DOI: 10.5281/zenodo.24158838
Proceedings of CCSN2025: (14th International Conference on Computing, Communication and Sensor Network)

AI-DWTE: Learning-Guided Wavelet-Domain Enhancement of Low-Contrast Remote Sensing Images with Gain-Consistent Detail Preservation

Rajni Sharma, Ravinder M

Dept. of CSE
Indira Gandhi Delhi Technical University for Women, Kashmere Gate, New Delhi, India.
Corresponding Author: rajnisharma.hp@gmail.com

ABSTRACT

Low contrast and poor illumination limit how well land cover, roads and settlements can be interpreted in optical remote sensing images. Wavelet–SVD methods tuned by nature-inspired algorithms (NIA) such as PSO and GWO are fast and interpretable, but they search per image for a proxy fitness (mean-intensity matching). We show that their SVD correction reduces to a single scalar gain on the approximation band, and that the detail bands are left untouched, so weak detail is attenuated relative to the brightened background. We propose AI-DWTE, which replaces the per-image search with a learned predictor. A deep ensemble of multilayer perceptrons maps 100 wavelet-domain statistics to the eight parameters of a Haar-DWT operator in a single forward pass. The operator combines black-point removal, channel-wise power-law mapping, a learned singular-value gain, local contrast, and a new gain-consistent detail transfer that scales each detail coefficient by the local slope of the tone curve after wavelet shrinkage. The network is trained on parameters found by an offline oracle that has access to ground truth. On 420 held-out UC Merced images under three degradations (1,260 cases), AI-DWTE is compared against ten techniques, from HE and CLAHE through the PSO/GWO base method to the pretrained deep models Zero-DCE, Zero-DCE++ and SCI. It improves average PSNR from 17.95 dB (PSO base method) to 21.71 dB, SSIM from 0.787 to 0.851, and CIEDE2000 colour error from 13.85 to 9.75. It runs in 27 ms per image on one CPU thread, 4.3× faster than the PSO search. Under an unseen non-uniform illumination degradation, its gain over the base method is small (+0.15 dB) and not statistically significant, a limitation we report and analyse.

Keywords: Remote sensing image enhancement, discrete wavelet transform, singular value decomposition, deep learning, neural parameter prediction, particle swarm optimization, contrast enhancement

I. INTRODUCTION

Optical satellite and aerial images are often captured under unfavourable conditions: low sun elevation, haze and thin cloud, sensor limitations and quantisation produce images that are dark, washed out or low in contrast. Enhancement is therefore a routine pre-processing step before visual interpretation, mapping and automated analysis. For remote sensing, an enhancer must raise contrast *and* preserve the radiometric and spatial detail that later analysis relies on.

Transform-domain methods are popular for this purpose. The discrete wavelet transform (DWT) separates an image into an approximation (LL) band carrying illumination and three detail bands (LH, HL, HH) carrying edges and texture. Demirel *et al.* [2] equalised the LL band through its singular values (DWT-SVD). Later work tuned the transformation with nature-inspired algorithms (NIA), such as cuckoo search [3]. In our earlier work [1], a power-law operator on the LL band was optimised per image with PSO or GWO to match the mean intensity of the input, and then transferred through an SVD correction factor. That method outperformed HE-, DCT- and brute-force-based enhancement in terms of the contrast assessment function (CAF).

This paper re-examines that design and makes four observations: (i) the NIA runs an iterative search for every image, although images with similar statistics need similar parameters; (ii) the fitness function (mean matching) is a proxy and does not measure fidelity or detail; (iii) as we prove in Section III, the SVD step reduces exactly to multiplying LL by a scalar, so the tone *shape* is never changed; and (iv) because the detail bands are passed through unchanged while LL is brightened, weak details lose contrast relative to their surroundings, and sensor noise is not addressed.

At the same time, deep learning has advanced low-light enhancement. Zero-DCE [10], Zero-DCE++ [11], SCI [12], URetinex-Net [13] and Retinexformer [14] learn enhancement from data, and remote-sensing-specific networks such as GSA-Net [16] and LeRSGAN [17] have appeared. However, most are trained on natural-scene photographs and have tens of thousands to millions of parameters. As our experiments show, they can fail badly on remote sensing degradations such as haze, where the image is bright but low in contrast.

We propose **AI-DWTE** (AI-guided DWT Enhancement), which keeps the interpretable, lightweight wavelet pipeline of [1] and replaces the per-image NIA search with a learned predictor. Our contributions are:

- An analysis showing that the SVD correction of DWT-SVD/NIA methods is equivalent to a scalar LL gain, which explains their limited tonal control and the “weak detail” problem.

