Journal of Research High School
Vol. 2026 (6) 501 – 507
ISSN 2688-3651 [online]

Improving Low-Light Object Detection Without More Data: A Pre-Processing Ablation for YOLOv11 Under Tight Memory Budgets

Annie Ye[1] *

[1]Princeton International School of Mathematics and Science, Princeton, NJ, USA
*Corresponding Author: annieyxy081018@gmail.com

Advisor: Chenliang Zhou, chenliang.zhou@qq.com

Received October 27, 2025; Revised February 13, 2026; Accepted February 25, 2026

Abstract

Object detection in low-light scenes is degraded by low signal-to-noise ratio, reduced local contrast, color cast, and motion blur. This study tests whether offline image pre-processing can improve low-light detection without increasing model size or data volume. Using YOLOv11-Nano as a fixed detector and ExDARK (7,363 images across 12 object classes) as the benchmark, a reproducible pipeline was built that includes robust ExDARK-to-YOLO conversion and four same-size pre-processing variants: unsharp sharpening, gamma-based brightening, CLAHE on luminance, and CLAHE followed by sharpening. All runs used identical hyperparameters (imgsz=512, batch=8, mosaic off, 5 epochs, AMP+half) under strict memory constraints. CLAHE improved accuracy relative to the baseline, with mAP@50 increasing from 0.277 to 0.290 and mAP@50–95 increasing from 0.171 to 0.179, while sharpening and global brightening underperformed. The findings indicate that conservative, domain-aware contrast enhancement can provide a compute-neutral gain for low-light detection.

Keywords: Low-light enhancement; Object detection; YOLOv11; CLAHE; ExDARK.

1. Introduction

Reliable object detection in low-light scenes is critical for intelligent transportation, robotics, mobile systems, and surveillance. One-stage detectors in the YOLO family have shown strong real-time performance in well-lit images (Redmon et al., 2016; Redmon & Farhadi, 2018), but night-scene benchmarks report clear accuracy drops under darkness (Loh & Chan, 2018; Neumann et al., 2018). The degradation is especially severe for small or distant objects in low illumination, where weak gradients and sensor noise reduce feature separability (Loh & Chan, 2018; Zheng et al., 2024).

A common strategy is to enhance images before detection. Prior low-light image enhancement methods include classical contrast operators and learning-based approaches such as Retinex-based models, Zero-DCE, and GAN-based pipelines (Pizer et al., 1990; Zuiderveld, 1994; Wei et al., 2018; Guo et al., 2020; Jiang et al., 2019; Ma et al., 2022). However, enhancement optimized for human perception can amplify noise, introduce artifacts, or alter color statistics in ways that do not align with detector optimization (Zheng et al., 2024; Ma et al., 2022). Joint enhancement-detection systems can reduce this mismatch, but they generally increase model complexity and compute requirements (Ma et al., 2022).

The research gap is that controlled evidence remains limited on whether lightweight offline pre-processing alone can improve low-light detection when detector architecture, parameter count, and dataset size are held fixed on ExDARK. Most prior studies focus on designing stronger enhancers rather than isolating the contribution of simple pre-processing under strict resource constraints (Guo et al., 2020; Jiang et al., 2019; Wei et al., 2018; Zhang et al., 2019).

This study addresses that gap by fixing the detector (YOLOv11-Nano), data volume, split, and training schedule, then comparing four low-cost photometric transforms in a reproducible ablation on ExDARK. The goal is to identify

Vol. 2026 (6) 501 – 507
ISSN 2688-3651 [online]

whether a practical pre-processing policy can deliver measurable gains without additional model parameters or data.

1.1 Related Work

ExDARK provides low-light images with object annotations and has become a standard benchmark for nighttime visual recognition analysis (Loh & Chan, 2018; Exclusively Dark (ExDARK) Image Dataset, 2018–). NightOwls complements this setting with real-world nighttime pedestrian detection and is widely used for night-scene evaluation (Neumann et al., 2018).

