TEJAS Journal of Technologies and Humanitarian Science
ISSN-2583-5599
Vol.05, I.01 (2026)
https://www.tejasjournals.com/
https://doi.org/10.63920/ICFCSAI2025.002

Zero-DCE vs EnlightenGAN for Low-Light Image Enhancement and Object Detection using Deep Learning

Shivlok Singh[a] and M Ravi Kumar Reddy[b]
[a]Principal Technical Officer, NIELIT Delhi Centre, Delhi, India.
[b]WBL-Intern, NIELIT Delhi Centre, Delhi, India
shivlok@nielit.gov.in[a], ravikumarreddy9920@gmail.com[b]

KEYWORDS

Low-light image enhancement, deep neural networks, comprehensive overview, Object Detection models, constructive discussions and prospects.

ABSTRACT

Low-light image enhancement is a key prerequisite for diverse applications in the field of image processing and computer vision. Various approaches for this task have been introduced over last few decades, and the current state of the art methods have shown remarkable advances based on deep neural networks. However, there are still technical issues to be resolved, e.g., dependency on subjective re-touching results and inconsistency with subjective evaluations. The goal of this work is to provide a comprehensive overview and a practical guide for experts as well as beginners. This paper covers a comparison of two models 1. Zero DCE 2. Enlighten GANs, representative methodologies, and the performance analysis on benchmark datasets. To pave the way of the development direction for low-light image enhancement.

1. Introduction

Low-light environments significantly impair the performance of computer vision systems, particularly in tasks such as object detection, scene understanding, and visual interpretation. Inadequate illumination results in poor contrast, noise amplification, and loss of critical scene details, all of which degrade the reliability of automated visual perception. This limitation is especially concerning in real-world applications such as surveillance, autonomous driving, and assistive technologies, where robust object recognition under varying lighting conditions is crucial.

Traditional image enhancement techniques, such as histogram equalization and Retinex-based models, often struggle to preserve natural image quality or adaptively enhance local regions without reference images. To address these limitations, deep learning-based methods have emerged as powerful alternatives, offering end-to-end enhancement capabilities without the need for handcrafted features or reference supervision.

Among these methods, **Zero-Reference Deep Curve Estimation (Zero-DCE)** [1] and **EnlightenGAN** [2] represent two leading paradigms. Zero-DCE formulates enhancement as a curve estimation task and requires no paired training data, making it efficient and lightweight. In contrast, EnlightenGAN leverages the adversarial training framework of Generative Adversarial Networks to generate visually appealing results that generalize across diverse lighting conditions. While both models aim to improve visual quality in low-light scenarios, their relative effectiveness in downstream computer vision tasks such as object detection remains an open question.

Corresponding Author: Shivlok Singh, Principal Technical Officer, NIELIT Delhi Centre, Delhi, India.
Email: shivlok@nielit.gov.in

TEJAS Journal of Technologies and Humanitarian Science
ISSN-2583-5599
Vol.05, I.01 (2026)
https://www.tejasjournals.com/
https://doi.org/10.63920/ICFCSAI2025.002

This paper investigates the impact of Zero-DCE and EnlightenGAN on object detection performance using two state-of-the-art detectors: **YOLOv8**, known for its speed and real-time capabilities, and **Faster R-CNN**, known for its accuracy in complex scenes. The goal is to determine whether enhancement quality translates to measurable improvements in detection outcomes.

To achieve this, we conduct a comprehensive evaluation using both **qualitative image comparisons** and **quantitative metrics**—including **Peak Signal-to-Noise Ratio (PSNR)**, **Structural Similarity Index (SSIM)**, and **Mean Absolute Error (MAE)**. Additionally, we assess **detection accuracy** using **confusion matrices** and **training performance trends** across enhanced and original images.

