Published by :
https://www.ijert.org/
An International Peer-Reviewed Journal

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

DSN-RD: Domain-Statistics-Driven Residual Diffusion for Data-Efficient, Compute-Aware Super-Resolution in Frequency-Shifted Imagery

C. Kowsika T. Kaaviya Varshini, Sajana Parveen, Annanya R K
Mr. N. Ravi, Mrs. G. Aishwaryalakshini

Abstract—Diffusion models are well-suited for image super-resolution (SR), but suffer from challenges in domains with non-natural image frequency statistics, limited paired training data, and limited sampling budget. Adversarial methods can handle some of these conditions, but can suffer from instability during training and augmentation dependence. We introduce DSN-RD, a conditional residual-diffusion image super-resolution method built with the following insight: compute only a few domain statistics once and use them in three ways: (1) as domain shift and data scarcity prior for the diffusion model (radially averaged residual frequency spectrum), (2) as a mechanism to avoid spurious detail in flat regions (stability without a discriminator), and (3) as a sampling budget allocation to image tiles based on their level of detail using DDIM. A common label is used to condition on augmentation and on real vs. simulated pairs, and is identity at inference. We describe the model, training and inference methods and present a pre-defined comparison and ablation methodology along with decision rules. No experiments have been performed and thus all potential advantages are presented as hypotheses. We also point out that the normalization used is equivalent to a transformation of the standard DDPM and thus any results are likely to be empirical rather than theoretical. Additionally, further research is needed into related work in order to verify novelty.

Index Terms - Image super-resolution, diffusion models, domain shift, frequency statistics, residual learning, adaptive sampling, data-efficient learning, remote sensing.

I. Introduction

Existing super-resolution approaches tackle domain-specific and limited-data challenges through fragmented methods, including GANs prone to adversarial instability [2, 3, 5-7], augmentations that risk distribution bias [5], and frequency-specific analyses of domain shift [6, 7]. Meanwhile, current diffusion models rely on fixed sampling budgets and remain largely untested in low-data, domain-shifted scenarios [1, 8], while deployment costs and real-time constraints are frequently overlooked or unreported in prior literature [3, 5-7, 9, 10].

Research gap: The reviewed papers do not combine (i) a diffusion generator, (ii) a generative target normalized by measured domain statistics, and (iii) a sampling budget driven by the same statistics, in a setting diagnosed as frequency-shifted [6]. Whether such a combination improves stability, data efficiency, and average cost is an open empirical question.

Motivation:DSN-RD (Domain-Statistics-Driven Residual Diffusion) estimates two quantities once, a residual spectrum and a detail-scale map, and uses them for the noise prior, a flatregion guard, and a per-tile compute allocator. A shared tag handles augmentation and simulated pairs.

The contributions of this paper are:

1. A formulation of conditional residual diffusion for SR in which the diffusion noise is shaped by an estimated domain spectrum and a predicted detail-scale map, trained with a plain noise-prediction loss and no discriminator.

2. A tile-adaptive DDIM step budget driven by the same detail map, so that flat tiles receive fewer sampling steps than dense tiles.

3. A shared tag that routes augmentation and simulated pairs through one conditioning mechanism, with spectrum estimation restricted to unaugmented data and identity tagging at inference.

4. . A re-specified validation and ablation protocol with decision rules, including explicit criteria under which individual modules should be dropped.

5. The components are potentially novel only in their combination; broader literature verification is required before any novelty claim (Section X)

II. RELATED WORK

Papers [1–10] are the primary literature basis. Papers [4,6,9,10] were read at abstract level only, and their characterization here is correspondingly limited. Table I summarizes the.

A. Diffusion-based vision and SR methods.

Diffusion methods [1,8] provide the conditional DDPM/DDIM generative core used by DSN-RD. Their strengths are a simple regression-style loss and stable optimization. Their limitations for our set-ting are a single global sampling-step count and the absence of evaluation on domain-shifted and low-data tasks.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

Table 1: Comparison of Existing Methods (as characterized in the reviewed papers)

Ref.	Theme	Limitation relevant to DSN-RD
[1, 8]	Diffusion	Global step count; untested on domain-shifted, low-data SR
[2, 3]	Adversarial	Training instability; deployment cost largely unreported [3]
[5]	GAN, augmentation	Augmentation can bias learned distribution; instability
[6]	Frequency-aware SR (CDGAN)	Remedy remains adversarial; abstract-level reading
[7]	Simulated pairs (underwater)	Adversarial; other domain; synthetic–real gap
[9]	One-class detection	Discriminative; not an SR method
[10]	DRL resource allocation	Networking problem; not an SR method
4]	Trust in GenAI	Not algorithmic