Classical CLAHE remains a practical baseline for local contrast recovery in dark regions (Pizer et al., 1990; Zuiderveld, 1994). Learning-based low-light enhancement methods include Zero-DCE, EnlightenGAN, Deep Retinex, and recent efficient variants (Guo et al., 2020; Jiang et al., 2019; Wei et al., 2018; Ma et al., 2022; Zhang et al., 2019). Recent surveys emphasize that enhancement quality for human viewing does not always translate to downstream task gains, motivating detector-oriented evaluation protocols (Zheng et al., 2024).

2. Materials and Methods

Figure 1 summarizes the workflow: ExDARK-to-YOLO conversion, frozen stratified split creation, preprocessing-variant generation, fixed-configuration training, and evaluation on a common test domain.

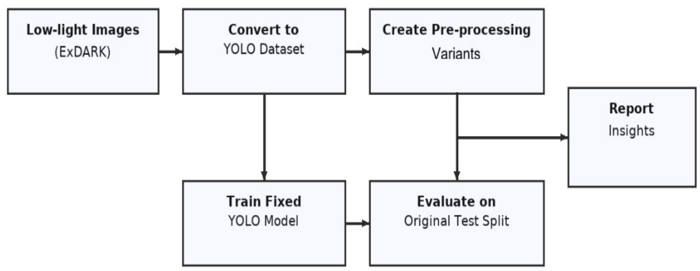


Figure 1. Overview of the end-to-end study pipeline.

2.1 ExDARK Dataset and Conversion Pipeline

The ExDARK dataset was used with 12 categories from real low-light scenes (Loh & Chan, 2018; Exclusively Dark (ExDARK) Image Dataset, 2018–). BBGt annotations were converted to YOLO format with a robust script that re-moves AppleDouble sidecar files (._*), matches labels by filename stem, and writes normalized labels in (class, x_c, y_c, w, h) format. A stratified 70/15/15 split with seed=42 was fixed once and reused for all runs. The resulting local copy contained 7,363 images and 7,364 label files after de-duplication.

Figure 2. Examples of the four preprocessing modules.

2.2 Low-Light Pre-Processing Modules

Four offline transforms were tested while preserving image count, labels, and split identity:

- SHARP: unsharp masking with Gaussian blur ($\sigma \approx 1.0$) and amount $\alpha \approx 1.0$–1.2.
- BRIGHT: gamma lookup table with $\gamma = 0.8$ and mild gain/bias (approximately 1.05 and +5).
- CLAHE: contrast-limited adaptive histogram equalization on the luminance channel in LAB space (clip limit 2.0, tile grid 8 by 8).
- COMBO: CLAHE followed by gentle unsharp masking.

2.3 Fixed Training Protocol

To isolate preprocessing effects, model capacity and optimization settings were held constant across runs. The detector was YOLOv11-Nano (yolo11n.pt, Ultralytics 8.3.199). Common settings were irngsz=512, batch=8, rnosaic=O, 5 epochs, and AMP+half precision.

Journal of Research High School

Vol. 2026 (6) 501 – 507
ISSN 2688-3651 [online]

2.4 Evaluation Protocol and Metrics

After training, each model was evaluated with Ultralytics val. Metrics were precision, recall, mean Average Precision at IoU 0.50 (mAP@50), and mean Average Precision averaged over IoU 0.50:0.95 (mAP@50–95). Run-level metrics were parsed from results.csv; confusion matrices and prediction visualizations were also reviewed.

In the reported experiments, preprocessing was applied to the training split used for weight updates, while evalution used the original unprocessed ExDARK test split through a common validation script. This was a deliberate stress test of robustness under a train-to-evaluation photometric distribution shift.

2.5 Reproducibility and Variants

Conversion and preprocessing scripts were organized so that file names, label indices, and split membership remained synchronized across variants. This enables direct retraining and controlled extensions under the same memory budget.

