM), which evaluates structural similarity, is defined as:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

where μ_x and μ_y are the mean intensities, σ_x and σ_y are the variances, σ_{xy} is the covariance, and c_1 and c_2 are stability constants. Additionally, the following metrics were used for object detection assessment:

Accuracy: Accuracy is defined as the ratio of the correctly predicted positive and negative instances over the total number of predictions. Mathematically, accuracy can be represented as:

$$AS = \frac{TP + TN}{TP + FN + TN + FP}$$

Precision: Positive predictive accuracy is the precision, a ratio of true positive predictions over the total number of positive predictions. It gives the number of confident predictions the model has for positive instances and finds most use in applications where false positives are expensive. Precision is calculated using the following formula:

$$PS = \frac{TruePositives}{TruePositives + FalsePositives}$$

Recall: Recall, or true positive rate, measures the proportion of a model's ability to detect all relevant positive cases. It is given by the formula for the number of true positives over total actual positives. Recall is useful when capturing each case positive is of paramount importance-even if it results in false positives.

$$RS = \frac{TruePositives}{TruePositives + FalseNegatives}$$

F1 Score: The F1 score balances both precision and recall and is a useful measure for when the trade-off is required. It is actually the harmonic mean of precision and recall, and that penalizes extreme values, being given by:

$$F1 = \frac{2 * PS * RS}{PS + RS}$$

J. Evaluation and Comparative Analysis:

Performance on all these metrics has been tested for these models and points out their strengths and weaknesses in low-light urban scenes. A higher PSNR and SSIM value represents better enhancement quality, and high accuracy, precision, recall, and F1 score sign a strong object recognition capability. Optimal model identification for use in applications such as autonomous driving and urban monitoring can be done through comparative analysis.

Batch size and the learning rate: The training process included careful consideration of batch size and learning rate, both of which are essential for effective model training. A batch size of 32 was chosen to balance memory efficiency and stable gradient estimates, enabling the model to learn from a representative sample of the dataset. The learning rate was initially set to 0.001, scheduling the learning rate to drop by a factor of 0.1 after a predetermined number of epochs with little or no improvement in validation loss. This dynamic adjustment can fine-tune the learning process of the model.

Optimization Methods: To enhance the generalization capabilities of the models, during training, numerous optimization strategies were used in an effort to reduce overfitting. Data augmentation was used, including randomly applying transformations for further enhancing the generalization ability from the training data set. Furthermore, dropout layers are incorporated into the model architectures by randomly shutting off the neurons to reduce reliance on certain features. Early stopping was used so that training would stop when validation loss stopped improving. That way, the model stays generalizable to new data.

Hardware and Software: The training of the models was carried out in a high-performance computing environment, utilizing NVIDIA GeForce RTX 3080 GPUs to ensure rapid processing and effective results. This powerful hardware setup allows for handling large datasets and complex model architectures efficiently. The models were developed using the PyTorch framework, which provides the flexibility necessary for implementing deep learning strategies. In addition, libraries like NumPy and OpenCV supported numerical operations and image processing tasks, making the overall training workflow much smoother.

IV. Results

*a) **Enhancement Effects on Image Quality**:* Improving the aesthetic appeal of an image is just one benefit-such techniques enhance the performance of computer vision algorithms that are at the core of many applications, such as object detection and scene analysis. Techniques aimed at addressing inherent problems in low-light images produce a more favorable operating environment for models and consequently lead to better results for several applications. However, quantitative analysis using such as PSNR and SSIM metrics can be based in concrete terms to grade before and after enhancement qualities for images. For all different kinds of enhancement techniques involved within this paper, that includes tone mapping and gamma compression, average PSNR scores and SSIM values would be provided to further evidence the quality improvements over image enhancement. Visual comparisons through a series of images—featuring dark, enhanced, and final outputs—will further emphasize the differences in clarity, contrast, and detail. These side-by-side comparisons will annotate key improvements in object visibility and overall image clarity, underscoring the effectiveness of the enhancement techniques used.

*b) **Object Detection Performance**:* The impact of different enhancement techniques on the detection accuracy will be shown using a table or graph comparing the detection