Table 2: DSN-RD Components and Corresponding Research Gaps

Problem	Component	Status
P1 Scarce data	Unpaired spectrum; simulated pairs (tag)	Partial; hypothesis
P2 Domain shift	Module 1 normalizer Ψ_y	Addressed by design; unverified
P3 Augmentation shift	Module 2 tag	Unproven; no guarantee
P4 Instability	No discriminator	By removal; to be verified
P5 Slow sampling	Modules 1 and 3	Hypothesized
P6 Deployment cost	Module 3; efficiency protocol	Partial
P7 Flat-region hallucination	Detail map m_φ guard	Hypoth

than a new diffusion theory.

B. GAN-based and adversarial SR. Adversarial approaches [2, 3, 5–7] target realistic detail and are the reference approach for limited-data and domain-specific imaging. Their recurring limitation is training instability. DSN-RD responds by removal: it uses no discriminator and instead constrains hallucination through the detail-scale map and consistency projection.

C. Domain- and frequency-aware methods. The frequency-aware analysis in [6] (CDGAN) diagnoses a frequency mismatch as a cause of SR failure in remote sensing, but the remedy remains adversarial. Paper [7] handles a different domain (underwater color correction) and contributes the idea of simulated pairs. DSN-RD moves the frequency principle from the discriminator into the noise prior, and reuses simulated pairs through a tag; it does not claim to solve color correction.

D. Data-efficient and augmentation-based approaches. Augmentation for data-limited GAN training raises the concern that augmentations leak into the learned distribution [5]. Simulated pairs [7] expand paired data at the risk of a synthetic-real gap. One-class detection without attacker data [9] addresses scarcity in a discriminative form that DSN-RD does not target. DSN-RD uses a single tag τ for both, estimates the spectrum only on unaugmented data, and offers no guarantee against leakage.

E. Efficient sampling and deployment cost. Existing diffusion studies fix one step count for the whole image [1, 8]; GAN-based studies rarely report deployment cost [3, 5–7, 9]; and real-time constraints are stressed in a networking context [10]. Trust in generative outputs is a separate, non-algorithmic concern [4]. DSN-RD allocates steps per tile and specifies a standard efficiency protocol; it does not address communication-constrained resource allocation [10] or software-practice trust [4].

III. Problem Formulation and Research Gap

III.1 Conditional SR

Let $x \in \mathrm{R}^{H\times W}$ denote an HR image and $y = D(x)$ its LR observation under a known or estimated degradation operator $D(\cdot)$ (noise omitted for brevity). Let $U(\cdot)$ be an upsampling operator and $\hat{y} = U(y)$. The goal is to learn $p(x \mid y)$ from a small paired set $P = \{(y_i, x_i)\}$ an unpaired HR set U, and optionally a degradation simulator Deg(·), in a domain whose residual spectrum differs from that of natural images.

III.2 Recurring problems

Spot on. Those seven bottlenecks form the exact checklist of challenges that the DSN-RD framework—with its four-stage training pipeline, domain-statistics normalization, and adaptive sampling—is engineered to solve. Let me know if you would like to dive deeper into any specific stage or component!

III.3 Scope of DSN-RD

Table II maps each problem to the DSN-RD mechanism and status. DSN-RD addresses P1–P5 and P7 partially or hypothetically, and P6 only through an efficiency protocol and reduced average step count. It does not address non-Euclidean or discrete diffusion theory [1, 2, 8], spoof detection [9], communication-constrained resource allocation [10], or trust in generative AI [4]. It does not solve underwater color correction [7]; only the simulated-pair idea is reused, and transfer to other domains is a future test.

IV. Proposed DSN-RD Method

DSN-RD performs conditional SR in the residual domain with a diffusion model whose noise is shaped by domain statistics. The processing chain is: LR image → upsampling → detailscale prediction → domain-statistics normalization → tagged conditioning → conditional diffusion U-Net → tile-adaptive DDIM sampling → residual refinement/projection → HR image (Fig. 1). Upsampling and tiling. The LR image is upsampled to ŷ, which serves as a low-band anchor, and split into overlapping tiles for memory-bounded, per-tile processing. Residual anchor. The network models only the residual R0 = x −ŷ, so low-band content is inherited from ŷ rather than