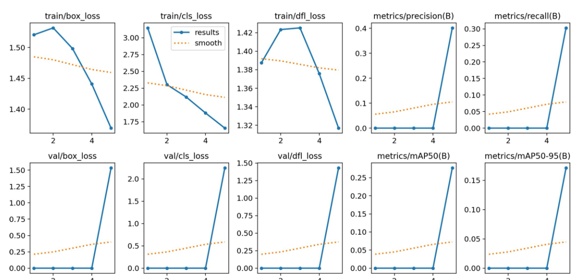
3. Results

3.1 Aggregate Quantitative Results

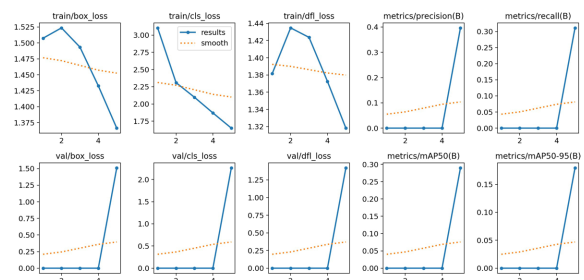
Table 1 shows that CLAHE produced the highest mAP@50 and mAP@50–95 among all variants. Relative to the unprocessed baseline, CLAHE increased mAP@50 by 0.013 and mAP@50–95 by 0.008. SHARP, BRIGHT, and COMBO did not exceed the baseline in aggregate mAP.

Table 1: Ablation summary on ExDARK (identical model and hyperparameters; 5 epochs; evaluation on original test split).

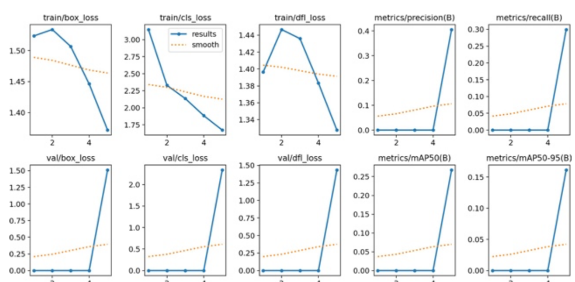
Run	Precision	Recall	mAP@50	mAP@50–95
base	0.401	0.303	0.277	0.171
sharp	0.406	0.300	0.268	0.161
bright	0.384	0.250	0.258	0.155
clahe	0.397	0.311	0.290	0.179
combo	0.367	0.280	0.271	0.165



