American Journal of Advanced Multidisciplinary Innovation and Research (AJAMIR)

E-ISSN: XXXX-XXXX ● Website: www.ajamir.com ● Email: editor@ajamir.com

Robust Computer Vision Under Low-Light and Low-Bandwidth Conditions

Dr. Lucas M. Anderson

Associate Professor, KTH Royal Institute of Technology, Sweden

Abstract

Computer-vision systems deployed in surveillance, intelligent transportation, remote monitoring, robotics, autonomous platforms, environmental sensing, and emergency-response environments frequently operate under conditions that differ substantially from conventional vision benchmarks. Two particularly important constraints are inadequate illumination and limited communication bandwidth. Low-light imaging reduces signal-to-noise ratio, contrast, color fidelity, edge visibility, and recoverable object detail, while bandwidth limitations force aggressive compression, reduced spatial or temporal resolution, selective frame transmission, or partitioned edge-cloud inference. ExDark contains 7,363 low-light images spanning ten illumination conditions and twelve object categories, while the DarkVision benchmark further demonstrates the importance of evaluating machine perception across multiple illumination levels and camera systems rather than assuming that conventional daylight performance generalizes to photon-limited environments.

This paper proposes an integrated methodological framework for robust computer vision under simultaneous low-light and low-bandwidth constraints. The framework combines illumination-aware sensing, noise-sensitive enhancement, task-oriented feature preservation, adaptive compression, edge-cloud split inference, confidence monitoring, and degradation-aware fallback policies. Existing research provides evidence that lightweight low-light enhancement can improve dark-scene representations and that feature compression optimized directly for downstream detection can outperform conventional image-compression approaches at constrained transmission rates. A transparent simulation compares four pipeline strategies: a naïve baseline, low-light enhancement followed by conventional compression, task-aware feature compression, and a joint low-light/bandwidth-aware pipeline. Across simulated stress conditions, the mean robustness score increases from 50.83 for the baseline to 79.17 for the integrated strategy. These scores are explicitly illustrative and do not represent measurements from a real deployed system. The study argues that robustness should be optimized jointly across sensing, enhancement, representation, transmission, inference, and uncertainty management rather than treating illumination restoration and bandwidth reduction as independent engineering problems.

Keywords: low-light computer vision, bandwidth-constrained inference, robust object detection, image enhancement, edge computing, feature compression, split computing, task-aware compression, visual perception, reliable artificial intelligence

American Journal of Advanced Multidisciplinary Innovation and Research (AJAMIR)
E-ISSN: XXXX-XXXX ● Website: www.ajamir.com ● Email: editor@ajamir.com

1. Introduction

Computer-vision systems increasingly leave controlled laboratory settings and operate in environments where image formation and communication resources are highly variable. Nighttime transportation monitoring, remote infrastructure inspection, wearable assistance, mobile robotics, disaster response, autonomous navigation, wildlife monitoring, and security applications can require visual inference when illumination is weak and network connectivity is unstable or expensive. Under low illumination, cameras receive fewer photons, increasing the relative influence of sensor noise and reducing the visibility of texture, boundaries, and object-specific information. The DarkVision benchmark was developed specifically around photon-limited perception and includes registered bright-dark observations collected at multiple illuminance levels with different cameras, illustrating that low-light degradation can directly challenge downstream detection and recognition.

The low-light problem has traditionally been approached through enhancement. Methods such as Zero-Reference Deep Curve Estimation treat enhancement as a learned adjustment of image-specific intensity curves and demonstrate that lightweight networks can improve poorly illuminated images without requiring conventional paired-reference supervision. ExDark provides an important complementary benchmark because it focuses explicitly on visible-light scenes ranging from very dark environments to twilight and supplies object-level annotations suitable for studying recognition and detection under realistic illumination degradation. More recent low-light detection research has increasingly moved beyond enhancement for visual appearance alone and toward representations that preserve information important for downstream tasks. FeatEnHancer, for example, reports task-driven improvements across low-light object detection, face detection, semantic segmentation, and video detection benchmarks, indicating that perceptual enhancement and machine-perception enhancement should not automatically be treated as equivalent objectives.