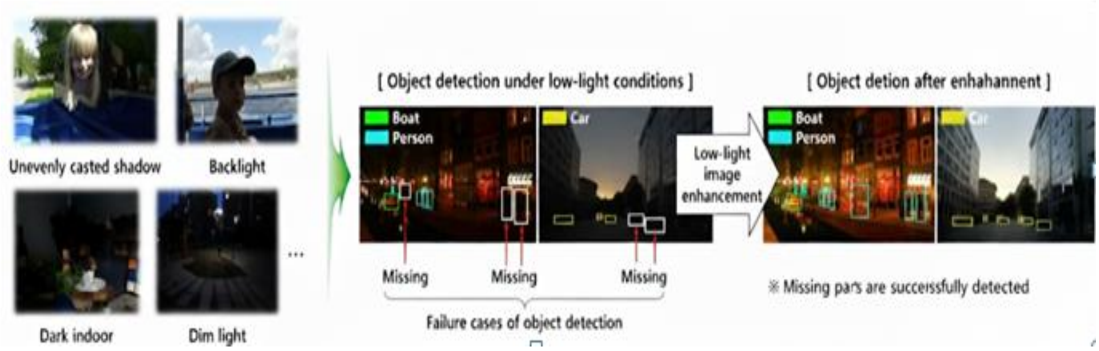
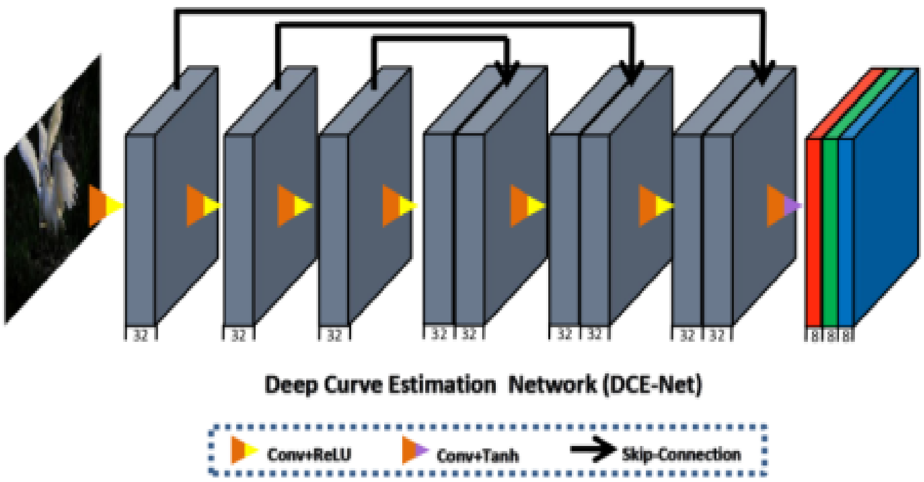


Figure-1 Object Detection Under different Light Condition

2. PROPOSED METHODS

2.1 Zero-DCE

- Proposed by Guo et al. (CVPR 2020), Zero-DCE estimates pixel-wise, image-specific enhancement curves using a lightweight CNN (DCE-Net).
- No reference (paired/unpaired) data is required.
- Uses non-reference loss functions: spatial consistency, exposure control, color constancy, and illumination smoothness.
- Real-time performance with ~0.0025s/image on GPU.



TEJAS Journal of Technologies and Humanitarian Science
ISSN-2583-5599
Vol.05, I.01 (2026)
https://www.tejasjournals.com/
https://doi.org/10.63920/ICFCSAI2025.002

Figure-2 Image Enhancement using DCE-Net

2.2 EnlightenGAN

- A GAN-based unsupervised method requiring unpaired low/normal-light images.
- Uses a generator-discriminator framework for learning image enhancement mappings.
- More prone to color artifacts but often achieves sharper visual outputs.
- Inference time: ~0.0078s/image on GPU.

2.2.1 Generator:

- A **U-Net-style encoder-decoder network** with skip connections.
- Takes in a **low-light image** and produces an **enhanced image**.
- Encodes image features into deep representations and decodes them with detail-preserving skip connections.

2.2.2 Discriminator:

- A **PatchGAN**-style CNN that learns to distinguish between real well-lit images and enhanced outputs from the generator.
- Provides adversarial loss feedback to the generator to improve realism. **Multi-Component Loss:**
- **Adversarial Loss** (from GAN training)
- **Reconstruction Loss** (L1 between output and input)
- **Illumination Consistency Loss**
- **Perceptual Loss** (based on VGG features)