ISBN: 978-81-985770-4-7

- A differentiable-in-form eight-parameter wavelet operator with a new *gain-consistent detail transfer*: detail coefficients are shrunk to suppress noise and then scaled by the local slope of the tone curve, so that weak details are amplified along with the background.
- A neural parameter predictor (a deep ensemble of MLPs on 100 wavelet statistics), trained by distilling oracle parameters. It predicts all parameters in one forward pass and removes the iterative search.
- A reproducible benchmark on 1,260 test cases from the UC Merced dataset [19], covering three degradations (one unseen in training) and ten comparison methods, including pretrained deep networks run with their official weights. We report full-reference, no-reference, runtime, ablation and significance results.

II. RELATED WORK

A. *Histogram and Retinex methods:*

Global histogram equalisation (HE) maximises entropy but over-enhances and shifts brightness. CLAHE [5] limits local histogram slopes. MSRCR [6] and LIME [7] estimate an illumination component and divide it out.

B. *Transform-domain and NIA methods*:

DWT-SVD [2] scales the LL band by the ratio of the largest singular values of the equalised and original LL. Bhandari *et al.* [3] optimised a related DWT-SVD scheme with cuckoo search, and Sharma and Verma [4] used gamma correction in the DWT-SVD domain. Our base method [1] searches for a per-channel power-law exponent β with PSO [8] or GWO [9].

C. *Deep enhancement:*

Zero-DCE [10] predicts pixel-wise quadratic curves with a seven-layer CNN trained with non-reference losses. Zero-DCE++ [11] makes it lightweight with depth-wise convolutions, and SCI [12] learns a self-calibrated illumination module. Retinex-guided transformers [14] and deep unfolding networks [13]reach higher fidelity when paired supervision is available. Wave-Mamba brings wavelets into state-space models [15]. For remote sensing, GSA-Net [16] trains a U-Net on gamma-darkened images, and LeRSGAN [17] uses adversarial learning. AI-DWTE is closer in spirit to curve-estimation networks: it predicts a small number of interpretable operator parameters. Unlike those networks, it works in the wavelet domain, explicitly treats the detail bands, and needs only about 6×10^4 parameters per ensemble member, applied to a 100-dimensional descriptor.

III. REVISITING THE NIA–DWT–SVD BASE METHOD

For each channel, the base method [1] computes $\{LL, LH, HL, HH\} = \mathrm{DWT}(I_c)$ and finds $\beta^\star$ by PSO/GWO so that the mean of LL^β matches a target mean intensity (Eq. 12 of [1]). With the SVD $LL = U\Sigma V^\top$ and $LL^{\beta^\star} = U'\Sigma'V'^\top$, the correction factor is $\sigma = \max(\Sigma')/\max(\Sigma)$, and the corrected band is

$$LL_{\text{new}} = U(\sigma\Sigma)V^\top = \sigma\, U\Sigma V^\top = \sigma\, LL. \qquad (1)$$

The output is then $\mathrm{IDWT}(\sigma LL, LH, HL, HH)$. Two consequences follow. First, the power law affects the result *only* through the scalar σ: the method is a per-channel global gain on LL and cannot, for instance, brighten shadows more than highlights. Second, because the detail bands are unchanged, the ratio of detail to local mean falls by $1/\sigma$ whenever $\sigma > 1$, so faint textures (field boundaries, narrow roads) lose relative contrast. Eq. (1) applies equally to the DWT-SVD method of [2]. These observations motivate the operator below.

IV. PROPOSED METHOD: AI-DWTE

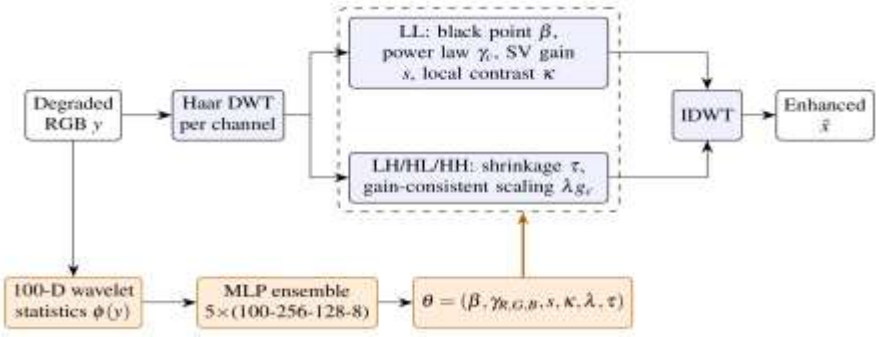


Fig. 1. AI-DWTE. A neural predictor (orange) maps wavelet statistics of the input to the eight parameters of a Haar-DWT enhancement operator (blue) in a single forward pass, replacing the per-image PSO/GWO search of [1].