Low-light and bandwidth constraints are commonly investigated separately, yet real systems frequently experience them simultaneously. Enhancing a dark frame may amplify noise and increase high-frequency information that is expensive to encode. Aggressive compression can subsequently remove enhanced details, while transmitting unenhanced dark imagery can waste bitrate on noise patterns that contribute little to downstream inference. More recent collaborative-perception research explicitly treats communication efficiency as part of perception-system design, and CVPR 2026 work such as WhisperNet continues this movement toward bandwidth-efficient collaboration rather than raw-data exchange. These interactions suggest that the correct engineering objective is not simply maximum image quality or minimum bitrate, but maximum task reliability per available sensing and communication resource.

The present paper develops a unified methodological framework for this combined problem. Rather than claiming results from an unavailable experimental dataset, the study uses a transparent simulation to demonstrate how a robust vision pipeline could be evaluated across varying illumination and bandwidth conditions. The proposed architecture integrates low-light processing with downstream-task objectives, rate-adaptive communication, edge-cloud partitioning, confidence-aware inference, and degradation monitoring. Its central proposition is that robust vision under resource constraints requires coordinated optimization across the complete perception pipeline.

American Journal of Advanced Multidisciplinary Innovation and Research (AJAMIR)

E-ISSN: XXXX-XXXX ● Website: www.ajamir.com ● Email: editor@ajamir.com

2. Objectives of the Study

2.1 To Identify the Joint Reliability Effects of Illumination and Bandwidth Degradation

The first objective is to examine low-light degradation and limited transmission capacity as interacting rather than independent sources of computer-vision failure. The study considers how photon scarcity, sensor noise, contrast loss, reduced spatial resolution, lossy compression, dropped frames, and compressed intermediate representations can jointly affect detection reliability.

The objective also distinguishes human-perceived image quality from machine-perception utility. An enhanced image may appear visually brighter while still suppressing or distorting features required by a downstream detector.

2.2 To Develop a Task-Oriented Robust Vision Pipeline

The second objective is to propose a modular computer-vision architecture capable of adapting simultaneously to illumination level and communication availability. The framework combines low-light normalization, denoising, feature preservation, rate adaptation, edge processing, split inference, and confidence-based fallback behavior.

The underlying principle is that the system should preserve information according to its downstream value rather than allocating equal transmission resources to every pixel or feature channel.

2.3 To Establish a Reliability-Oriented Evaluation Strategy

The third objective is to define how such systems should be evaluated. Instead of reporting accuracy alone, the paper incorporates robustness across repeated degraded conditions, bandwidth efficiency, latency, calibration, worst-condition performance, and failure detectability.

This wider evaluation approach is particularly important because a system that performs well on average but fails silently under extreme darkness or sudden bandwidth collapse may be unsuitable for safety-relevant deployment.

3. Review of Literature

3.1 Low-Light Perception and Task-Oriented Enhancement

Low-light computer vision differs fundamentally from ordinary illumination adjustment because information may be absent or heavily contaminated by noise rather than merely displayed too darkly. ExDark was introduced by Loh and Chan to support systematic investigation of low-light recognition and detection and contains 7,363 images across ten illumination categories and twelve object classes. DarkVision extends benchmark realism through multi-illumination, multi-camera image and video acquisition and explicitly connects image restoration with downstream object detection. Together, these datasets demonstrate that low-light robustness should be evaluated across a spectrum of exposure and sensor conditions rather than through one artificial darkness transformation.

Low-light image enhancement methods have progressed from conventional histogram operations toward learned decomposition, curve estimation, and task-aware representation learning. Guo and colleagues' Zero-DCE formulates enhancement as image-specific curve estimation and uses a lightweight network that does not depend on the conventional paired-reference training structure. Retinex-inspired methods have similarly attempted to disentangle illumination and reflectance, while

American Journal of Advanced Multidisciplinary Innovation and Research (AJAMIR)

E-ISSN: XXXX-XXXX ● Website: www.ajamir.com ● Email: editor@ajamir.com

more recent work explores latent disentanglement and task-sensitive features. These developments are relevant to constrained deployment because lightweight or zero-reference approaches can reduce data-collection and compute requirements.