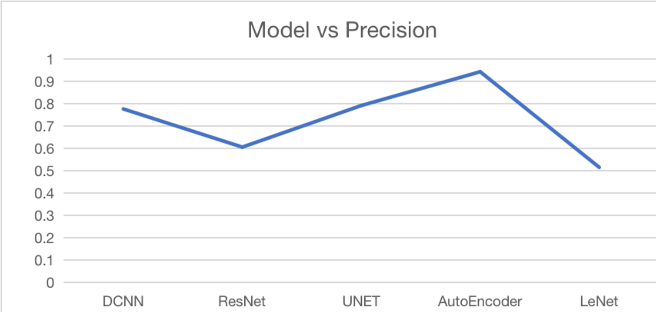

Fig. 1. Precision of Object Recognition

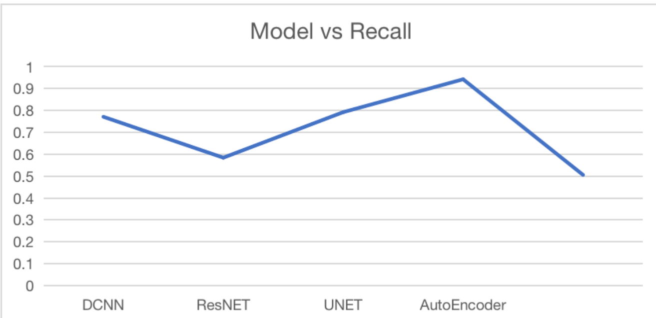

Fig. 2. Recall of Object Recognition

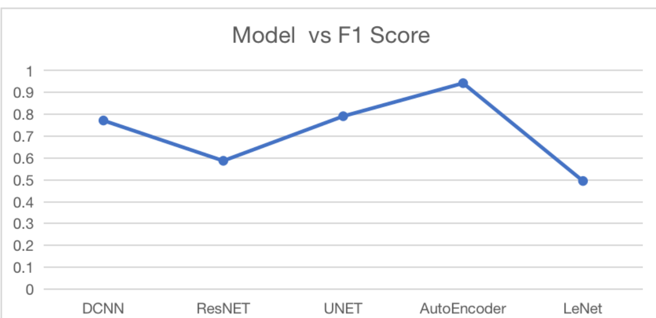

Fig. 3. F1 Score of Object Recognition

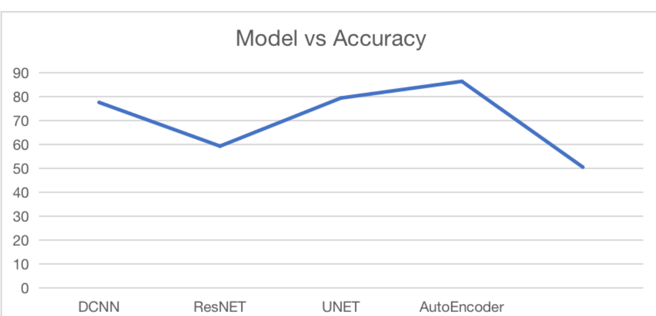

Fig. 4. Accuracy of Object Recognition

performance on non-enhanced (original) versus enhanced images. The analysis will discuss the effect of each enhancement technique on the accuracy of the detection, and if the enhancements have significantly improved identification

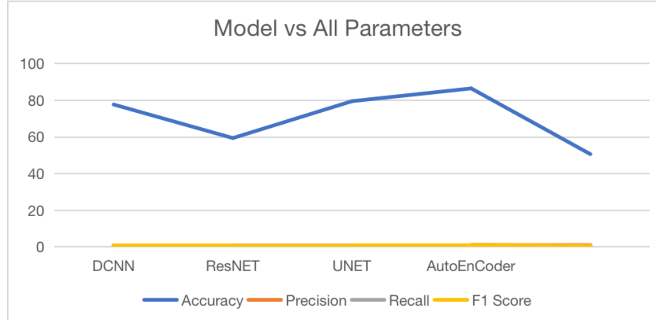

Fig. 5. Evaluation Metrics of Object Recognition

of objects in low illumination. A performance comparison among the different neural networks, U-Net, DCNN, ResNet, and Autoencoder, concerning their results of detection based on improved and non-enhanced images will also be performed. This comparison shall further be able to illustrate the best models with improvements in performance and a related justification with architectural strength that is possible to exploit enhanced features. To show model performance, precision-recall curves or confusion matrices will be used, thus representing a clearer view of effectiveness in object detection in low-light situations.

V. Conclusion and Future Work

In summary, the results of this study emphasize the importance of the several image enhancement techniques that improved object detection accuracy in low-light environments. Techniques like gamma compression and inverse white balance have been very useful in image preparation for better detection with models in such challenging conditions. The same architecture of each of the neural networks contributed differently in enhancing detection capabilities and revealed comparative advantages for selecting the appropriate model for specific tasks.

As future research directions, the enhancement techniques may be optimized to achieve the highest robustness and effectiveness. Alternative networks can help increase the detection performance more as well. Hybrid models based on the strengths of different models can be developed that seem to improve accuracy by far. Real-time improvement in detection and enhancement capabilities on embedded systems can provide avenues towards practical applications in a scope such as autonomous vehicle operations and surveillance. Appendices will contain supplementary information. These will include data set details, preprocessing scripts, and evaluation code besides charts, graphs, and sample outputs that will give better insight into the methodologies adopted in the current study.