Fig. 1 shows the pipeline. Images are in $[0,1]$. For each channel $c \in \{R, G, B\}$ we apply a one-level orthonormal Haar DWT and normalise the approximation band as $\ell_c = LL_c/2 \in [0,1]$.

A. *Enhancement Operator*

Black-point removal and adaptive power law. Haze adds a veil that a pure power law cannot remove, so we first remove an offset β and then apply a channel-

ISBN: 978-81-985770-4-7

wise exponent γ_c (which also corrects colour casts) and a gain s:

$$\tilde{\ell}_c = \text{clip}\left(\frac{\ell_c - \beta}{1-\beta}, \epsilon, 1\right), \quad t_c = s\,\tilde{\ell}_c^{\gamma_c}. \tag{2}$$

By Eq. (1), s plays the role of the SVD correction factor of [1]: it is a *learned singular-value gain*, $U(s\Sigma)V^\top$, applied after a non-linear tone curve rather than instead of one.

Local contrast. A band-limited unsharp term restores mid-frequency contrast: $t'_c = \text{clip}(t_c + \kappa\,(t_c - G_4 * t_c), 0, 1)$, where G_4 is a Gaussian with $\sigma = 4$ LL pixels.

Gain-consistent detail transfer. A detail δ riding on a background ℓ should be amplified by the local slope of the tone curve, since $f(\ell + \delta) \approx f(\ell) + f'(\ell)\,\delta$. We therefore compute the slope map

$$g_c = \mathcal{U}_3 * \text{clip}\left(\frac{s\,\gamma_c\,\tilde{\ell}_c^{\gamma_c - 1}}{1-\beta}, 0, 4\right), \tag{3}$$

($\mathcal{U}_3$ is a 3×3 box filter) and scale the soft-thresholded detail bands:

$$D'_c = \lambda\, g_c \odot \mathcal{S}_\tau(D_c), \quad D \in \{LH, HL, HH\}, \tag{4}$$

where $\mathcal{S}_\tau(u) = \text{sgn}(u)\max(|u| - \tau, 0)$ is wavelet shrinkage [18]. The shrinkage removes noise before it is amplified, and λ adds a global sharpening control. The output is $\hat{x}_c = \text{IDWT}\left(2t'_c, D'_{LH,c}, D'_{HL,c}, D'_{HH,c}\right)$. In the base method, by contrast, $g_c \equiv 1$ and $\tau = 0$.

The parameter vector $\theta = (\beta, \gamma_R, \gamma_G, \gamma_B, s, \kappa, \lambda, \tau)$ is bounded to $[0,0.6] \times [0.2,2.5]^3 \times [0.5,1.6] \times [0,2] \times [0,2.5] \times [0,0.06]$.

B. Neural Parameter Predictor

Descriptor. From the input we compute $\phi(y) \in \mathbb{R}^{100}$: for each channel, a 16-bin square-root histogram of ℓ_c, seven percentiles, its mean and standard deviation, and the log mean and log median magnitudes of the three detail bands. To these we add a median-absolute-deviation noise estimate from the finest diagonal band, the channel mean ratios (colour balance), mean saturation, and the mean and 1st percentile of the dark channel (veil cues).

Network. Each ensemble member is an MLP (100–256–128–8, ReLU, L_2 weight 10^{-1}, Adam, batch 64) that regresses θ normalised to $[0,1]$. Five members with different seeds are averaged (a deep ensemble), for a total of 298,920 parameters (59,784 per member). Prediction is one matrix-vector pass, which takes well under a millisecond.

Oracle distillation. Because the operator is cheap, we obtain training targets by solving, offline and for each training pair (y, x),

$$\theta^\star = \underset{\theta}{\text{argmin}}\; 10 \parallel \hat{x}(\theta) - \frac{x\|_2^2}{N} + \left(1 - \text{SSIM}_7(\hat{x}(\theta), x)\right), \tag{5}$$

Using bounded Powell search (at most 400 evaluations, best of three initial points), which takes about 8 s per image. The MLPs learn the mapping $\phi(y) \mapsto \theta^\star$. The expensive search is thus paid once, at training time, rather than for every image as in [1]. We also tried end-to-end fine-tuning of the ensemble on the image loss, back-propagating finite-difference gradients of the operator through the MLPs. It did not improve validation PSNR (25.22→24.81 dB after four epochs), so the distilled model is used.