3.2 Bandwidth-Constrained Visual Inference

The bandwidth problem emerges when visual sensing occurs on resource-constrained edge devices while computationally expensive inference is performed partly or entirely elsewhere. Sending raw or conventionally compressed images is straightforward but can be inefficient because conventional codecs are designed primarily around perceptual reconstruction rather than downstream machine interpretation.

Yuan and colleagues proposed rate-constrained feature compression for edge-assisted object detection. Their split-computation architecture compresses intermediate YOLO features and jointly trains compression with the detection network, achieving stronger detection accuracy than image-compression baselines at constrained rates.

Cohen, Choi, and Bajić developed lightweight feature-tensor compression for collaborative intelligence and reported compression of floating-point activations to well below one bit per value with limited accuracy degradation on tested models. These findings support transmitting task-relevant representations rather than reconstructed images when communication resources are scarce.

3.3 Joint Robustness Under Compound Degradation

Low-light and bandwidth constraints interact through the information content of the signal. Enhancement can expose useful edges and contrast but can also amplify sensor noise. Compression may suppress noise but can simultaneously erase small objects, boundaries, or low-contrast features. A pipeline optimized independently at each stage may therefore produce locally reasonable results while decreasing end-to-end perception reliability.

A more robust system should optimize the combined objective:

[\max_{\theta,\phi,\pi} ; \mathcal{R}_{vision}]

subject to:

[B(\phi,\pi) \leq B_{\max}]

and:

[C(\theta,\phi,\pi) \leq C_{\max}]

where (\theta) represents low-light enhancement parameters, (\phi) represents the task model and feature representation, (\pi) represents compression or communication policy, (B_{\max}) is the available bandwidth, and (C_{\max}) is the edge-computation budget.

The objective should include detection or recognition performance rather than only reconstruction metrics. This direction is compatible with semantic and task-oriented compression research designed for machine vision. More recent bandwidth-efficient collaborative systems further demonstrate that communication architecture and task representation can be co-designed rather than treated as independent components.

4. Research Methodology

4.1 Research Design

The study adopts a conceptual-methodological design supported by simulation-based robustness analysis. No proprietary camera stream, human-subject dataset, or completed field deployment was used. Accordingly, the numerical results reported in Section 5 are synthetic and are intended to demonstrate an evaluation architecture rather than claim measured performance.

The methodological framework was constructed from primary research on low-light enhancement, low-light detection, dark-scene benchmarks, task-oriented image and feature compression, split computing, and bandwidth-efficient collaborative inference.

4.2 Integrated Low-Light and Bandwidth-Aware Pipeline

The proposed pipeline contains six stages.

The first stage estimates environmental illumination and sensor degradation. The second performs lightweight, noise-aware normalization rather than unconditional brightening. The third extracts features on the edge device. The fourth determines the available communication rate and selects a transmission policy. The fifth performs local, split, or remote inference depending on computational and bandwidth conditions. The sixth evaluates confidence and activates fallback behavior when the system moves outside its validated operating region.

Table 1: Proposed Robust Computer-Vision Framework

Pipeline Stage	Primary Challenge	Proposed Strategy	Reliability Function
Illumination Assessment	Darkness varies substantially between frames and locations	Estimate illumination, exposure, noise, and signal quality	Prevents unnecessary or excessive enhancement
Low-Light Preprocessing	Brightening may amplify noise or create artifacts	Lightweight task-aware enhancement and denoising	Improves recoverable object structure
Edge Feature Extraction	Edge devices have limited computation	Compact backbone or partial detector execution	Reduces local computational burden
Adaptive Communication	Available bitrate changes over time	Feature compression, channel selection, quantization, frame prioritization	Preserves task-relevant information under bandwidth limits
Split/Remote Inference	Cloud access may be intermittent or delayed	Dynamic split point and local fallback	Maintains service continuity
Reliability Monitoring	Severe degradation can invalidate predictions	Confidence calibration, shift detection, abstention, human/system escalation	Reduces silent high-confidence failure

Source: Author-developed synthesis informed by low-light enhancement, task-aware low-light detection, feature compression, and collaborative inference research.

4.3 Simulation and Evaluation Metrics

Four pipeline strategies were constructed for comparative demonstration.