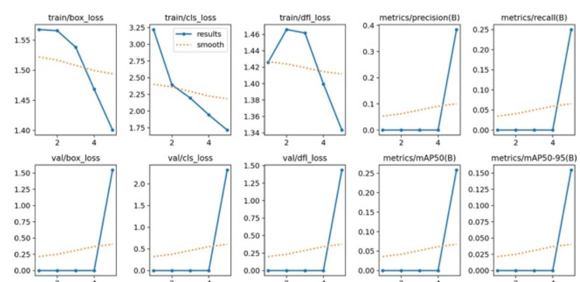
(a) Base



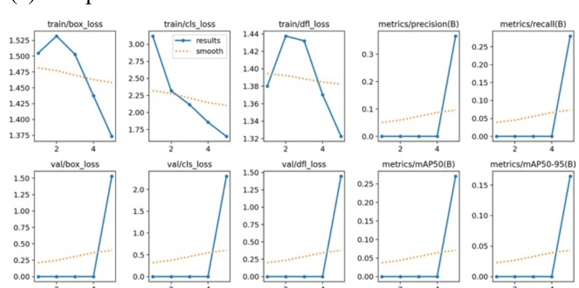
(b) CLAHE



(c) Sharpen



(d) Brighten

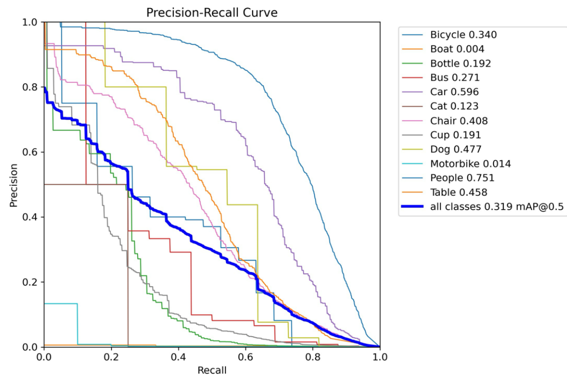


(e) Combo

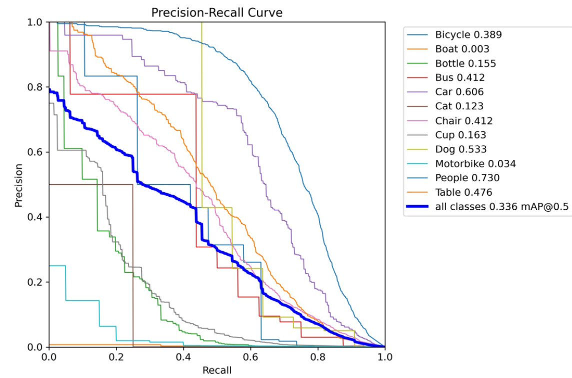
Figure 3. Training curves across variants under identical settings.

Vol. 2026 (6) 501 – 507
ISSN 2688-3651 [online]

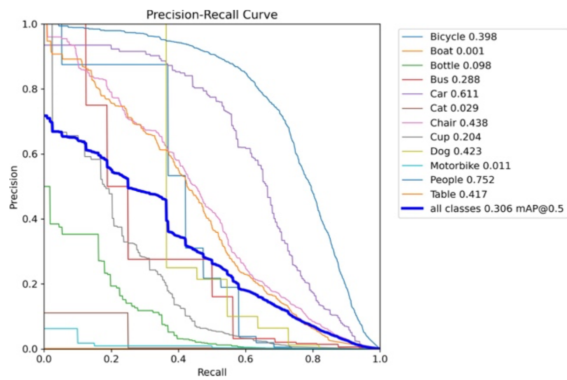
Figure 3 indicates similar optimization behavior across runs, with modest metric separation within the 5-epoch schedule. CLAHE curves remain slightly above baseline in validation metrics over late epochs.



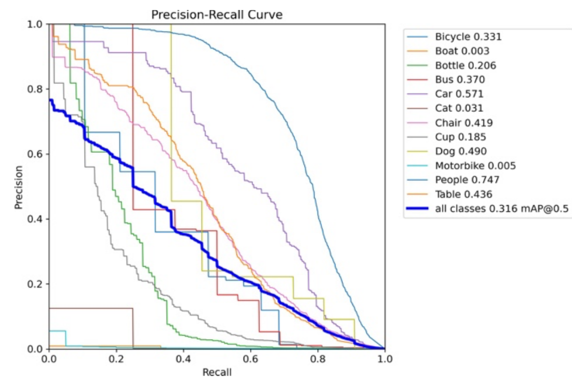
(a) Base



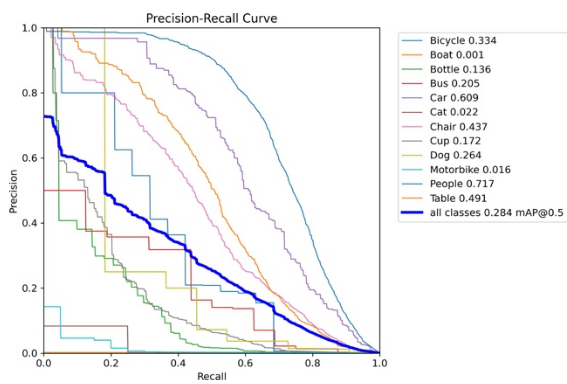
(b) CLAHE



(c) Sharpen



(d) Brighten



(e) Combo

Figure 4. Precision-recall curves for each variant.