V. EXPERIMENTAL SETUP

- **Data:** We use the UC Merced Land Use dataset [19]: 2,100 RGB aerial images (256×256, 0.3 m, 21 classes). Images 00–49 of each class (1,050) are used for training, with 15% held out for model selection, and images 80–99 (420) are used for testing, so there is no overlap. The 44 images that are not exactly 256×256 pixels are resampled to that size.
- **Degradations:** As in GSA-Net [16], paired data are synthesised from clean images. *D1 (under-illumination):* $y = (a x^g + b)\,\mathbf{c} + n$, with $g \sim U(1.3,2.6)$, $a \sim U(0.45,0.85)$, $b \sim U(0,0.06)$ and colour cast $\mathbf{c} \sim U(0.9,1.1)^3$. *D2 (haze):* $y = x\,t + A(1-t) + n$, with $t \sim U(0.35,0.7)$ and $A \sim U(0.55,0.9)$. In both, $n \sim \mathcal{N}(0, \sigma^2)$ with $\sigma \sim U(0,0.02)$. Training uses one random D1 or D2 degradation per image. *D3 (unseen)* is used only for testing: a smooth random illumination field in $[0.25,0.85]$ (cloud shadow or vignetting) combined with a power law, mild haze and noise. Every test image receives D1, D2 and D3, giving 1,260 test cases with fixed seeds.
- **Compared methods:** HE; CLAHE [5]; DWT-SVD [2]; the base NIA method [1] with PSO and with GWO (20 agents, 30 iterations, $\beta \in [0.05,10]$). Eq. (12) of [1] admits several readings, so we implement it on the $[0,1]$-normalised LL with a mid-grey target of 0.5 and follow Eqs. (13)–(16) exactly. We also compare MSRCR [6], LIME [7] (Gaussian-refined illumination map), and three deep networks run with their *official pretrained weights*: Zero-DCE [10] (79,416 parameters), Zero-DCE++ [11] (10,561) and SCI [12] ("medium" model). All methods are implemented in Python/NumPy. The deep models are re-implemented for CPU inference and load the authors' released weights.
- **Metrics:** Full-reference: PSNR, SSIM [20] and mean CIEDE2000 colour difference [21] (ΔE_{00}, lower is better). No-reference, as used in [1]: mean luminance, entropy H, CAF $= H \cdot C^{1/4}$ with C the RMS contrast of the luminance, and average gradient (AG). Run time is measured per image on a single CPU thread. Paired differences are tested with the Wilcoxon signed-rank test.

ISBN: 978-81-985770-4-7

VI. RESULTS AND DISCUSSION

TABLE I
FULL-REFERENCE COMPARISON ON 420 UC MERCED TEST IMAGES PER DEGRADATION (MEAN). BEST AMONG METHODS IN BOLD. ΔE: CIEDE2000, LOWER IS BETTER. D3 IS NOT SEEN IN TRAINING.

	D1: under-illumination			D2: haze			D3: non-uniform (unseen)			All 1,260 cases		
Method	**PSNR↑**	**SSIM↑**	**ΔE ↓**	**PSNR**	**SSIM**	**ΔE**	**PSNR**	**SSIM**	**ΔE**	**PSNR**	**SSIM**	**ΔE**
Degraded input	11.97	0.601	23.76	17.01	0.776	13.81	13.92	0.686	18.98	14.30	0.688	18.85
HE	14.74	0.540	20.26	14.81	0.536	20.09	13.59	0.600	20.40	14.38	0.559	20.25
CLAHE [5]	16.36	0.650	15.23	17.97	0.625	11.91	16.51	0.700	13.43	16.94	0.658	13.52
DWT-SVD [2]	18.46	0.786	13.38	17.47	0.758	14.54	17.35	0.813	14.10	17.76	0.786	14.01
Base NIA–PSO [1]	18.63	0.792	13.42	17.57	0.755	14.31	17.64	0.815	13.83	17.95	0.787	13.85
Base NIA–GWO [1]	18.63	0.792	13.43	17.57	0.755	14.32	17.64	0.815	13.83	17.95	0.787	13.86
LIME [7]	14.95	0.646	18.56	7.86	0.639	32.53	11.60	0.707	23.12	11.47	0.664	24.74
MSRCR [6]	16.26	0.681	16.54	17.22	0.684	15.69	17.44	0.712	14.40	16.97	0.692	15.54
Zero-DCE [10]	20.89	0.827	11.01	11.20	0.667	25.46	16.17	0.753	15.85	16.08	0.749	17.44
Zero-DCE++ [11]	20.04	0.817	13.49	10.30	0.643	28.04	14.97	0.744	19.61	15.10	0.735	20.38
SCI [12]	14.29	0.730	18.22	5.42	0.326	39.67	8.67	0.621	29.60	9.46	0.559	29.16
AI-DWTE (ours)	**23.04**	**0.844**	**9.07**	**24.31**	**0.898**	**8.21**	**17.79**	**0.812**	**11.96**	**21.71**	**0.851**	**9.75**