References

[1] Z. Xiao, X. Zhou, and P. Ru. "Dark light enhancement for dark scene urban object recognition". In: *IET Image Processing* (2023). DOI: 10.1049/ipr2.12771.

[2] S. Panagiotou and A. S. Bosman. "Denoising diffusion post-processing for low-light image enhancement". In: *Pattern Recognition* 156 (2024), p. 110799.

[3] Y. Zhang, J. Zhang, and X. Guo. "Kindling the darkness: A practical low-light image enhancer". In: *Proceedings of the 27th ACM International Conference on Multimedia*. 2019, pp. 1632–1640.

[4] Y. Li, F. Cai, Y. Tu, and Y. Ding. "Low-light image enhancement under non-uniform dark". In: (2023), pp. 190–201.

[5] K. G. Lore, A. Akintayo, and S. Sarkar. "LLNet: A deep autoencoder approach to natural low-light image enhancement". In: *Pattern Recognition* 61 (2017), pp. 650–662.

[6] D. Ye et al. "Glow in the dark: Low-light image enhancement with external memory". In: *Pattern Recognition* ().

[7] L. Tao et al. "LLCNN: A convolutional neural network for low-light image enhancement". In: *2017 IEEE Visual Communications and Image Processing (VCIP)*. 2017, pp. 1–4.

[8] Y. Wang et al. "Low-light image enhancement with normalizing flow". In: *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 36. 3. 2022, pp. 2604–2612.

[9] I. Jeong and C. Lee. "An optimization-based approach to gamma correction parameter estimation for low-light image enhancement". In: *Multimedia Tools and Applications* 80 (2021), pp. 18027–18042.

[10] Y. Wang, Z. Liu, J. Liu, S. Xu, and S. Liu. "Low-light image enhancement with illumination-aware gamma correction and complete image modeling network". In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2023, pp. 13128–13137.

[11] S. Liu, W. Long, Y. Li, and H. Cheng. "Low-light image enhancement based on membership function and gamma correction". In: *Multimedia Tools and Applications* 81.16 (2022), pp. 22087–22109.

[12] J. J. Jeon, J. Y. Park, and I. K. Eom. "Low-light image enhancement using gamma correction prior in mixed color spaces". In: *Pattern Recognition* 146 (2024), p. 110001.

[13] Z. Ying, G. Li, Y. Ren, R. Wang, and W. Wang. "A new low-light image enhancement algorithm using camera response model". In: *Proceedings of the IEEE International Conference on Computer Vision Workshops*. 2017, pp. 3015–3022.

[14] F. Lv, Y. Li, and F. Lu. "Attention guided low-light image enhancement with a large-scale low-light simulation dataset". In: *International Journal of Computer Vision* 129.7 (2021), pp. 2175–2193.

[15] Y. Guo, Y. Lu, R. W. Liu, M. Yang, and K. T. Chui. "Low-light image enhancement with regularized illumination optimization and deep noise suppression". In: *IEEE Access* 8 (2020), pp. 145297–145315.

[16] Y. Zhang et al. "Self-supervised low-light image enhancement and denoising". In: *arXiv preprint arXiv:2103.00832* (2021).

[17] Y. Wu et al. "Learning semantic-aware knowledge guidance for low-light image enhancement". In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2023, pp. 1662–1671.

[18] L. Zhu, W. Yang, B. Chen, F. Lu, and S. Wang. "Enlightening low-light images with dynamic guidance for context enrichment". In: *IEEE Transactions on Circuits and Systems for Video Technology* 32.8 (2022), pp. 5068–5079.

[19] H. Wang, X. Yan, X. Hou, K. Zhang, and Y. Dun. "Extracting noise and darkness: Low-light image enhancement via dual prior guidance". In: *IEEE Transactions on Circuits and Systems for Video Technology* (2024).

[20] L. Ma, R. Liu, J. Zhang, X. Fan, and Z. Luo. "Learning deep context-sensitive decomposition for low-light image enhancement". In: *IEEE Transactions on Neural Networks and Learning Systems* 33.10 (2021), pp. 5666–5680.

[21] W. Yang, S. Wang, Y. Fang, Y. Wang, and J. Liu. "Band representation-based semi-supervised low-light image enhancement: Bridging the gap between signal fidelity and perceptual quality". In: *IEEE Transactions on Image Processing* 30 (2021), pp. 3461–3473.

[22] G. Zhai, W. Sun, X. Min, and J. Zhou. "Perceptual quality assessment of low-light image enhancement". In: *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)* 17.4 (2021), pp. 1–24.

[23] W. Lin et al. "No-reference quality assessment for low-light image enhancement: Subjective and objective methods". In: *Displays* 78 (2023), p. 102432.

[24] B. Chen et al. "The loop game: Quality assessment and optimization for low-light image enhancement". In: *arXiv preprint arXiv:2202.09738* (2022).