- The Baseline Raw/Naïve Pipeline applies no specialized low-light processing or task-aware communication strategy.
- The Low-Light Enhancement + Standard Compression Pipeline applies enhancement before transmitting conventionally compressed images.
- The Task-Aware Feature Compression Pipeline performs partial network inference on the edge and transmits compressed task-specific features.
- The Joint Low-Light + Bandwidth-Aware Pipeline adapts enhancement, feature preservation, compression rate, and inference location jointly according to environmental conditions.

5. Analysis

5.1 Simulated Robustness Comparison

The simulation evaluates repeated combined stress conditions rather than a single average illumination level or bitrate.

Table 2: Simulated Performance of Robust Vision Pipeline Strategies

Pipeline Strategy	Mean RVS	Median RVS	Minimum	Maximum	Principal Interpretation
Baseline Raw/Naïve Pipeline	50.83	51.50	41	58	Highly vulnerable when illumination and communication deteriorate simultaneously
Low-Light Enhancement + Standard Compression	61.92	62.50	52	70	Enhancement improves recoverability but conventional compression remains task-agnostic
Task-Aware Feature Compression	69.83	70.00	61	78	Task-relevant feature transmission provides stronger rate efficiency
Joint Low-Light + Bandwidth-Aware Pipeline	79.17	79.50	70	88	Coordinated adaptation provides strongest simulated robustness

Note: All values are simulated methodological examples and must not be interpreted as measurements of commercial or research systems.

American Journal of Advanced Multidisciplinary Innovation and Research (AJAMIR)
E-ISSN: XXXX-XXXX ● Website: www.ajamir.com ● Email: editor@ajamir.com

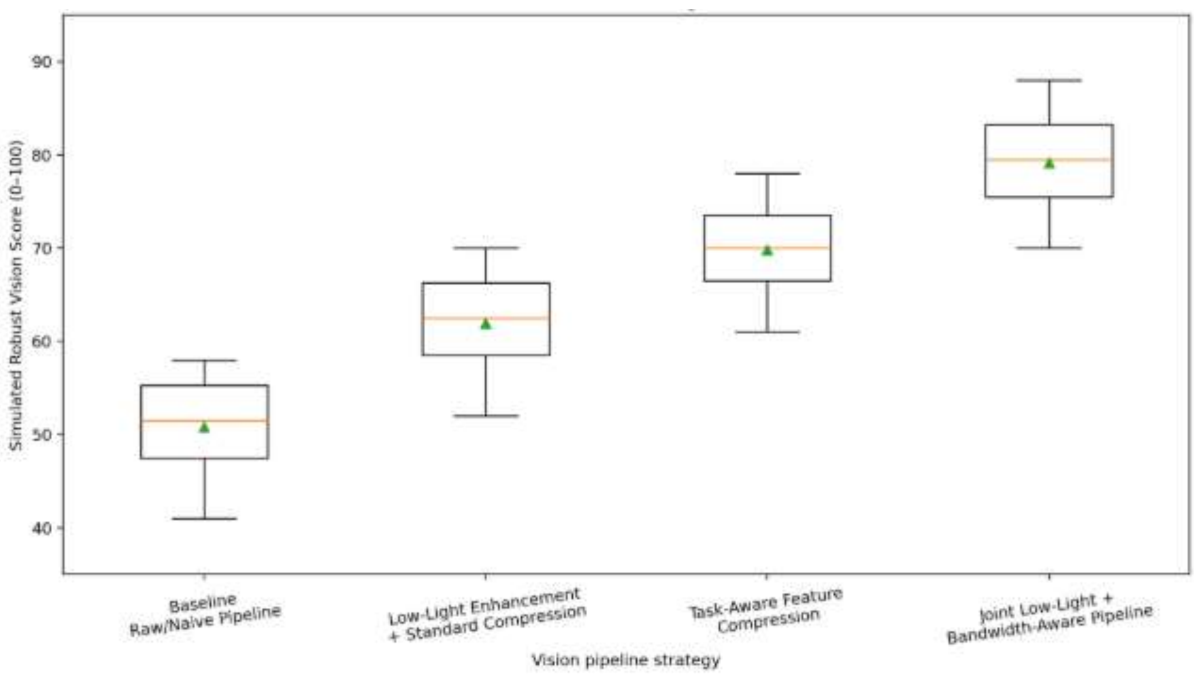
The baseline produces the weakest mean and worst-condition results. This outcome represents the expected vulnerability of a system that attempts to transmit degraded imagery without optimizing either image formation or communication for machine perception.