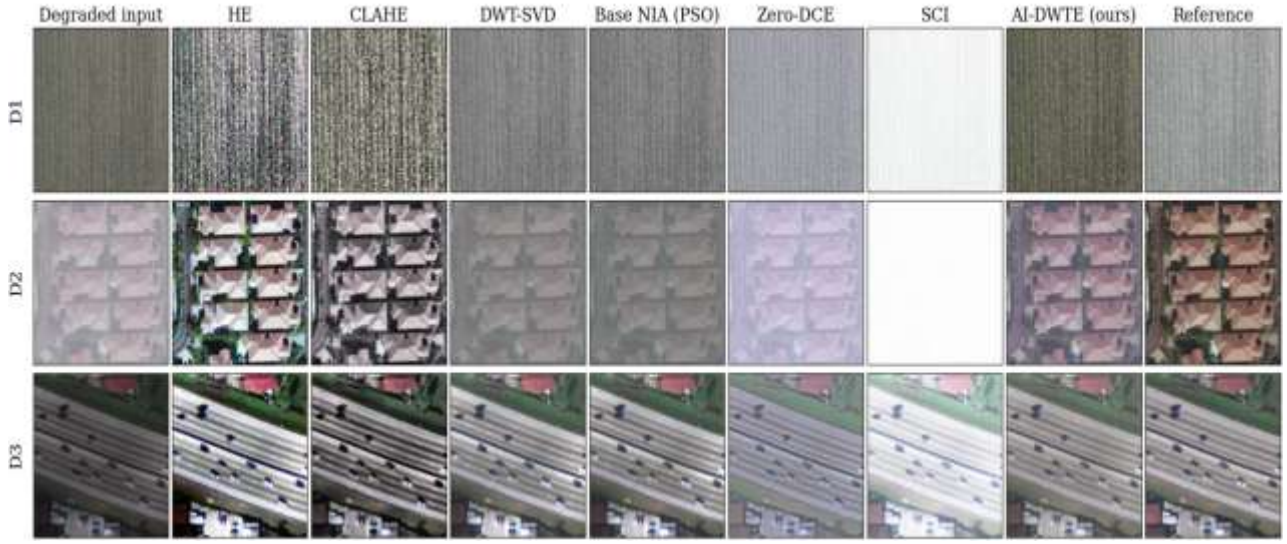


Fig. 2. Visual comparison on test images (agricultural80/D1, denseresidential80/D2, freeway80/D3; examples fixed before the results were computed). AI-DWTE removes the haze veil and colour cast without the noise amplification of HE/CLAHE or the saturation of SCI. In the D1 example the clean scene is unusually bright, and AI-DWTE, like any blind method, under-estimates the target brightness.

- **Full-reference quality.** Table I summarises the results. On the two degradation families seen in training, AI-DWTE is best on every metric. For D1 it reaches 23.04 dB and SSIM 0.844, against 20.89 dB and 0.827 for the strongest competitor (Zero-DCE). This is +2.15 dB higher on average and better on 73.1% of images ($p = 4 \times 10^{-21}$). Compared with the PSO base method, the gain is +4.41 dB (better on 84.3% of images). Haze (D2) separates the methods most clearly. Low-light networks assume dark inputs and therefore over-brighten already bright hazy images: Zero-DCE falls to 11.20 dB and SCI to 5.42 dB. The global-gain methods (DWT-SVD, base NIA) cannot remove the veil because Eq. (1) only rescales it. AI-DWTE reaches 24.31 dB, which is +6.74 dB over the base method and +6.34 dB over the best baseline (CLAHE), with better SSIM on 98.3% of images than the base method. Over all 1,260 cases, AI-DWTE improves on the PSO base method from 17.95 to 21.71 dB, from 0.787 to 0.851 SSIM, and from 13.85 to 9.75 ΔE_{00}. PSO and GWO give practically identical results, because the 1-D search converges for both.
- **Generalisation (D3).** Under the unseen spatially varying illumination, the advantage largely disappears.

