Journal of Computer Science and Artificial Intelligence
ISSN: 3078-8242 (Print), ISSN: 3078-2309 (Online) | Vol. 8, No. 1, 2026

An Adaptive Portrait Image Enhancement Algorithm Based on Facial and Scene Statistics in Complex Lighting

Guozhi Hu [1, 2, *], Robiatul A'dawiah Jamaluddin [1]

[1] Kuala Lumpur University of Science and Technology (KLUST), Kuala Lumpur, Malaysia
[2] Beijing Kingsha Technology Co., Ltd., Beijing, China
[*] **Corresponding author**: Guozhi Hu (Email: 242924843@s.klust.edu.my)

Abstract: Portrait image quality in user-provided photographs is frequently compromised by highly variable capture conditions, including underexposure, facial shadows cast by side lighting, and color casts introduced by indoor illuminants. This paper presents an interpretable, fully ablatable workflow for portrait-perception-based color normalization. The algorithm builds a multidimensional statistical profile of the face and the surrounding scene in the LAB and HSV color spaces, and uses this profile to drive adaptive color-cast correction with a dead-zone-and-damping mechanism, localized skin-tone equalization through large-scale Gaussian diffusion, nonlinear luminance mapping with highlight suppression, and detail-aware texture restoration. By separating luminance from chrominance, the method avoids the color crosstalk typical of RGB-space adjustments. On a dataset of 120 portraits covering 12 typical complex-lighting scenarios, our method raises the central L-mean from 144.64 to 165.55 while constraining the overall highlight ratio to 0.72%. In the overexposed group, the method suppresses the highlight ratio to 1.12%, avoiding the failure state (8.57%) produced by conventional brightness enhancement, and in the low-light group it lifts the central L-mean from 128.37 to 170.05. Ablation experiments confirm that the color-normalization module is the dominant contributor to these gains. The workflow offers a robust, reproducible, and interpretable alternative to black-box enhancement filters for portrait normalization under complex lighting.

Keywords: Color normalization, complex lighting, highlight suppression, portrait enhancement, statistical profiling.

1. Introduction

The processing of user-uploaded portraits for official documentation is a typical "small image, large constraint" problem. Unlike studio photography, consumer-grade uploads are captured under uncontrolled illumination: side lighting produces heavy facial shadows, indoor luminaires introduce warm color shifts, and backlit scenes combine overexposed backgrounds with underexposed subjects. Generic global operators such as gamma correction or histogram equalization can improve overall contrast, but they lack any prior awareness of the portrait subject. As a result, they frequently distort skin tones, amplify highlight clipping in bright regions while attempting to recover shadow detail, and produce unstable results across heterogeneous inputs.

This paper presents a systematic and reproducible workflow for portrait-perception-based color normalization. In contrast to black-box learning models, our approach follows a constrained-optimization philosophy in which every adjustment is driven by explicit statistics and bounded by explicit protection mechanisms, balancing overall brightness, skin-tone naturalness, and texture retention.

Specifically, we introduce a multidimensional statistical profiling framework in the LAB/HSV spaces that generates robust exposure anchors from L-channel percentiles, detects side lighting through a left–right half-face balance metric, and characterizes the scene context (backlight, white background) to trigger scenario-specific constraints. Building on this profile, we design a white-balance and color-cast correction strategy with a mathematically explicit dead-zone-and-damping rule and a white-background protection score, which together prevent over-correction, desaturated "grey faces," and background contamination. Finally, we develop a localized skin-tone equalization mechanism that diffuses correction seeds through a large-scale Gaussian kernel into a smooth, spatially varying correction surface, lifting facial shadows and side-lit regions without disturbing the background.

2. Related Work

2.1. Traditional Image Enhancement

Classical enhancement methods follow either histogram-based global mapping or Retinex theory [1]. Histogram equalization and its contrast-limited adaptive variant CLAHE [3, 4] redistribute intensity values to improve visibility, while multiscale Retinex with color restoration (MSRCR) [2] estimates and removes the illumination component. LIME [5] refines an illumination map with structure priors for low-light enhancement. These methods are effective at raising global visibility but are agnostic to the semantics of the image: applied to portraits, they cannot distinguish facial skin from background, and therefore cannot enforce skin-tone naturalness or protect bright backgrounds from clipping.