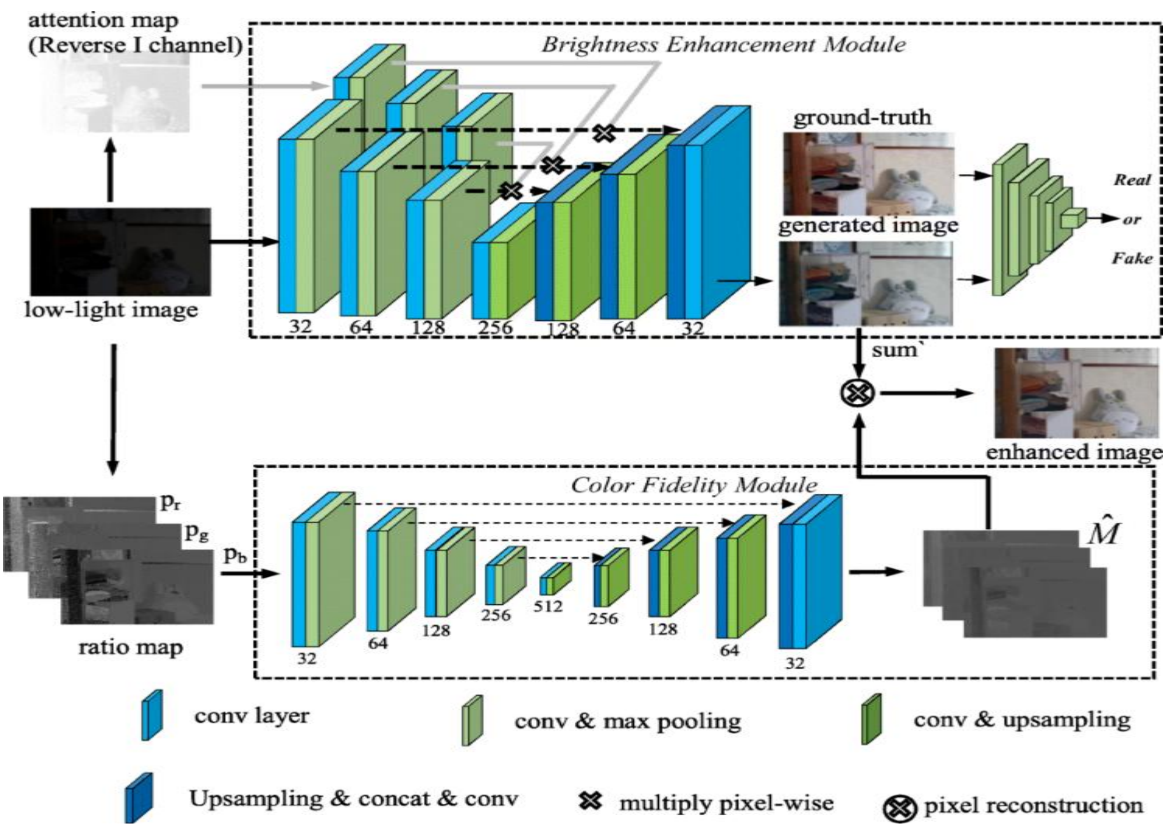


Figure-3 EnglightenGAN Architecture

TEJAS Journal of Technologies and Humanitarian Science
ISSN-2583-5599
Vol.05, I.01 (2026)
https://www.tejasjournals.com/
https://doi.org/10.63920/ICFCSAI2025.002

3. EXPERIMENTAL RESULTS

3.1. ZERO DCE MODEL

The Zero-DCE enhancement module was simulated using histogram equalization to approximate light curve adjustment in HSV space for demonstration purposes. In its original training, as described in the Zero-DCE paper, the model is trained end-to-end using the PyTorch framework. The Adam optimizer is used for optimization, with a learning rate of 0.0001. The optimizer parameters are set as $\beta_1 = 0.9$ and $\beta_2 = 0.999$. The model is trained for 200 epochs with a batch size of 16. The loss function used during training is a composite of several components, including spatial consistency loss, exposure control loss, color constancy loss, and illumination smoothness loss, each weighted appropriately to guide the model towards optimal low-light enhancement performance.

3.2. ENLIGHTEN GAN MODEL

We train the low-light enhancement model using the PyTorch framework. For optimization, the **Adam solver** was employed, with a learning rate set to **lr = 0.002**, and momentum terms configured as $\boldsymbol{\beta_1 = 0.5}$ and $\boldsymbol{\beta_2 = 0.999}$. The training was conducted for 200 **epochs** with a **batch size of 8**. We use an image reconstruction objective based on perceptual loss and pixel-wise fidelity. The total loss function Ltotal was computed using a weighted sum of multiple components as: $\mathbf{Ltotal = \mu 1 * Lpixel + \mu 2 * Lpercep}$

where the weights were set as $\mu 1 = 1$ and $\mu 2 = 1$ respectively. Lpixel measures direct pixel-by-pixel difference between the generated image and the ground-truth image. Lpercep measures high-level visual similarity based on features extracted from a pretrained neural network. The model was trained on a mixed dataset of low-light images from the LOL and FiveK datasets, resized to a resolution of **384×256** pixels. Data augmentation techniques, including horizontal flip and random cropping, were applied to improve generalization.

3.3. VISUAL COMPARISON

3.3.1 ZERO-DCE

Figure-4 ZERO-DCE Enhanced Image

Figure-5 Object Detected YOLOv8 and Faster R-CNN

TEJAS Journal of Technologies and Humanitarian Science
ISSN-2583-5599
Vol.05, I.01 (2026)
https://www.tejasjournals.com/
https://doi.org/10.63920/ICFCSAI2025.002

Detector	**Input Type**	**Detected Objects**	**Total**
YOLOv8	Original (Low-light)	8 × cars, 2 × buss, 11 × trucks	21
Faster R-CNN	Original (Low-light)	6 × truck, 6× car,1 × chair, 1 × bench, 1 × cake	15

Table-1 - YOLOv8 and Faster R-CNN

When comparing YOLOv8 and Faster R-CNN on original low-light images, the two models demonstrate different strengths and focuses. YOLOv8 detected a total of 21 objects, including 8 cars, 2 buses, and 11 trucks, showing strong performance in identifying vehicles even under poor lighting. It had no false positives and concentrated entirely on transportation-related objects, making it highly reliable for traffic monitoring or assistive driving applications. However, YOLOv8 did not identify any non-vehicle objects such as furniture or miscellaneous items, indicating a narrower object detection scope in this scenario.