ISBN: 978-81-985770-4-7

AI-DWTE is still best in PSNR (17.79 dB) and colour error (11.96), but its PSNR gain over the base method (+0.15 dB) is not significant ($p = 0.16$), and its SSIM (0.812) is statistically tied with the base method (0.815) and DWT-SVD (0.813). The parameters are global per image (except κ), so a single tone curve cannot correct both the shadowed and the lit parts of a scene. Spatially varying parameters (for example, predicted on a coarse grid) are the natural extension. AI-DWTE still outperforms all three deep networks on D3, by +1.62 dB over Zero-DCE.

- **Visual quality.** Fig. 2 shows that HE and CLAHE produce strong contrast but amplify noise and create blotchy textures. SCI and LIME saturate hazy scenes. DWT-SVD and the base method brighten without removing the veil. AI-DWTE restores contrast and colour, and its shrinkage keeps flat regions clean.

TABLE II
NO-REFERENCE METRICS (ALL 1,260 CASES) AND RUN TIME PER 256 × 256 IMAGE ON ONE CPU THREAD.

Method	Lum.	H	CAF	AG	ms
HE	128.8	7.96	23.13	25.90	10.9
CLAHE [5]	105.1	7.54	20.35	23.06	27.6
DWT-SVD [2]	130.4	6.56	15.93	9.23	20.3
Base NIA–PSO [1]	124.7	6.51	15.70	9.06	114.3
Base NIA–GWO [1]	124.7	6.51	15.70	9.06	117.0
LIME [7]	188.1	6.35	14.49	15.11	4.8
MSRCR [6]	136.6	7.49	20.49	20.54	304.0
Zero-DCE [10]	148.0	5.99	12.71	8.64	2796.5
Zero-DCE++ [11]	154.2	6.01	12.77	8.78	480.1
SCI [12]	208.9	5.20	11.71	7.99	7.5
AI-DWTE (ours)	124.5	6.57	15.78	9.60	26.6
Clean reference	120.0	6.71	16.78	11.29	–

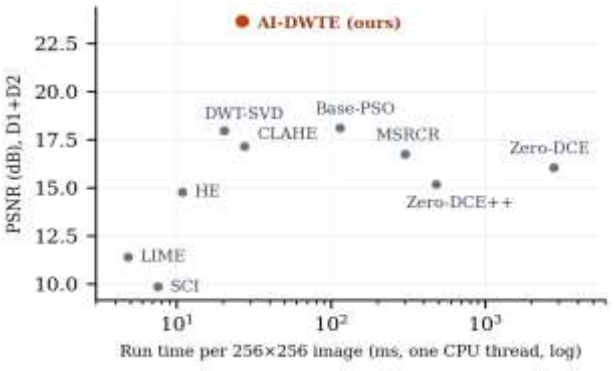


Fig. 3. Quality versus run time (PSNR averaged over D1 and D2). AI-DWTE achieves the highest fidelity at a cost comparable to CLAHE and 4.3 × lower than the PSO search of [1].

- **No-reference metrics and a caution.** Table II shows that HE, CLAHE and MSRCR obtain the highest entropy, CAF and average gradient, far *above* the clean reference (CAF 16.78). Their low PSNR and high ΔE in Table I show that these values reflect over-enhancement and amplified noise, not recovered information. AI-DWTE's no-reference values (Lum. 124.5, H 6.57, CAF 15.78, AG 9.60) are the closest to the clean reference in luminance, entropy and average gradient. Only DWT-SVD is marginally closer in CAF. We therefore recommend reporting CAF-type scores alongside full-reference fidelity, never alone.

- **Run time.** AI-DWTE needs 26.6 ms per image (Fig. 3): the descriptor, the MLP pass and a single application of the operator. The PSO base method needs 114.3 ms, because it evaluates the fitness about 1,860 times per image. Our NumPy CPU implementation of Zero-DCE takes 2.8 s per image, and it would be much faster on a GPU. AI-DWTE needs no GPU at all, which suits on-board or edge processing.

TABLE III
ABLATION OF THE OPERATOR COMPONENTS (ALL 1,260 TEST CASES). COMPONENTS ARE DISABLED AT INFERENCE WITH THE SAME PREDICTOR.

Variant	PSNR	SSIM	ΔE
Full model	21.71	0.851	9.75
w/o gain-consistent detail (HF untouched)	21.36	0.835	9.85
w/o local contrast ($\kappa = 0$)	21.46	0.841	9.85
w/o wavelet shrinkage ($\tau = 0$)	21.62	0.836	10.07
shared γ for R,G,B	21.69	0.851	9.77

- **Ablation.** In Table III, the gain-consistent detail transfer contributes the largest improvement: removing it, as in the base method, costs 0.36 dB and 0.016 SSIM. Removing shrinkage costs 0.015 SSIM and raises the colour error. This confirms that noise must be treated before the details are amplified. Local contrast adds 0.25 dB. Sharing γ across channels changes little on this dataset, so the colour correction comes mainly from the black point and the gain.