2.2. Learning-Based Enhancement

Deep models such as RetinexNet [6], Zero-DCE [7], EnlightenGAN [8], and URetinex-Net [9] achieve strong perceptual quality on low-light benchmarks. However, they behave as black boxes: their failure modes are difficult to diagnose, their outputs are not guaranteed to be stable across the narrow but strict requirements of document photography, and they typically require GPU inference and paired or curated training data. For portrait normalization pipelines that must be auditable and deterministic, an interpretable statistical approach remains attractive.

2.3. Color Constancy and Portrait-Specific Constraints

Color-cast estimation is closely related to computational color constancy, from the classical gray-world assumption [10] to edge-based methods [12] and the systematic treatment in [11]. Portrait-oriented enhancement imposes stricter physical constraints than generic color correction: skin chromaticity must remain within a natural range, facial illumination should be spatially balanced, and any correction must preserve identity-related texture. Our method uses face priors (location from a standard detector [13] and skin-region statistics) to anchor all adjustments, and borrows the spirit of edge-preserving regularization [14] and adaptive unsharp masking [16] for texture restoration. The same design integrates cleanly with modern portrait matting front-ends such as BiRefNet [15].

3. Methodology

Figure 1 shows the overall pipeline. Given an input BGR image, the system first extracts facial key regions, converts the image to the LAB and HSV spaces, and builds a multidimensional statistical profile. The profile then drives four adjustment stages executed in sequence: color-cast correction, localized fill light, nonlinear luminance mapping with highlight suppression, and texture restoration. All chrominance operations are confined to the LAB a/b channels and all luminance operations to the L channel, which avoids the color crosstalk typical of RGB-space adjustments.

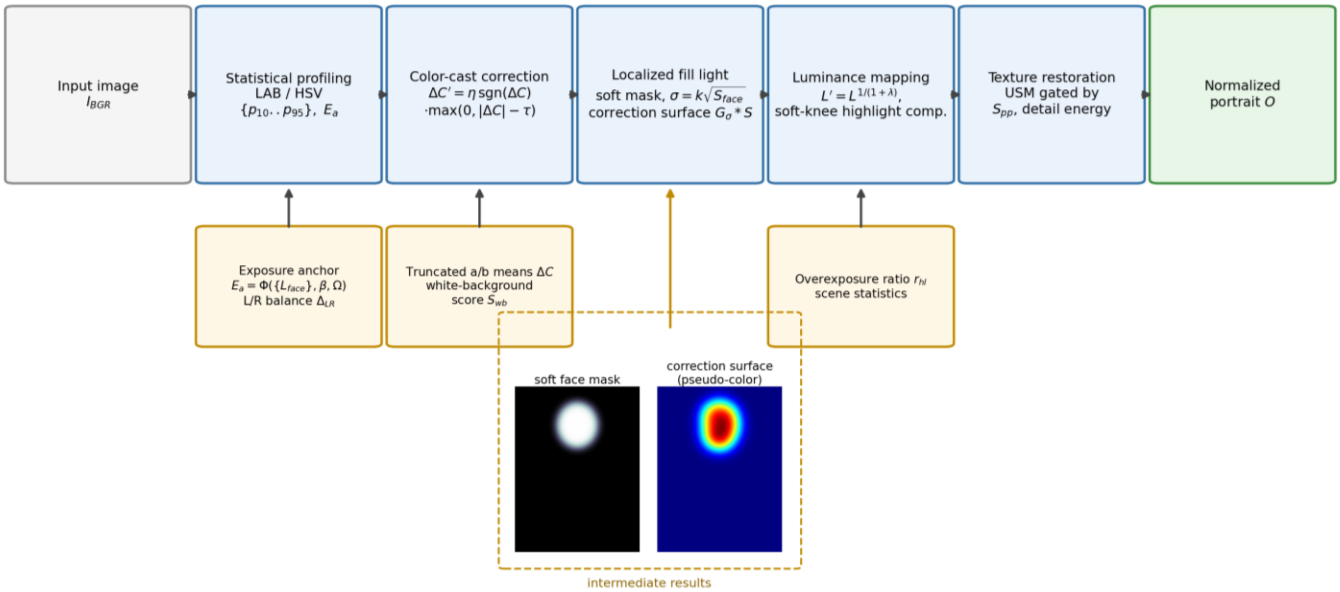

Figure 1. Data flow of the proposed portrait-perception-based color normalization. Each module is annotated with its governing quantities (Ea, ΔC′, η, τ, σ, Spp); the insets show two intermediate results of the localized fill-light stage — the soft face mask and the diffusion-based correction surface (pseudo-color).