Figure 4 shows larger precision-recall area for frequent classes and comparatively weaker curves for rare classes across all variants. The mean curve for CLAHE is higher than baseline over much of the recall range.

3.2 Class Distribution and Error Patterns

Figure 5 confirms a strongly imbalanced training distribution, with People and Car dominant and Boat and Motorbike relatively scarce.

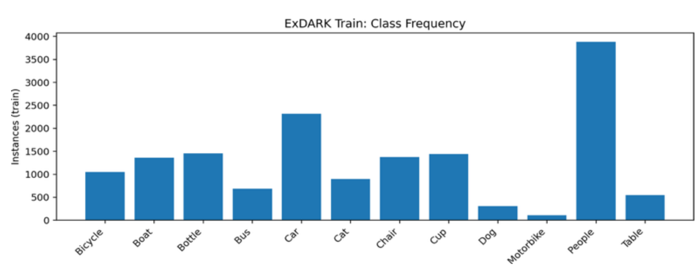
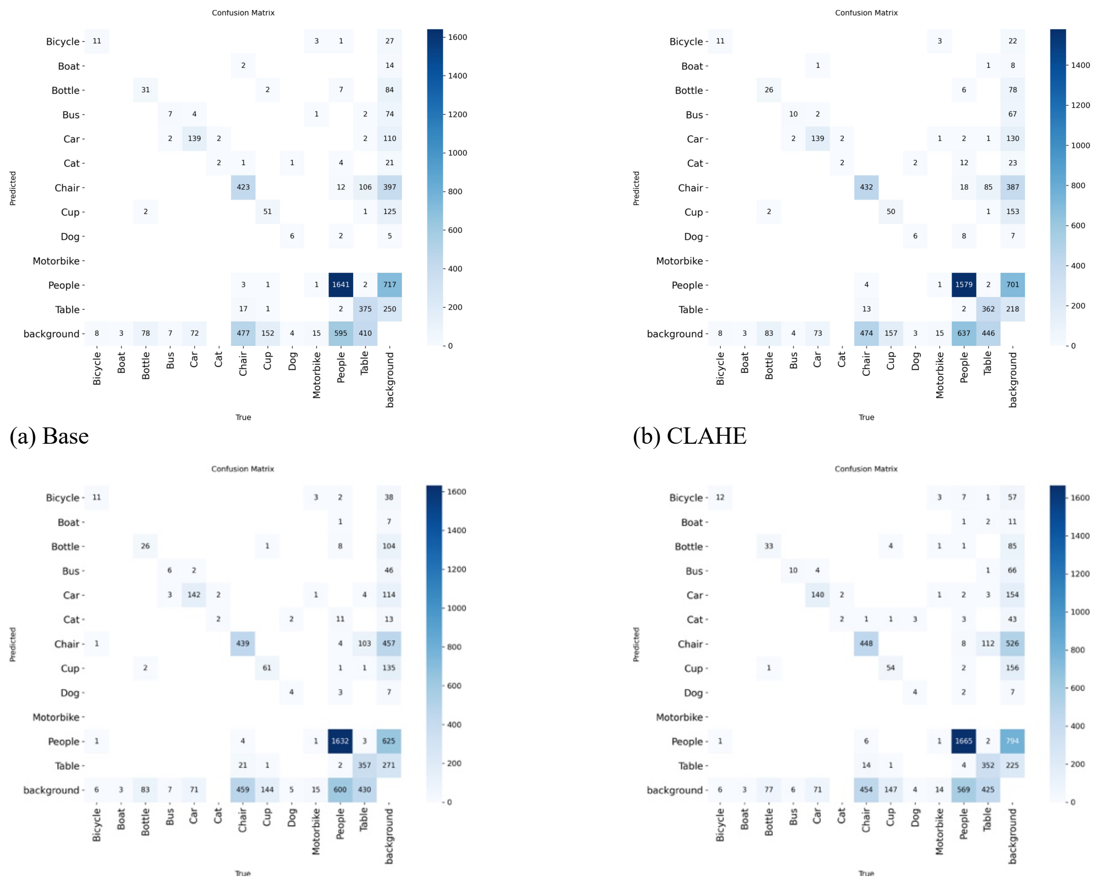