TABLE IV
CHOICE OF PARAMETER PREDICTOR (157 HELD-OUT TRAINING-SPLIT IMAGES, D1/D2) AND THE ORACLE UPPER BOUND.

Predictor	PSNR	SSIM	ΔE
Per-image PSO search (base [1])	18.97	0.796	12.33
Ridge regression	24.44	0.877	7.66
k-NN (k=7)	24.22	0.890	7.68
Random forest	24.57	0.890	7.54
MLP (single)	24.98	0.900	7.36
MLP ensemble (5)	25.22	0.904	7.18
Oracle parameters	32.31	0.929	3.83

- **Predictor and data size.** Table IV shows that all learned predictors beat the per-image search, which confirms that θ can be predicted from image statistics. The MLP ensemble is the most accurate. The oracle parameters, which require the ground truth, reach 32.31 dB, so the operator itself has substantial headroom. The remaining gap comes from the predictor, and part of it cannot be closed: a blind

ISBN: 978-81-985770-4-7

method cannot know whether a dark scene is dark or darkened (Fig. 2, first row). Validation PSNR is still rising with training set size (22.76, 23.75, 24.72 and 25.21 dB for 89, 223, 446 and 893 images), so more, and more diverse, training data should help further.

TABLE V
NO-REFERENCE METRICS ON THE 30 DARKEST ORIGINAL (UNDEGRADED) TEST IMAGES: 19 FOREST, 9 RIVER, 2 GOLF COURSE.

Method	Lum.	*H*	CAF	AG
Original (dark)	54.9	6.07	12.99	8.75
HE	129.8	7.75	22.59	31.17
CLAHE [5]	104.9	7.33	19.41	22.29
DWT-SVD [2]	131.6	7.23	18.96	16.74
Base NIA–PSO [1]	126.3	7.18	18.68	16.35
LIME [7]	168.7	7.12	18.32	25.68
MSRCR [6]	167.7	7.24	19.18	20.91
Zero-DCE [10]	126.3	6.43	14.67	13.98
Zero-DCE++ [11]	136.0	6.49	15.01	14.60
SCI [12]	184.9	6.85	17.50	18.17
AI-DWTE (ours)	120.7	6.72	16.04	13.98

- **Real low-brightness images.** On the 30 darkest original test images, which have no synthetic degradation (Table V), AI-DWTE raises mean luminance from 54.9 to 120.7 and entropy from 6.07 to 6.72. It is deliberately more conservative than HE, CLAHE or MSRCR, which reach higher CAF. Because these scenes (dense forest, water) are dark in reality, stronger enhancement is not necessarily more faithful. Without ground truth this experiment only shows that the method behaves sensibly outside the synthetic protocol.

Limitations. (i) Training pairs are synthetic, and real sensor degradations (e.g. with atmospheric correction residuals or compression) may differ. (ii) The parameters are global per image, which limits performance under non-uniform illumination (D3). (iii) Only three-band RGB imagery at 0.3 m was tested. Multispectral bands and other resolutions and sensors remain to be evaluated. (iv) Downstream benefits, for example to classification or detection, were not measured.

VII. CONCLUSION

We showed that the SVD correction of NIA-tuned DWT-SVD enhancement is equivalent to a scalar gain on the approximation band, which explains both its limited tonal control and its attenuation of weak details. AI-DWTE replaces the per-image PSO/GWO search with a neural parameter predictor and extends the operator with black-point removal, a learned singular-value gain, local contrast and gain-consistent, noise-aware detail transfer. On 1,260 UC Merced test cases it outperformed ten classical, NIA-based and pretrained deep methods. It improved PSNR over the PSO base method by 3.8 dB on average (4.4 dB and 6.7 dB on the two in-distribution degradations) at a quarter of the run time, and it runs without a GPU. Future work will predict spatially varying parameters to handle non-uniform illumination, extend the method to multispectral bands, and evaluate downstream tasks.

REFERENCES

[1] R. Sharma, M. Ravinder, N. Sharma, and K. Sharma, “An optimal remote sensing image enhancement with weak detail preservation in wavelet domain,” manuscript, 2021. [Authors: insert final publication venue]

[2] H. Demirel, C. Ozcinar, and G. Anbarjafari, “Satellite image contrast enhancement using discrete wavelet transform and singular value decomposition,” *IEEE Geosci. Remote Sens. Lett.*, vol. 7, no. 2, pp. 333–337, 2010.