3.1. Facial and Scene Statistical Profiling

The profile is built on five percentiles of the facial L-channel distribution, p_{10}, p_{25}, p_{50}, p_{85}, and p_{95}. Rather than using a simple mean, the exposure anchor E_a is defined as a robust estimate over a dynamically selected statistical window,

$$\mathrm{Ea} = \Phi(\{\mathrm{Lface}\},\ \beta,\ \Omega) \tag{1}$$

Where β is the shadow bias measured from the low percentiles and Ω encodes the dark-recovery demand. When shadows dominate, the window shifts toward the low–mid percentiles; when the face is bright, high percentiles are discounted so that local highlights cannot skew the anchor. Side lighting is detected by the left–right half-face brightness difference ΔLR computed across the sagittal plane of the detected face box; a large |ΔLR| triggers asymmetric compensation in Section 3.3. Scene-level statistics — the non-facial mean luminance, dark- and light-pixel ratios — identify environmental contexts such as backlit subjects (background much brighter than the face) or white backgrounds. A white-background protection score combines background brightness and background chroma; a high score limits both global enhancement and color correction so that a light background is not crushed into grey.

3.2. Color-cast Correction with Dead Zone and Damping

The global color cast ΔC is estimated from the truncated means (10–90%) of the LAB a and b channels, which resists contamination by saturated pixels. The applied correction ΔC′ follows a dead-zone-and-damping rule:

$$\Delta C' = \eta \cdot \mathrm{sgn}(\Delta C) \cdot \max(0, |\Delta C| - \tau) \tag{2}$$

Where τ is a dead-zone threshold that preserves faint, natural ambience (e.g., mild skin warmth), and η ∈ (0, 1] is a damping coefficient that caps the correction strength. The correction is further scaled down in proportion to the white-background protection score. Together these mechanisms prevent the desaturated “grey-face” artifact common to aggressive white-balance methods.

3.3. Skin-tone Equalization Via Large-Scale Gaussian Diffusion

Rather than applying a global fill light, the algorithm identifies correction seeds — underexposed pixels inside the facial region, augmented on the darker half-face when |ΔLR| indicates side lighting — and diffuses them with a large-scale Gaussian kernel to form a smooth, spatially varying correction surface. The surface is applied through a soft face mask, a binary facial mask blurred with a radius that adapts to the face scale: $\sigma = k \cdot \sqrt{S_{face}}$, where S_{face} is the face area in pixels. This design guarantees a natural transition between face and background, raises visibility in facial shadows and side-lit regions, and leaves the global background undisturbed. A highlight-protection weight fades the gain to zero in already-bright areas.

3.4. Nonlinear Luminance Mapping and Highlight Suppression

Shadow recovery applies a profile-driven nonlinear (gamma-type) lift whose strength is proportional to the gap between the exposure anchor and the target luminance, and attenuated by the white-background protection score. Highlight suppression is governed by the measured overexposure ratio: a soft-knee compression is applied above a luminance threshold so that gains in the dark and mid tones cannot push bright regions into clipping. In over-saturated bright areas the LAB b-mean is additionally lowered to suppress warm-hue drift.

3.5. Texture Restoration with Portrait Protection

Finally, unsharp masking driven by local detail energy restores edge clarity lost during smoothing. To avoid destroying natural skin grain or amplifying noise on smooth gradients, a portrait protection score S_{pp} suppresses sharpening on low-frequency skin regions (cheeks, forehead) while permitting it on high-frequency structures such as eyes and hair [16].