Figure 5. Class frequency in the training split.

Journal of Research High School

Vol. 2026 (6) 501 – 507
ISSN 2688-3651 [online]



(a) Base (b) CLAHE

(c) Sharpen (d) Brighten

Figure 6. Confusion matrices (part 1 of 2).

Figures 6 and 7 show concentration on diagonal entries for frequent classes and broader off-diagonal scattering for rarer classes. Common confusions include Bottle-Cup, Bicycle-Motorbike, and Chair-Table.

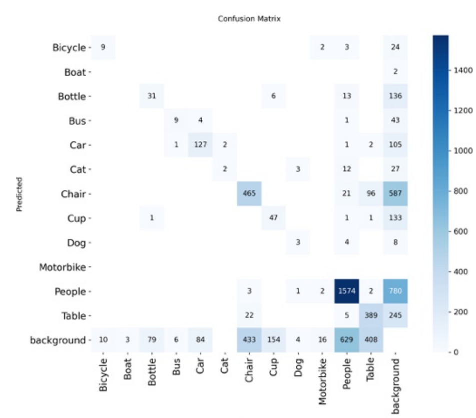


Figure 7. Confusion matrices (part 2 of 2): COMBO.

3.3 Qualitative Comparisons

Figure 8 compares baseline and CLAHE detections on identical scenes. Additional true positives and cleaner localization are observed in several dim, high-glare, or high-shadow regions for the CLAHE model.

4. Discussion

Under fixed model capacity and fixed data volume, CLAHE on luminance provided the most consistent aggregate improvement. The result supports the use of conservative local contrast normalization when retraining detector weights for low-light domains.

The train-to-evaluation mismatch in this design was intentional: models were trained with preprocessed inputs but evaluated on unprocessed ExDARK test images. This stress test indicates robustness to moderate distribution shift, but it does not measure the maximum performance achievable when train and inference preprocessing are matched.

Several limitations remain. First, the study used a single random seed (42), so variance across random initializations was not estimated. Second, training used a short 5-epoch schedule and one detector scale (YOLOv11-Nano).

Third, class imbalance in ExDARK can increase per-class metric instability, especially for rare categories. Because multi-seed reruns were not completed in the current resource window, these results should be treated as preliminary and need confirmation with repeated trials.

(a) Base predictions. (b) CLAHE predictions

Figure 8. Qualitative detections on identical test images at a fixed score threshold.

5. Conclusion

This study performed a controlled pre-processing ablation for low-light detection on ExDARK with YOLOv11-Nano under tight memory limits. Among four tested transforms, CLAHE on luminance produced the best aggregate perfor-mance, improving mAP@50 from 0.277 to 0.290 and mAP@50–95 from 0.171 to 0.179 without adding parameters or data. The findings are most applicable to lightweight one-stage detectors trained under similar compute and memory constraints. Given the single-seed, short-schedule setting, conclusions should be interpreted as indicative rather than definitive.