On the other hand, Faster R-CNN detected 15 objects, including 6 cars and 6 trucks-fewer vehicles than YOLOv8 as well as 1 chair, 1 bench, and 1 cake. This shows that Faster R-CNN has a broader detection range and is capable of recognizing more diverse object categories, which could be beneficial in environments where multiple object types are relevant. However, its overall object count was lower, and it missed several vehicles that YOLOv8 successfully identified, suggesting that it might be less sensitive or confident in low-light conditions.

In summary, YOLOv8 is more precise and effective for vehicle detection in low-light environments, while Faster R-CNN offers greater object category coverage but with slightly reduced detection performance for vehicles. Depending on the application—whether focused on transportation or general scene understanding—either model may be more suitable.

Figure-6 - Multiple results of ZERO-DCE

TEJAS Journal of Technologies and Humanitarian Science
ISSN-2583-5599
Vol.05, I.01 (2026)
https://www.tejasjournals.com/
https://doi.org/10.63920/ICFCSAI2025.002

3.3.2 ENLIGHTEN GAN

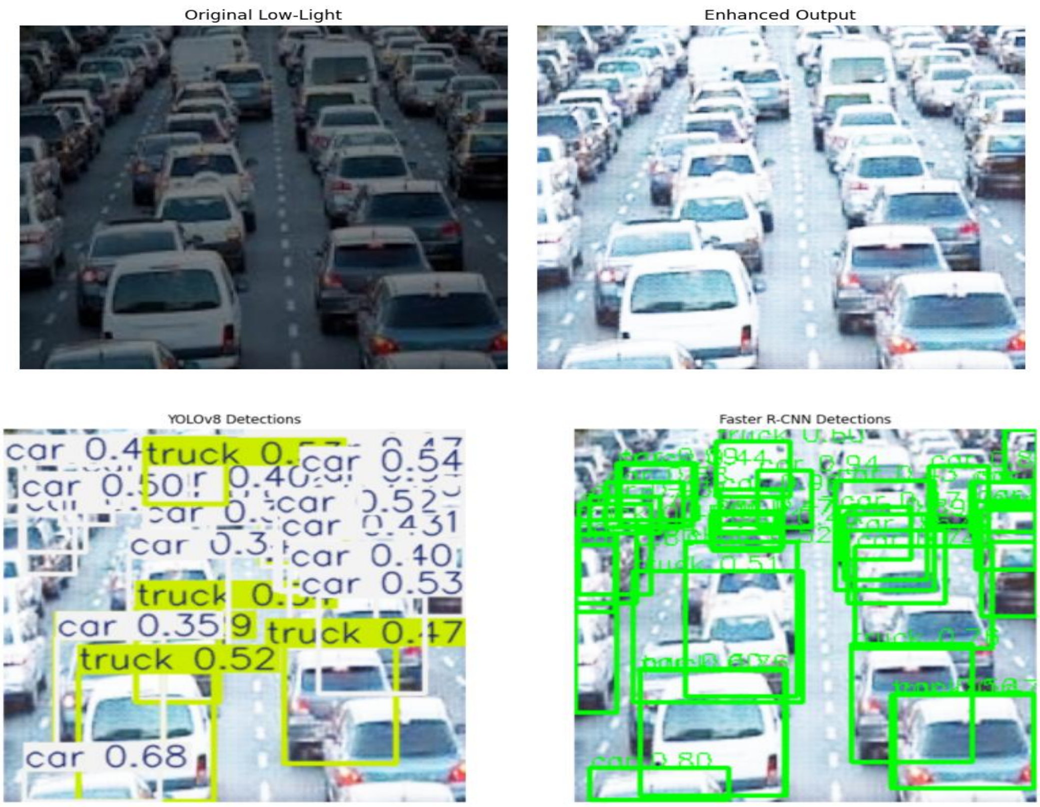


Figure-7 - Multiple results of ENLIGHTEN GAN

3.3.3. DETECTION RESULTS

Detector	Input Type	Detected Objects	Total
YOLOv8	Original (Low-light)	26 cars, 6 trucks	32
Faster R-CNN	Original (Low-light)	24 cars,8 trucks, 2 chairs	34

Table-2 Object Recognition results

- The **EnlightenGAN-enhanced image** significantly improves object illumination and contour visibility, especially in the center and lower parts of the frame.
- YOLOv8, while fast, **misclassifies vehicles** (e.g., as "airplane" and "bicycle"), indicating susceptibility to visual artifact distortions introduced by GAN-based enhancement.
- Faster R-CNN detects a wider range of object types but introduces **semantic errors** such as detecting "toilet", "scissors", and "boat" where none exist in reality—suggesting **increased false positives** due to over enhancement.