4. Experiments and Results

4.1. Dataset and Experimental Setup

The method was validated on a purpose-built dataset of 120 consumer-grade portrait photographs organized into 12 categories of 10 images each, covering six scenario groups: illumination/exposure (low light, overexposed), color deviation (yellow indoor light, reddish skin tone), background complexity (complex background, light background), clothing color (dark clothing, light clothing), body geometry (wide shoulders, narrow shoulders), and hair boundary (long hair, wispy hair). All images were processed by the same standardization front-end (face detection, BiRefNet foreground segmentation [15], and geometric composition), so that the compared conditions differ only in the color-processing stage:

B0 — no color processing (unprocessed baseline);

B1 — simple global brightness/contrast enhancement;

B2 — the proposed full color normalization with fixed geometric composition;

Proposed — the proposed full color normalization within the complete adaptive pipeline.

Four statistics are computed over the central face-covering region of each output: the L-channel mean (central L-mean), the dark-pixel ratio (share of pixels with $L < 60$), the highlight ratio (share of pixels with $L > 245$), and the mean HSV saturation. The central L-mean measures achieved brightness, the dark and highlight ratios measure failure at the two ends of the tonal range, and saturation reflects color-cast removal.

4.2. Quantitative Analysis

Table 1 reports means over all 120 samples. Our method achieves the highest central L-mean among all conditions while keeping the highlight ratio essentially at the unprocessed level (0.72% vs. 0.42% for B0), whereas simple enhancement (B1) increases the highlight ratio sevenfold to 3.00%. The near-identical figures of B2 and the full pipeline confirm that these color statistics are produced by the color-normalization module itself and are robust to the surrounding geometric processing.

Table 1. Overall performance comparison (mean over 120 samples).

Method	L-mean	Dark (%)	HL (%)	Sat.
B0	144.64	6.58	0.42	100.94
B1	160.10	5.54	3.00	95.42
B2	165.89	3.78	0.70	84.91
Ours	165.55	3.95	0.72	85.03

4.3. Category-specific Performance

The value of the constrained design is most evident in the extreme-lighting groups (Table 2, Figure 2). In the overexposed group, simple enhancement pushes the central L-mean to 178.46 at the cost of an 8.57% highlight ratio — a failure state in which facial detail is irreversibly lost to clipping — while ours suppresses it to 1.12%. In the low-light group, the algorithm lifts the central L-mean from 128.37 to 170.05, well above its global average, demonstrating targeted compensation for severely underexposed samples. In the two color-deviation groups, saturation is reduced from 103.39 to 87.86 (yellow indoor light) and from 108.49 to 87.34 (reddish skin tone), indicating effective removal of the respective casts.

Table 2. Performance in the extreme-lighting categories.

Category	Metric	B0	B1	Proposed
Low light	L-mean	128.37	143.22	170.05
Low light	Sat.	119.18	113.33	89.07
Overexposed	HL (%)	0.78	8.57	1.12
Overexposed	L-mean	162.98	178.46	169.55

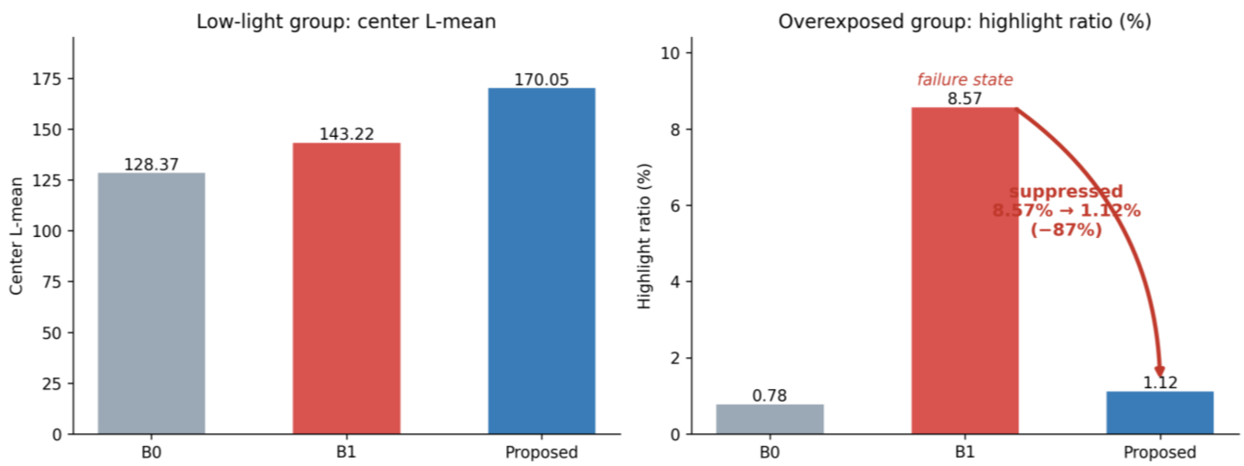

Figure 2. Category-level comparison: low-light brightness recovery (left) and overexposure control (right). Simple enhancement (B1) enters a highlight-clipping failure state at 8.57%, which the proposed method suppresses by 87% to 1.12%.