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

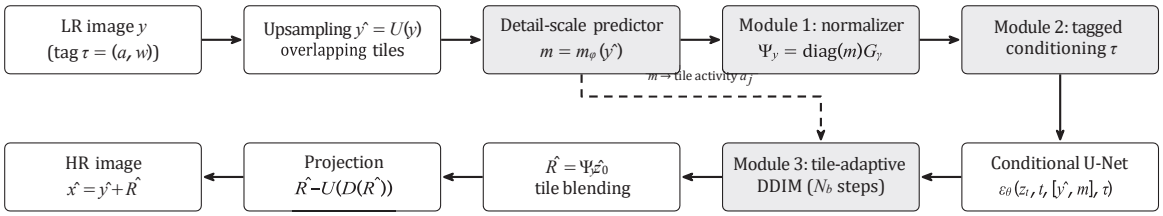


Figure 1: Overall DSN-RD architecture (schematic; shaded blocks are the proposed modules). The same detail map m is used as noise prior, flat-region guard, and compute allocator. **[Replace with final vector figure.]**

generated. This residual-plus-projection device is known outside the reviewed corpus and is reused, not claimed.

Detail-scale predictor. A network m_φ predicts, from $\hat{y}$ alone, a per-pixel residual scale map m. It is trained once (stage 1) and frozen. Because m is computed from the LR input, it is available at inference.

Domain-statistics normalization (Module 1). The radially averaged residual power spectral density (PSD) S_R is estimated once from unaugmented data and converted to a spectral gain G_γ. Combined with m, it forms the normalizer $\Psi_y = \mathrm{diag}(m)\, G_\gamma$, which scales and colors the diffusion noise to match the measured residual statistics. Diffusion then operates in the normalized space, where the target has approximately unit scale.

Tagged conditioning (Module 2). A tag $\tau = (a, w)$ records the applied augmentation a and a real/synthetic flag w for simulated pairs. Augmentations and simulated pairs thus enter training through one mechanism. The tag is dropped with probability p during training and set to identity at inference, with the aim that the model can separate augmentation effects from the clean-domain distribution. Spectrum-altering augmentations are excluded from spectrum estimation. Whether the tag prevents leakage is unproven.

Conditional diffusion U-Net. The denoiser ε_θ receives the noisy normalized residual z_t, timestep t, the conditioning $[\hat{y}, m]$, and τ, and predicts the noise with a plain ℓ_2 loss.

Tile-adaptive DDIM sampling (Module 3). Each tile's mean detail scale defines an activity a_j; tiles are grouped into B buckets and sampled with bucket-specific DDIM step counts.

Residual refinement and projection. The sampled normalized residual is mapped back by Ψ_y, tiles are blended, and the component of $\hat{R}$ visible to D is removed before adding $\hat{y}$.

Three uses of the domain statistics. (i) *Noise prior*: Ψ_y shapes the noise so that the prior matches the domain's residual statistics, addressing domain shift and data scarcity (P1, P2). (ii) *Hallucination guard*: where m is small, the reconstructed residual $\Psi_y \hat{z}_0$ is small in amplitude, limiting spurious detail in flat regions (P7) without a discriminator (P4). (iii) *Compute allocator*: the same m defines tile activity and step budgets (P5, P6).

V. Mathematical Formulation

Residual target.

$$R_0 = x - \hat{y}, \qquad \hat{y} = U(y).$$

Domain-statistics normalizer.

$$\Psi_y = \mathrm{diag}\left(m_\varphi(\hat{y})\, G_\gamma, \quad g(\kappa) = \max\left(S_R(\kappa)/\bar{S}_R\right)^{\gamma/2}, g_{\min}\right), \tag{2}$$

where $m_\varphi(\hat{y})$ is the predicted per-pixel detail scale, clamped to $[m_{\min}, m_{\max}]$; $S_R(\kappa)$ is the radially averaged residual PSD at radial frequency κ, estimated once from unaugmented data (from $R = x - U(D(x))$ for HR images; **[confirm]**); $\bar{S}_R$ is the PSD reference level (**[confirm definition]**); $\gamma \geq 0$ controls the strength of spectral shaping; $g_{\min} > 0$ is a gain floor; and G_γ is the spectral gain operator with radial response $g(\kappa)$. With the clamps, Ψ_y is invertible.