Enhancement followed by conventional compression improves the simulated mean by approximately 11 points. This improvement reflects greater visibility of task information but remains constrained because the communication stage optimizes reconstructed image quality rather than the requirements of the downstream detector.

Task-aware feature compression raises simulated robustness further. This result is methodologically consistent with rate-constrained object-detection research showing advantages for jointly learned feature compression relative to conventional image transmission at constrained rates.The integrated strategy produces the strongest simulated minimum, median, and mean. Its advantage arises because it does not assume a fixed enhancement strength or compression policy. Instead, the system modifies preprocessing and transmission according to the actual degradation state.

5.2 Box-Plot Analysis of Robustness

Graph 1: Distribution of Simulated Robust Vision Scores Under Combined Low-Light and Low-Bandwidth Conditions



Source: Author-generated simulation.

The box plot provides additional information beyond the mean values because robustness is partly determined by variability. The baseline pipeline exhibits low central performance and remains concentrated within a weak operating region. Low-light enhancement shifts the score distribution upward, but performance continues to deteriorate materially under the more demanding simulated conditions.

Task-aware feature compression raises the entire distribution and improves the lower tail. The joint strategy produces the strongest simulated median while also keeping its minimum above the median performance of the baseline. This lower-tail behavior is particularly important because critical computer-vision deployments should be evaluated according to difficult operating conditions rather than only average-case accuracy.

The simulated pattern illustrates an important design principle: a robust system should improve not merely peak performance but the consistency of performance across darkness, bitrate reduction, and combined degradation.

5.3 Operational Robustness Strategy

The analysis supports adaptive rather than static operation. Under adequate lighting and stable bandwidth, the platform should avoid unnecessary enhancement and may transmit higher-quality visual or feature information. Under reduced illumination but adequate communication, stronger low-light processing or richer intermediate features may be justified. When bandwidth falls, the system should preserve task-critical regions or features and suppress redundant information.

Under simultaneous severe darkness and low bandwidth, the system should not simply increase compression. It should consider whether the information content remains adequate for reliable inference. If confidence drops below a validated threshold, the correct action may be to reduce inference frequency, request additional exposure, switch sensing modality where available, delay the decision, transmit selected high-value regions, or abstain from an automatic conclusion.

6. Discussion

6.1 Visual Enhancement Should Serve Machine Perception Rather Than Appearance Alone

A central implication of the analysis is that the visually most attractive enhancement is not necessarily the most useful input for an object detector. Low-light enhancement traditionally emphasizes brightness, color balance, contrast, perceptual quality, or reconstruction fidelity. Machine vision instead depends on whether task-relevant spatial and semantic information survives processing. Excessive brightening can amplify noise, alter texture statistics, create halos, distort color relationships, or generate features that appear visually plausible but are inconsistent with the detector's training distribution. Research connecting low-light processing directly with object detection and hierarchical feature enhancement reinforces the value of downstream-task evaluation rather than perceptual evaluation alone.

The appropriate enhancement strategy should therefore be conditional on the application. A system developed for human viewing may legitimately prioritize perceptual quality, while an automated detector should optimize the recoverability of features that determine detection performance. In some deployments, the most reliable approach may involve modest illumination normalization followed by feature-level enhancement rather than producing a fully reconstructed bright image. This principle also reduces unnecessary edge computation because the system need not solve a more difficult photographic restoration problem when the operational requirement is simply reliable object classification or localization.

Task-oriented enhancement should also be trained and validated over multiple illumination states. DarkVision is particularly valuable in this respect because its multi-illuminance structure permits analysis across graded photon-limited conditions. ExDark similarly demonstrates that “low light” is not one homogeneous domain but encompasses multiple illumination patterns. A robust model should therefore report performance according to illumination severity rather than collapsing all dark-scene images into a single test category.