4.4. Ablation study

Two ablations isolate the contribution of the color-normalization module within the full pipeline. A1 removes the module entirely; A2 replaces it with simple brightness/contrast enhancement. As Table 3 shows, A1 collapses to the unprocessed statistics (central L-mean 144.78) and A2 reproduces the failure pattern of B1 (highlight ratio 3.07%), confirming that neither omission nor a naive substitute can achieve the balanced brightness–highlight–saturation profile of the full module.

Table 3. Ablation of the color-normalization module.

Variant	L-mean	Dark (%)	HL (%)	Sat.
A1 (no color)	144.78	6.67	0.43	100.69
A2 (simple enh.)	160.20	5.63	3.07	95.17
Ours	165.55	3.95	0.72	85.03

4.5. Qualitative Discussion

Figure 3 presents representative results in four scenarios; the depicted subjects are consenting volunteers photographed under the corresponding lighting conditions and processed with settings identical to the main experiments. Our method avoids the low-contrast “grey fog” typical of global brightness shifts: by reducing mean saturation from 100.94 to 85.03 it renders skin in a neutral, professional tone, while the combination of side-light detection and Gaussian-diffusion equalization preserves local tonal hierarchy and edge clarity. In the low-light example the face is lifted dramatically without background artifacts; in the overexposed example facial highlights remain controlled; in the two color-cast examples the yellow and red shifts are removed while skin remains natural rather than grey.

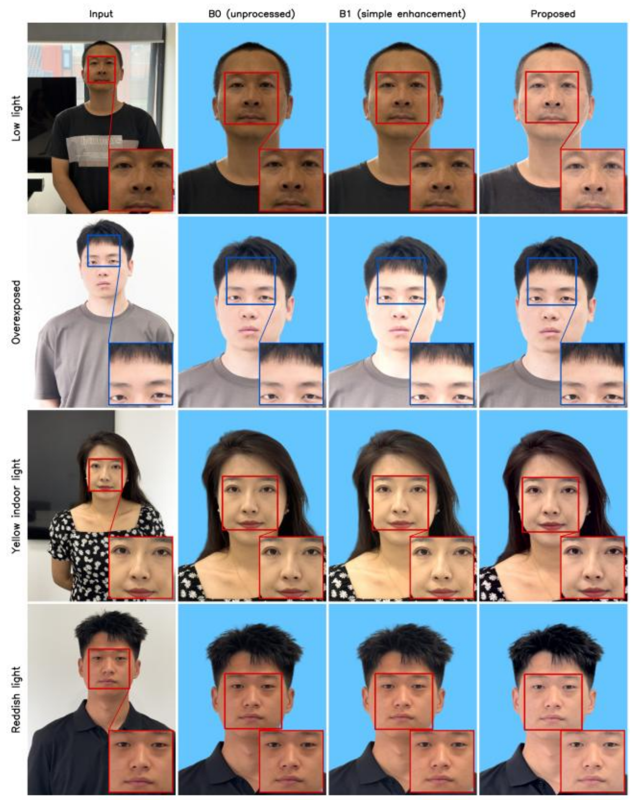

Figure 3. Visual comparison in four representative scenarios (rows) for the input and three processing conditions (columns). Red boxes and insets magnify the facial region; blue boxes in the overexposed row magnify the highlight-prone forehead region, where B1 clips to white while our method preserves detail.