Forward diffusion. With noise schedule $\{\bar{\alpha}_t\}_{t=1}^{T}$ and $\varepsilon \sim N(0, I)$,

$$R_t = \sqrt{\bar{\alpha}_t}\, R_0 + \sqrt{1 - \bar{\alpha}_t}\, \Psi_y \varepsilon. \tag{3}$$

Normalized diffusion space. With $z_t = \Psi_y^{-1} R_t$ and $z_0 = \Psi_y^{-1} R_0$,

$$z_t = \sqrt{\bar{\alpha}_t}\, z_0 + \sqrt{1 - \bar{\alpha}_t}\, \varepsilon, \tag{4}$$

which is exactly the standard DDPM marginal. Module 1 is therefore a conditional preconditioning of the target, not a new diffusion process (see Section X).

Noise-prediction loss.

$$L_\varepsilon = \mathrm{E}\left\| \varepsilon - \varepsilon_\theta(z_t, t, [\hat{y}, m], \tau) \right\|^2, \tag{5}$$

where $[\hat{y}, m]$ is channel-wise concatenation and $\tau = (a, w)$ is replaced by a null tag with probability p.

Detail-map loss (stage 1 only).

$$L_{\text{map}} = \mathrm{E}\left\| \log m_\varphi(\hat{y}) - \log(s + \varepsilon_0) \right\|^2, \quad s = \sqrt{B * R_0^2}, \tag{6}$$

where B is a local averaging kernel, $*$ denotes convolution, R_0^2 is the elementwise square, and $\varepsilon_0 > 0$ is a small constant. Thus s is the local residual scale.

Tile activity. For tile j with pixel set Ω_j,

$$a_j = \frac{1}{|\Omega_j|} \sum_{u \in \Omega_j} m_\varphi(\hat{y})(u), \tag{7}$$

where $\tilde{m}$ is the median activity over the training set.

Adaptive step budget. Thresholds $\{\theta_b\}$ set as quantiles of a_j on the training set, define B buckets; tile j is in bucket b if

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

Algorithm 1 DSN-RD Training

Input: Paired set P, unpaired HR set U, operators U, D, optional Deg, schedule$\{\bar{\alpha}_t\}T$, γ, $g_{\min}$, $[m_{\min}, m_{\max}]$, augmentation family A, tag-dropout p, buckets B

Output: θ, φ, G_γ, $\{\theta_b\}$

1: **Stage 0:** estimate S_R from unaugmented $x \in P \cup U$; build G_γ (Eq. 2)
2: **Stage 1:** train m_φ with L_{map} (Eq. 6); freeze
3: **Stage 2: for** each iteration **do**
4: draw $(y, x) \sim P$, or $(\text{Deg}(x), x)$ with $x \sim U$ and $w =$ synthetic
5: draw $a \sim A$, apply jointly; $\tau = (a, w)$; with prob. p set $\tau = \emptyset$
6: $\hat{y} = U(y)$; $R_0 = x - \hat{y}$; $m = \text{clamp}(m_\varphi(\hat{y}))$
7: $\Psi_y = \text{diag}(m)G_\gamma$; $z_0 = \Psi_y^{-1} R_0$
8: $t \sim U\{1..T\}$; $\varepsilon \sim N(0, I)$; $z_t = \sqrt{\bar{\alpha}_t} z_0 + \sqrt{1-\bar{\alpha}_t}\varepsilon$
9: update θ on $\|\varepsilon - \varepsilon_\theta(z_t, t, [\hat{y}, m], \tau)\|^2$ (Eq. 5)
10: **Stage 3:** compute a_j (Eq. 7) on training tiles; set $\{\theta_b\}$ as quantiles

Algorithm 2 DSN-RD Inference

Input: LR image y; θ, φ, G_γ, $\{\theta_b\}$ step counts $\{N_b\}$

Output: $\hat{x}$

1: $\hat{y} = U(y)$; tile $\hat{y}$ with overlap; $m = \text{clamp}(m_\varphi(\hat{y}))$; $\Psi_y = \text{diag}(m)G_\gamma$
2: compute a_j per tile (Eq. 7); assign bucket $b(j)$ (Eq. 8); group tiles by bucket
3: **for** each bucket b **do**
4: $z \sim N(0, I)$ for its tiles; $\tau =$ identity
5: for t in DDIM schedule with N_b steps **do**
6: $\hat{\varepsilon} = \varepsilon_\theta(z, t, [\hat{y}, m], \tau)$; $\hat{z}_0 = (z - \sqrt{1-\bar{\alpha}_t}\hat{\varepsilon})/\sqrt{\bar{\alpha}_t}$
7: $z = \sqrt{\bar{\alpha}_{t'}}\hat{z}_0 + \sqrt{1-\bar{\alpha}_{t'}}\hat{\varepsilon}$ (Eqs. 9–11)
9: **end for**
10: $\hat{R} = \Psi_y \hat{z}_0$ per tile; blend tiles; $\hat{R} \leftarrow \hat{R} - U(D(\hat{R}))$; $\hat{x} = \hat{y} + \hat{R}$ (Eq. 12)

$\theta_{b-1} \le a_j < \theta_b$ (with $\theta_0 = 0, \theta_B = \infty$) and is sampled with N_b DDIM steps,

$$N_1 \le N_2 \le \cdots \le N_B. \tag{8}$$

The network is trained across all noise levels, so any DDIM grid is admissible. That few steps suffice for low-detail tiles is a hypothesis.