[3] A. K. Bhandari, V. Soni, A. Kumar, and G. K. Singh, “Cuckoo search algorithm based satellite image contrast and brightness enhancement using DWT–SVD,” *ISA Trans.*, vol. 53, no. 4, pp. 1286–1296, 2014.

[4] N. Sharma and O. P. Verma, “Gamma correction based satellite image enhancement using singular value decomposition and discrete wavelet transform,” in *Proc. IEEE ICACCCT*, 2014, pp. 1286–1289.

[5] K. Zuiderveld, “Contrast limited adaptive histogram equalization,” in *Graphics Gems IV*, Academic Press, 1994, pp. 474–485.

[6] D. J. Jobson, Z. Rahman, and G. A. Woodell, “A multiscale retinex for bridging the gap between color images and the human observation of scenes,” *IEEE Trans. Image Process.*, vol. 6, no. 7, pp. 965–976, 1997.

[7] X. Guo, Y. Li, and H. Ling, “LIME: Low-light image enhancement via illumination map estimation,” *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 982–993, 2017.

[8] J. Kennedy and R. Eberhart, “Particle swarm optimization,” in *Proc. IEEE ICNN*, 1995, pp. 1942–1948.

[9] S. Mirjalili, S. M. Mirjalili, and A. Lewis, “Grey wolf optimizer,” *Adv. Eng. Softw.*, vol. 69, pp. 46–61, 2014.

[10] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, and R. Cong, “Zero-reference deep curve estimation for low-light image enhancement,” in *Proc. IEEE/CVF CVPR*, 2020, pp. 1780–1789.

[11] C. Li, C. Guo, and C. C. Loy, “Learning to enhance low-light image via zero-reference deep curve estimation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 8, pp. 4225–4238, 2022.

[12] L. Ma, T. Ma, R. Liu, X. Fan, and Z. Luo, “Toward fast, flexible, and robust low-light image enhancement,” in *Proc. IEEE/CVF CVPR*, 2022, pp. 5637–5646.

ISBN: 978-81-985770-4-7

[13] W. Wu, J. Weng, P. Zhang, X. Wang, W. Yang, and J. Jiang, "URetinex-Net: Retinex-based deep unfolding network for low-light image enhancement," in *Proc. IEEE/CVF CVPR*, 2022, pp. 5891–5900.

[14] Y. Cai, H. Bian, J. Lin, H. Wang, R. Timofte, and Y. Zhang, "Retinexformer: One-stage Retinex-based transformer for low-light image enhancement," in *Proc. IEEE/CVF ICCV*, 2023, pp. 12504–12513.

[15] W. Zou, H. Gao, W. Yang, and T. Liu, "Wave-Mamba: Wavelet state space model for ultra-high-definition low-light image enhancement," in *Proc. ACM Multimedia*, 2024, pp. 1534–1543.

[16] M. Zhao, R. Yang, M. Hu, and B. Liu, "Deep learning-based technique for remote sensing image enhancement using multiscale feature fusion," *Sensors*, vol. 24, no. 2, Art. 673, 2024.

[17] T. Li, T. Guo, and D. Xiang, "LeRSGAN: A GAN-based model for low-light remote sensing image enhancement," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 18, pp. 26489–26504, 2025.

[18] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation by wavelet shrinkage," *Biometrika*, vol. 81, no. 3, pp. 425–455, 1994.

[19] Y. Yang and S. Newsam, "Bag-of-visual-words and spatial extensions for land-use classification," in *Proc. ACM SIGSPATIAL GIS*, 2010, pp. 270–279.

[20] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, 2004.

[21] G. Sharma, W. Wu, and E. N. Dalal, "The CIEDE2000 color-difference formula: Implementation notes, supplementary test data, and mathematical observations," *Color Res. Appl.*, vol. 30, no. 1, pp. 21–30, 2005.

Cite this article as: Rajni, Ravinder "AI-DWTE: Learning-Guided Wavelet-Domain Enhancement of Low-Contrast Remote Sensing Images with Gain-Consistent Detail Preservation", Proceedings of CCSN2025: (14th International Conference on Computing, Communication and Sensor Network).

Displayed as online in 12th Novemberr 2026.

AOI: 10.100.232025.00101

Link: https://actsoft.org/science/ccsn2025-pro/59-ccsn2026.pdf

@Copyright to 'Applied Computer Technology', Kolkata, WB, India. Website: https://actsoft.org, Email: info@actsoft.org,

ISBN: 978-81-985770-4-7