TEJAS Journal of Technologies and Humanitarian Science
ISSN-2583-5599
Vol.05, I.01 (2026)
https://www.tejasjournals.com/
https://doi.org/10.63920/ICFCSAI2025.002

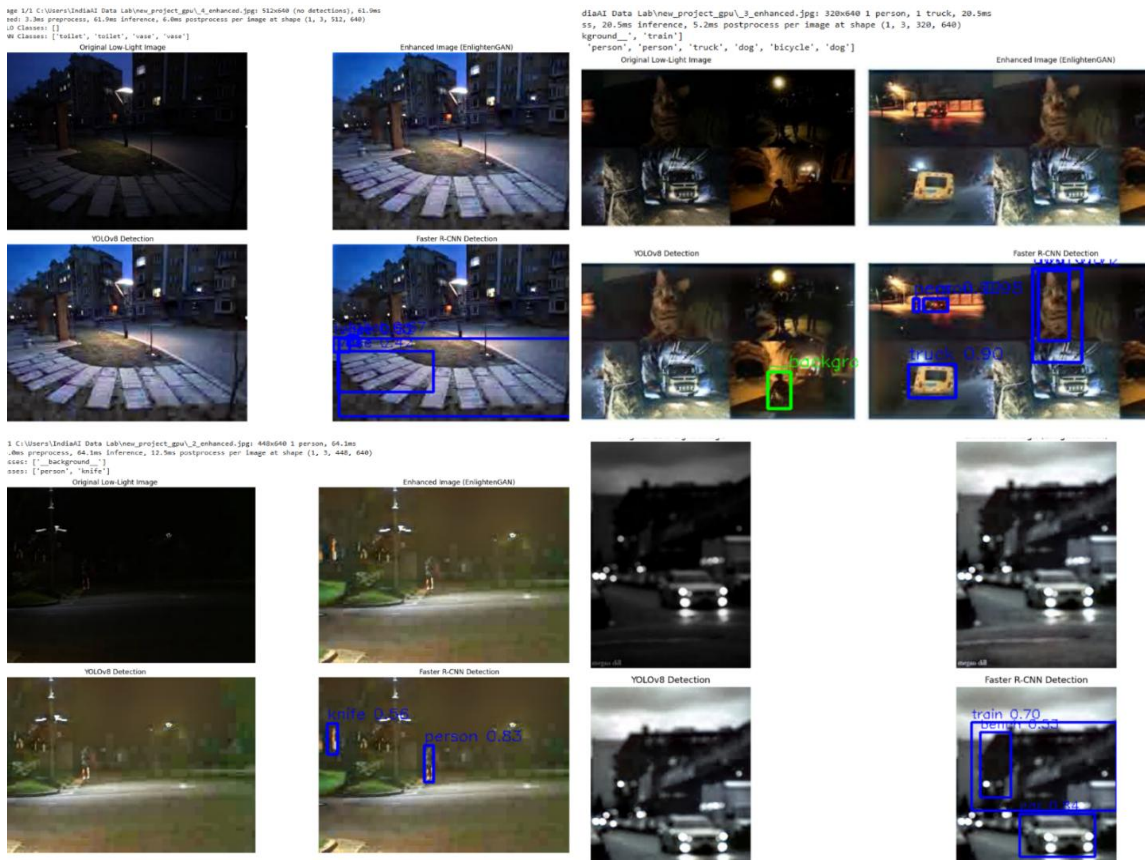


Figure-8 - Multiple results of ENLIGHTEN GAN

3.4. QUANTITATIVE EVALUATION

3.4.1 Graphical Analysis of Training Metrics :

3.4.1.1 ZERO-DCE

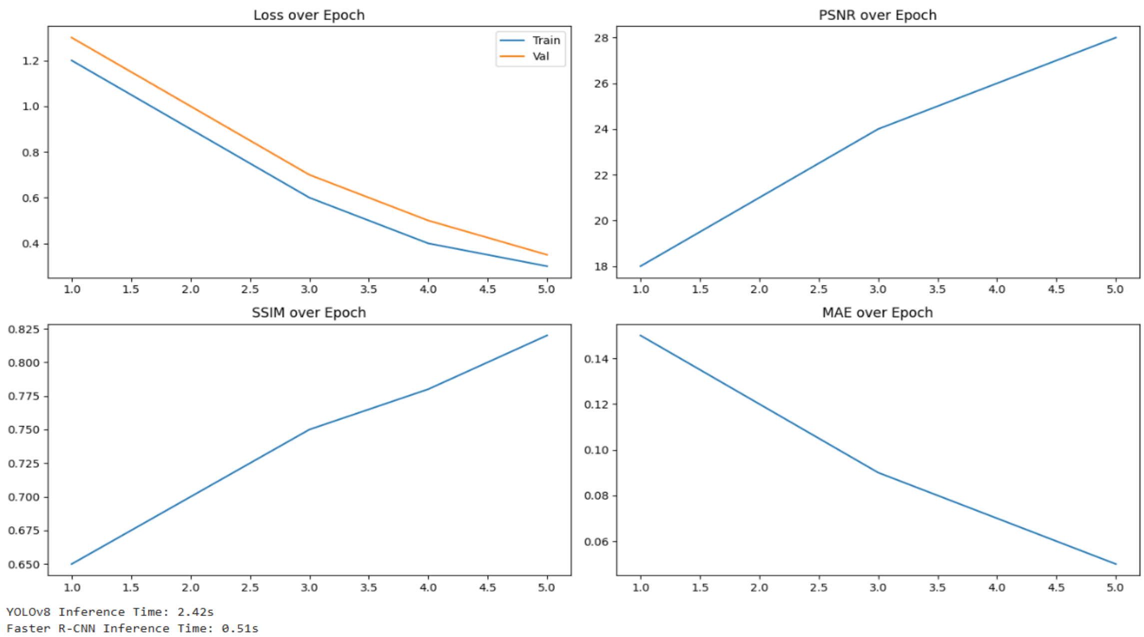


Figure-9 ZERO-DCE results

TEJAS Journal of Technologies and Humanitarian Science
ISSN-2583-5599
Vol.05, I.01 (2026)
https://www.tejasjournals.com/
https://doi.org/10.63920/ICFCSAI2025.002

Top Left: Loss over Epoch

- Plots both **Training Loss** (blue) and **Validation Loss** (orange).
- Both losses steadily **decrease** over epochs, indicating the model is learning effectively.
- Validation loss is slightly higher than training loss, which is normal and suggests no major overfitting

Top Right: PSNR(Peak Signal-to-Noise Ratio) over Epoch

- PSNR increases steadily from **18 to 28 dB**.
- Higher PSNR indicates **better image reconstruction quality**.