6.2 Bandwidth Reduction Should Preserve Semantics Before Pixels

A second implication concerns what should be transmitted when bandwidth becomes scarce. Conventional image compression was designed principally to reconstruct signals efficiently for human consumption. In machine-to-machine vision, the receiving system may not need a visually faithful image if compact intermediate features preserve the information required for detection. Rate-constrained feature-compression research supports this distinction by directly optimizing compressed neural representations for downstream object detection. Lightweight activation coding provides further evidence that intermediate features can be compressed aggressively while keeping tested task degradation relatively small.

This does not imply that feature transmission is universally preferable. Raw or reconstructed imagery may still be required for human review, forensic interpretation, multi-task reuse, or legal recordkeeping. Feature-space approaches can also make systems more dependent on particular network architectures and may complicate interoperability. A robust architecture should therefore preserve multiple operating modes. When bandwidth is high, the system may transmit richer visual data. Under constrained conditions, task-aware feature communication can provide a fallback. In safety-sensitive cases, selected original image regions may be transmitted alongside features for verification.

Adaptive communication is preferable to one fixed compression setting because network conditions change continuously. Newer collaborative-inference systems increasingly emphasize rate adaptation, selective activation transmission, and communication-efficient feature exchange. The relevant engineering objective is therefore not the smallest possible bitrate, but the lowest rate that preserves a required level of task reliability.

6.3 Robustness Requires Joint Degradation Monitoring and Failure-Aware Control

The third implication is that low-light and low-bandwidth systems require monitoring at the level of the complete inference pathway. An image-quality metric alone cannot detect communication-induced feature loss, while a network-throughput monitor cannot determine whether poor illumination has already removed essential visual information. The platform should therefore jointly monitor exposure, estimated noise, image or feature quality, bitrate, packet loss where relevant, inference confidence, latency, and disagreement among available models or sensors.

This monitoring should determine system behavior. For example, a modest decrease in illumination may justify lightweight enhancement without changing the transmission policy. A sharp bandwidth reduction may trigger stronger feature compression. If both conditions deteriorate simultaneously, the model's confidence may no longer remain calibrated because the input lies outside the conditions used during validation. Continuing to output high-confidence predictions under such circumstances would create a reliability risk.

A failure-aware system should therefore support abstention. Robustness is not equivalent to always producing a prediction. In high-consequence settings, the ability to recognize that visual evidence is inadequate can be more valuable than preserving nominal real-time operation. Low-light benchmarks demonstrate that severe photon scarcity can meaningfully reduce downstream object information, while low-rate collaborative-computing research shows that communication constraints create explicit rate–accuracy trade-offs. A responsible deployment strategy should make these trade-offs visible and define when automatic inference is no longer justified.

7. Conclusion

Robust computer vision under low-light and low-bandwidth conditions requires a wider engineering perspective than either conventional image enhancement or conventional network compression provides independently. Poor illumination reduces the amount and quality of information captured by the imaging sensor, while restricted communication limits how much of the remaining information can be transferred for downstream analysis. When these constraints occur simultaneously, independently optimized stages can work against one another: enhancement may amplify noise, compression may remove newly recovered detail, and remote inference may receive representations that are both visually degraded and communication-limited.

The methodological framework developed in this paper addresses the problem through coordinated adaptation across sensing, enhancement, feature extraction, communication, inference, and uncertainty management. The simulation indicates a progressive improvement from the baseline pipeline to low-light enhancement, task-aware feature compression, and finally the integrated low-light/bandwidth-aware strategy. The mean simulated robustness score rises from 50.83 to 79.17, while the minimum score increases from 41 to 70. These values are illustrative rather than empirical, but the box-plot distribution demonstrates the intended reliability objective: the strongest pipeline should improve both average performance and the lower tail of performance across adverse operating conditions.

The study further argues that machine-perception systems should preserve semantic utility before visual fidelity when communication resources become scarce. Low-light enhancement should be judged by its effect on the downstream task, and compression should allocate scarce bits toward task-relevant information. Feature compression, dynamic split computing, selective activation transmission, and task-aware video coding provide increasingly mature mechanisms for this objective. At the same time, full-image availability may remain necessary for human inspection or multi-purpose reuse, meaning that practical systems should support adaptive rather than exclusively feature-based communication.