6. Reproducibility (Commands)

Training commands were identical across variants except for the data= path.

```
# Baseline
yolo detect train model=yolo11n.pt data=exdark_yolo/exdark.yaml \
  imgsz=512 batch=8 mosaic=0 workers=0 half=True amp=True \
  deterministic=False auto_augment='' erasing=0.0 copy_paste=0.0 \
  epochs=5 val=False cache=False

# Example: CLAHE variant
DATA=/path/to/exdark_yolo_clahe/exdark.yaml
yolo detect train model=yolo11n.pt data=$DATA \
  imgsz=512 batch=8 mosaic=0 workers=0 half=True amp=True \
  deterministic=False auto_augment='' erasing=0.0 copy_paste=0.0 \
  epochs=5 val=False cache=False
```

Data and Code Availability

To keep the main text concise and align with journal policy, conversion and preprocessing scripts are available from the corresponding author and will be posted in a public repository for camera-ready submission.

Journal of Research High School
Vol. 2026 (6) 501 – 507
ISSN 2688-3651 [online]

References

Chen, C., et al. (2018). Learning to See in the Dark. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 3291–3300. https://openaccess.thecvf.corn/content_cvpr_2O18/ht rnl/Chen_Learning_to_See_CVPR_2O18_paper.htrnl.

Exclusively Dark (ExDARK) Image Dataset. (2018–). GitHub repository (cs-chan). https://github.corn/cs-cha n/Exclusively-Dark-Irnage-Dataset.

Guo, C., Li, C., Guo, J., et al. (2020). Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. https://openaccess.thecvf.corn/content_CVPR_2O2O/papers/Guo_Zero-Reference_Deep_Curve_Estirnation_for_Low-Light_Irnage_Enhancernent_CVPR_2O2O_paper.pdf.

Jiang, Y., et al. (2019). EnlightenGAN: Deep Light Enhancement without Paired Supervision. arXiv. https://arxiv.org/abs/19O6.O6972.

Loh, Y. P., & Chan, C. S. (2018). Getting to Know Low-Light Images with the Exclusively Dark Dataset. arXiv. https://arxiv.org/abs/18O5.11227.

Ma, L., et al. (2022). Toward Fast, Flexible, and Robust Low-Light Image Enhancement. Proceed-ings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 5637–5646. https://openac cess.thecvf.corn/content/CVPR2O22/papers/Ma_Toward_Fast_Flexible_and_Robust_Low-Light _Irnage_Enhancernent_CVPR_2O22_paper.pdf.

Neumann, L., et al. (2018). NightOwls: A Pedestrians at Night Dataset. Proceedings of the European Conference on Computer Vision. https://www.robots.ox.ac.uk/~vgg/publications/2O18/N eurnann18b/neurnann18b.pdf.

Pizer, S. M., et al. (1990). Contrast-Limited Adaptive Histogram Equalization. UNC Computer Science Technical Report TR90-035. https://www.cs.unc.edu/techreports/9O-O35.pdf.

Redmon, J., et al. (2016). You Only Look Once: Unified, Real-Time Object Detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. https://www.cv-foundat ion.org/openaccess/content_cvpr_2O16/papers/Redrnon_You_Only_CVPR_2O16_paper.pdf.

Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv. https://arxiv.org/abs/18O4 .O2767.

Wei, C., et al. (2018). Deep Retinex Decomposition for Low-Light Enhancement. Proceedings of the British Machine Vision Conference. https://arxiv.org/abs/18O8.O456O.

Zhang, Y., Zhang, J., & Guo, X. (2019). Kindling the Darkness: A Practical Low-Light Image Enhancer. Proceedings of the ACM International Conference on Multimedia, 1632–1640. https://dl.acrn.org/doi/1O.1145/334 3O31.335O926.

Zheng, S., et al. (2024). Low-Light Image and Video Enhancement: A Comprehensive Survey and Beyond. *arXiv*. https://arxiv.org/htrnl/2212.1O772v5.

Zuiderveld, K. (1994). Contrast Limited Adaptive Histogram Equalization. In *Graphics Gems IV*. https://theswiss bay.ch/pdf/Gentoornen%2OLibrary/Garne%2ODevelopment/Prograrnrning/Graphics%2OGerns%2O4.pdf.