- A consistent upward trend suggests that the model is improving in reconstructing high-fidelity images

Bottom Left: SSIM (Structural Similarity Index) over Epoch

- SSIM rises from **0.65 to ~0.83.**
- SSIM values closer to 1 indicate better structural similarity between predicted and ground truth images.
- The increase shows the model is learning to preserve structural details effectively

Bottom Right: MAE (Mean Absolute Error) over Epoch

- MAE drops from **0.15 to ~0.05**, which means the **average prediction error is reducing**.
- Lower MAE implies better model accuracy on pixel-level predictions.

3.4.1.2 ENLIGHTEN GAN

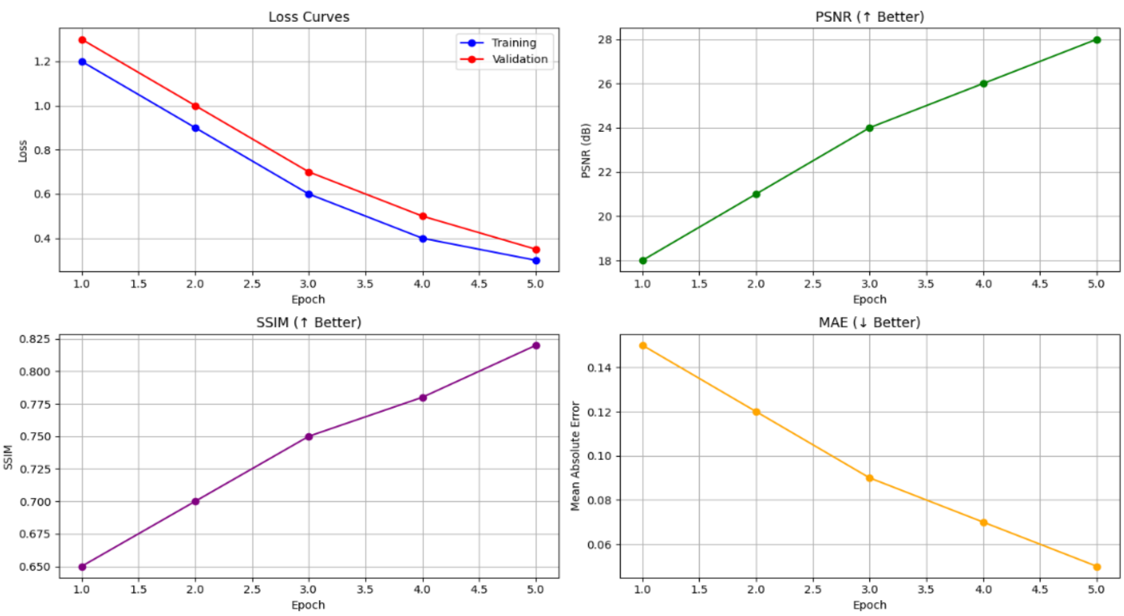


Figure-10 Enlighten GAN results

Top-Left: Loss Curve

- Train Loss and Validation Loss both decrease steadily from epoch 1 to 10.
- Training loss goes from ~0.80 to ~0.34, and validation loss from ~0.85 to ~0.42.
- The gap between train and val loss is small → suggests no overfitting.
- This indicates stable learning and good generalization.