The central conclusion is that a reliable vision platform must know when environmental conditions have exceeded its validated operating envelope. Severe darkness and aggressive compression can remove information that no downstream model can reconstruct reliably. Future research should therefore combine low-light benchmarks with explicit bitrate stress tests, jointly optimize enhancement and feature coding, evaluate calibration under compound degradation, investigate multimodal fallback using thermal or near-infrared sensing where appropriate, and establish standardized reporting of worst-condition performance. The most dependable low-resource computer-vision system will not be the one that always produces an answer, but the one that preserves useful visual evidence efficiently, adapts intelligently to changing constraints, and identifies when the available evidence is no longer sufficient for a trustworthy decision.

References

1. Jobson, D. J., Rahman, Z., & Woodell, G. A. (1997). A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image Processing*, *6*(7), 965–976. https://doi.org/10.1109/83.597272
2. Land, E. H., & McCann, J. J. (1971). Lightness and retinex theory. *Journal of the Optical Society of America, 61*(1), 1–11. https://doi.org/10.1364/JOSA.61.000001
3. Guo, X., Li, Y., & Ling, H. (2017). LIME: Low-light image enhancement via illumination map estimation. *IEEE Transactions on Image Processing, 26*(2), 982–993. https://doi.org/10.1109/TIP.2016.2639450
4. Lore, K. G., Akintayo, A., & Sarkar, S. (2017). LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognition, 61*, 650–662. https://doi.org/10.1016/j.patcog.2016.06.008
5. Chen, C., Chen, Q., Xu, J., & Koltun, V. (2018). Learning to see in the dark. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3291–3300.
6. Wei, C., Wang, W., Yang, W., & Liu, J. (2018). Deep retinex decomposition for low-light enhancement. *British Machine Vision Conference*, 1–12.
7. Wang, W., Wei, C., Yang, W., & Liu, J. (2019). GladNet: Low-light enhancement via global and local adaptive network. *IEEE Transactions on Image Processing, 29*, 4636–4648.
8. Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., Yang, J., Yuan, P., & Wang, Z. (2021). EnlightenGAN: Deep light enhancement without paired supervision. *IEEE Transactions on Image Processing, 30*, 2340–2349. https://doi.org/10.1109/TIP.2021.3051462
9. Zhang, Y., Zhang, J., & Guo, X. (2019). Kindling the darkness: A practical low-light image enhancer. *Proceedings of the 27th ACM International Conference on Multimedia*, 1632–1640.
10. Ren, W., Liu, S., Zhang, H., Pan, J., Cao, X., & Yang, M.-H. (2019). Low-light image enhancement via a deep hybrid network. *IEEE Transactions on Image Processing, 28*(9), 4364–4375.
11. Zamir, S. W., Arora, A., Khan, S., Hayat, M., Khan, F. S., & Yang, M.-H. (2020). Learning enriched features for real image restoration and enhancement. *European Conference on Computer Vision Workshops*, 492–511.
12. Zamir, S. W., Arora, A., Khan, S., Hayat, M., Khan, F. S., & Yang, M.-H. (2021). Restormer: Efficient transformer for high-resolution image restoration. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 5728–5739.
13. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 770–778.
14. Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.

15. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4510–4520.
16. Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. *Proceedings of the 36th International Conference on Machine Learning*, 6105–6114.
17. Sze, V., Chen, Y.-H., Yang, T.-J., & Emer, J. S. (2017). Efficient processing of deep neural networks: A tutorial and survey. *Proceedings of the IEEE*, *105*(12), 2295–2329. https://doi.org/10.1109/JPROC.2017.2761740
18. Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge computing: Vision and challenges. *IEEE Internet of Things Journal*, *3*(5), 637–646. https://doi.org/10.1109/JIOT.2016.2579198
19. Satyanarayanan, M. (2017). The emergence of edge computing. *Computer*, *50*(1), 30–39. https://doi.org/10.1109/MC.2017.9
20. Ballé, J., Minnen, D., Singh, S., Hwang, S. J., & Johnston, N. (2018). Variational image compression with a scale hyperprior. *International Conference on Learning Representations (ICLR)*.