5. Discussion

The gains reported in Section 4 come from adjustments that are individually traceable: every correction is driven by a named statistic and bounded by an explicit protection rule, so a specific lighting defect — a warm cast, a side-light shadow, a clipped highlight — can be diagnosed and tuned in isolation. This is difficult to achieve with black-box models and matters in auditable document-photography pipelines; it is also what makes the workflow fully ablatable, as Section 4.4 demonstrates. Although the algorithm behaves stably across all twelve categories, its performance naturally relies on the accuracy of the initial face localization — under severe occlusion or extreme pose the facial statistical profile may become unreliable, which suggests coupling the profiler with a more robust landmark detector. The evaluation is currently based on objective statistics over a 120-image dataset; a formal subjective study with blind ratings, broader demographic coverage of skin tones, and a comparison against learning-based enhancers under an identical protocol would further strengthen the aesthetic claims.

6. Conclusion

This paper presented an adaptive, perception-based approach to portrait enhancement under complex lighting. By integrating multidimensional statistical profiling with a dead-zone-and-damping correction rule, Gaussian-diffusion skin-tone equalization, and constraint-driven luminance mapping, the method normalizes portraits across disparate lighting conditions while enforcing strict guarantees on highlights and skin naturalness. On 120 samples across 12 scenarios it raised the central L-mean from 144.64 to 165.55 with an overall highlight ratio of only 0.72%, and in overexposed scenes reduced the highlight failure of naive enhancement from 8.57% to 1.12%. Future work will integrate human parsing models for joint photometric–geometric optimization and incorporate subjective blind testing to refine aesthetic weightings across diverse skin tones.

Ethics Statement

All example photographs shown in the figures depict volunteers who were photographed for this study and provided written informed consent for the publication of identifiable images. The 120-image evaluation dataset originates from a commercial portrait-standardization service under user agreements permitting research use; these images are used for aggregate statistical evaluation only and are not displayed in this paper.

References

[1] Land, E. H. (1977). The retinex theory of color vision. Scientific American, 237(6), 108–128. https://doi.org/10.1038/scientificamerican1277-108

[2] 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

[3] Pizer, S. M., et al. (1987). Adaptive histogram equalization and its variations. Computer Vision, Graphics, and Image Processing, 39(3), 355–368. https://doi.org/10.1016/0734-189X(87)90156-4

[4] Zuiderveld, K. (1994). Contrast limited adaptive histogram equalization. In P. S. Heckbert (Ed.), Graphics gems IV (pp. 474–485). Academic Press.

[5] 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.2016.2631944

[6] Wei, C., Wang, W., Yang, W., & Liu, J. (2018). Deep retinex decomposition for low-light enhancement. In Proceedings of the British Machine Vision Conference (BMVC). https://doi.org/10.5244/C.32.127

[7] Guo, C., et al. (2020). Zero-reference deep curve estimation for low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 1780–1789). https://doi.org/10.1109/CVPR42600.2020.01791

[8] Jiang, Y., et al. (2021). EnlightenGAN: Deep light enhancement without paired supervision. IEEE Transactions on Image Processing, 30, 2340–2349. https://doi.org/10.1109/TIP.2021.3052733

[9] Wu, W., et al. (2022). URetinex-Net: Retinex-based deep unfolding network for low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 5901–5910). https://doi.org/10.1109/CVPR52688.2022.00584

[10] Buchsbaum, G. (1980). A spatial processor model for object colour perception. Journal of the Franklin Institute, 310(1), 1–26. https://doi.org/10.1016/0016-0032(80)90033-8

[11] Ebner, M. (2007). Color constancy. John Wiley & Sons.

[12] van de Weijer, J., Gevers, T., & Gijsenij, A. (2007). Edge-based color constancy. IEEE Transactions on Image Processing, 16(9), 2207–2214. https://doi.org/10.1109/TIP.2007.903229

[13] Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (Vol. 1, pp. 511–518). https://doi.org/10.1109/CVPR.2001.990517

[14] Rudin, L. I., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. Physica D: Nonlinear Phenomena, 60(1–4), 259–268. https://doi.org/10.1016/0167-2789(92)90242-F

[15] Zheng, P., et al. (2024). Bilateral reference for high-resolution dichotomous image segmentation. CAAI Artificial Intelligence Research, 3, 9150038. https://doi.org/10.1016/j.caair.2024.9150038

[16] Polesel, A., Ramponi, G., & Mathews, V. J. (2000). Image enhancement via adaptive unsharp masking. IEEE Transactions on Image Processing, 9(3), 505–510. https://doi.org/10.1109/83.827816