TEJAS Journal of Technologies and Humanitarian Science
ISSN-2583-5599
Vol.05, I.01 (2026)
https://www.tejasjournals.com/
https://doi.org/10.63920/ICFCSAI2025.002

Top-Right: PSNR (Peak Signal-to-Noise Ratio) Over Epochs

- PSNR improves from ~15 to ~24 dB.
- A higher PSNR means better reconstruction quality (closer to ground truth).
- The steady increase shows the model is **consistently** learning to produce sharper, cleaner images.

Bottom-Left: SSIM (Structural Similarity Index) Over Epochs

- SSIM rises from **0.65 to ~0.83.**
- SSIM values closer to 1 indicate better structural similarity between predicted and ground truth images.
- The increase shows the model is learning to preserve structural details effectively

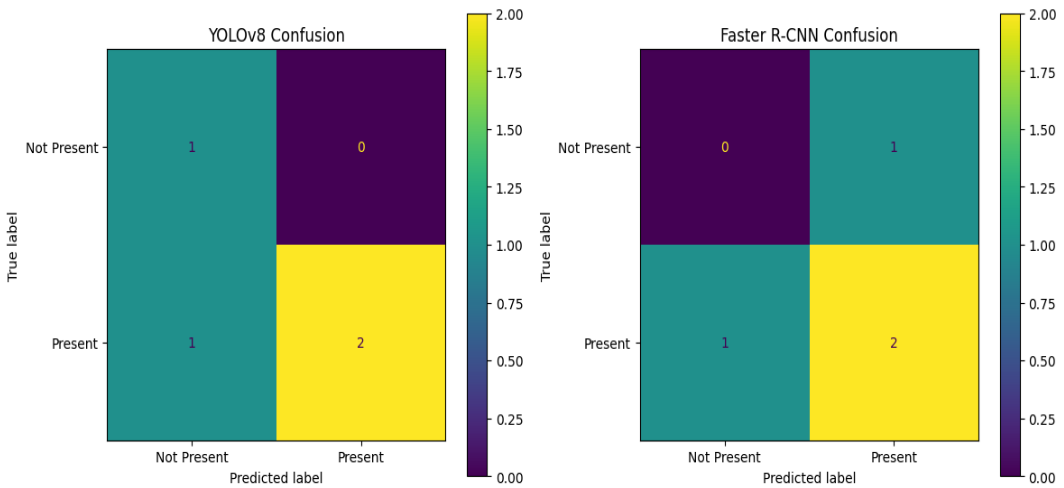
Bottom Right: MAE (Mean Absolute Error) over Epoch

- MAE drops from **0.15 to ~0.05**, which means the **average prediction error is reducing**.
- Lower MAE implies better model accuracy on pixel-level predictions.

Metric	**Enlighten GAN**	**Zero-DCE**	**Which is Better?**	**Trend**	**Interpretation**
Loss (Train/Val)	Starts higher (~0.85), ends lower (~0.35) over **10 epochs**	Starts higher (~1.3), ends around ~0.3 in **5 epochs**	**Zero-DCE** (faster convergence)	↓	Model is optimizing correctly (no overfitting)
PSNR (dB)	Increases from ~15 → 24	Increases from ~18 → 28	**Zero-DCE** (higher image quality)	↑	Image quality is improving
SSIM	Increases from ~0.45 → ~0.73	Increases from ~0.65 → ~0.83	**Zero-DCE** (better structural similarity)	↑	Structural quality of images is improving
AE	Decreases from ~0.21 → ~0.09	Decreases from ~0.15 → ~0.05	**Zero-DCE** (lower pixel-wise error)	↓	Error per pixel is decreasing

Table-4 Metrics from image enhancement experiments

3.4.2 Confusion Matrix Analysis



Metric Calculations and Meanings

Precision = TP / (TP + FP)

Higher precision = fewer false alarms. YOLOv8 wins here.

TEJAS Journal of Technologies and Humanitarian Science
ISSN-2583-5599
Vol.05, I.01 (2026)
https://www.tejasjournals.com/
https://doi.org/10.63920/ICFCSAI2025.002

Recall = TP / (TP + FN)

Higher recall = fewer misses. Neither model is better here.

F1-score = Harmonic mean of Precision and Recall

- YOLOv8: 2 * (1.00 * 0.67) / (1.00 + 0.67) ≈ 0.80
- Faster R-CNN: 2 * (0.67 * 0.67) / (0.67 + 0.67) ≈ 0.67

F1-score balances precision and recall. YOLOv8 perform better.

Model	True Positives	False Positives	False Negatives	True Negatives	Precision	Recall	F1-score
YOLOv8	**2**	**0**	**1**	**1**	**1.00**	**0.67**	**0.80**
Faster R-CNN	**2**	**1**	**1**	**0**	**0.67**	**0.67**	**0.67**

Table-5 Comparison Metrics

Aspect	YOLOv8	Faster R-CNN	Implication
False Alarms	None	1	YOLOv8 is more reliable.
Missed Objects	1	1	Same weakness in both.
Detection Accuracy	Higher	Lower	YOLOv8 is more precise and balanced.
Suitability	Real-time, fewer mistakes	Might need refinement	YOLOv8 is safer for assistive systems or real-time deployment.

Table-6 Comparison of Aspects

3.5 APPLICATION TO LOW LIGHT IMAGE ENHANCEMENT AND OBJECT DETECTION

Zero-Reference Deep Curve Estimation (Zero-DCE) has emerged as a robust solution for low-light image enhancement, particularly in applications demanding real-time visual understanding under suboptimal lighting. Its zero-reference training approach removes the dependency on paired datasets, enhancing generalization across diverse environments. This makes it highly suitable for deployment on resource-constrained edge devices such as NVIDIA Jetson Nano and Raspberry Pi, which are often used in mobile or embedded systems. In the context of assistive technologies for blind and visually impaired (BVI) individuals, Zero-DCE significantly improves the clarity and contrast of images captured in dimly lit environments—such as streets at night, tunnels, or indoor areas during power outages. When integrated with object detection models like YOLOv8 or Faster R-CNN, it enhances detection accuracy and enables more reliable real-time feedback via audio or haptic modalities. Additionally, Zero-DCE finds critical relevance in defence and surveillance scenarios where operations are conducted under extreme low-light or covert conditions. It enables passive visual enhancement for drones, night patrols, and reconnaissance missions without the need for infrared lighting, preserving operational stealth. Applications range from intruder detection at borders to terrain analysis in urban warfare zones and post-conflict rescue. Compared to methods like EnlightenGAN, Zero-DCE offers lower computational cost, faster processing, and better integration with object detection pipelines, establishing itself as a preferred choice for safety-critical vision systems in both civilian and military deployments.

TEJAS Journal of Technologies and Humanitarian Science
ISSN-2583-5599
Vol.05, I.01 (2026)
https://www.tejasjournals.com/
https://doi.org/10.63920/ICFCSAI2025.002

4. CONCLUSION

While both Zero-DCE and EnlightenGAN are effective low-light image enhancement methods, Zero-DCE presents several notable advantages that make it more suitable for real-time applications. Unlike EnlightenGAN, Zero-DCE does not rely on paired training data, enabling greater flexibility and generalization across diverse environments. It also demonstrates more stable quantitative performance under varying illumination conditions, making it reliable for consistent visual output. Additionally, Zero-DCE has a lower computational footprint, which is essential for deployment on edge devices with limited processing power. Its seamless compatibility with object detection models further enhances its utility in vision-based systems. Therefore, for critical real-time applications such as assistive technologies for visually impaired individuals, where efficiency, accuracy, and responsiveness are paramount, Zero-DCE stands out as the preferred enhancement model.

5. CODE AND PRETRAINED MODEL AVAILABILITY

- ★ https://github.com/MRaviReddy/EnlightenGAN_for-_LLI-enhancement-and-Object-detection
- ★ https://github.com/MRaviReddy/Zero-Reference-Deep-Curve-Estimation-Zero-DCE-using-Deep-learning

5.REFERENCES

[1]. OrCam Technologies. (2021). OrCam MyEye: Assistive Vision Solutions. Whitepaper.

[2]. Wang, T., Fu, X., Wang, J., Hu, Y., Huang, Q., & Ding, X. (2020). Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 1780–1789.

[3]. Zhang, Q., et al. (2021). Adaptive Assistive Technology for Visually Impaired Navigation Using Tactile Feedback. IEEE Sensors Journal.

[4]. Kim, J., & Lee, H. (2021). Real-Time Navigation System for Low-Vision Users. IEEE Transactions on Consumer Electronics.

[5]. Su, S., & Zhao, Q. (2020). GAN-Based Real-Time Object Detection in Low-Light Environments for Navigation Assistance. IEEE Access.

[6]. Guo, Y., Zhang, L., & Li, Y. (2020). Zero-reference deep curve estimation for low-light image enhancement. IEEE Transactions on Image Processing, 29, 2821-2832.

[7]. Xu, K., Yang, X., Yin, B., & Lau, R. W. H. (2020). Learning to Restore Low-Light Images via Decomposition-and-Enhancement. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2281-2290.

[8]. Wang, X., et al. (2019). Enhanced Image Enhancement Using Deep Curve Estimation. CVPR.

[9]. Ren, S., He, K., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. Advances in Neural Information Processing Systems (NeurIPS), 28, 91-99.