Reverse (DDIM) sampling. For consecutive timesteps $t \to t'$ on the bucket's schedule,

$$\hat{\varepsilon} = \varepsilon_\theta(z_t, t, [\hat{y}, m], \tau), \tag{9}$$

$$\hat{z}_0 = z_t - \sqrt{1-\bar{\alpha}_t}\,\hat{\varepsilon} \,/\, \sqrt{\bar{\alpha}_t}, \tag{10}$$

$$z_{t'} = \sqrt{\bar{\alpha}_{t'}}\,\hat{z}_0 + \sqrt{1-\bar{\alpha}_{t'}}\,\hat{\varepsilon}. \tag{11}$$

Residual reconstruction and projection.

$$\hat{R} = \Psi_y \hat{z}_0, \qquad \hat{R} \leftarrow \hat{R} - U\,D(\hat{R}), \qquad \hat{x} = \hat{y} + \hat{R}. \tag{12}$$

The projection removes the part of $\hat{R}$ that D can observe; it enforces low-band consistency when $D(U(\cdot)) \approx \text{id}$.

VI. Training and Inference

VI.1 Training

Training has four stages (Algorithm 1). Stage 0 estimates SR from unaugmented data and builds Gγ . Stage 1 trains and freezes mφ . Stage 2 trains εθ on paired data and, optionally, simulated pairs from U , applying augmentations jointly to both images and recording them in τ. Stage 3 calibrates bucket thresholds without gradients.

VI.2 Inference

At inference (Algorithm 2), the tag is set to identity, m is computed from $\hat{y}$, tiles are bucketed by activity, and each bucket is sampled with its own DDIM schedule. Tiles are blended, projected, and added to $\hat{y}$.

VII. Experimental Methodology

This section outlines a proposed experimental plan with all items marked [confirm] pending finalization, as no experiments have yet been executed. The evaluation begins with Experiment 0 as a motivation check, training plain conditional diffusion on the task from [6] to measure low-band consistency, flat-region spurious details, and radial PSD distance, where a lack of these problems would shift the contribution toward data and computational efficiency. The dataset selection includes two remote-sensing SR datasets from [6], a natural-image control such as DIV2K, and subsampled paired training sets at 100%, 25%, 10%, and 5% for low-data analysis, with scale factors, patch sizes, and splits still to be confirmed. Performance will be compared against multiple baselines, including CDGAN [6], an adapted DAG-style GAN [5], plain conditional diffusion on full images and residuals, and a standard regression SR network. Evaluation metrics encompass PSNR, SSIM, LPIPS, FID (used cautiously for small test sets), low-band consistency, radial PSD distance, flat-region spurious-detail energy, and run-to-run seed variance, alongside efficiency tracking for training time, inference latency, peak memory, FLOPs, parameter counts, and bucket step counts across unspecified hardware. Finally, specialized experiments will evaluate domain-shift robustness using matched versus mismatched sensor spectra, performance scaling across training fractions, augmentation impacts with and without conditioning tags, flat-region detail thresholds, and sampling efficiency comparing adaptive budgets against uniform allocations while verifying tile-boundary seam errors.

VIII. Expected Results and Discussion

No measured results are available. The following are hypotheses to be tested, not findings. H1 Low-band consistency improves through the residual anchor and projection.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

Table 3: Experimental Evaluation Plan (no results yet)

Experiment	Comparison	Measures
Exp. 0	Plain diffusion vs. [6] task	Low-band, flat-region, PSD
Low-data	100/25/10/5% paired data	PSNR, SSIM, LPIPS
Domain shift	Matched vs. mismatched spectrum	PSD distance, fidelity
Flat regions	Full DSN-RD vs. no m guard	Spurious-detail energy
Sampling	Adaptive vs. uniform, equal mean steps	Quality, latency, seams
Stability	Seeds; no discriminator vs. GAN baselines	Run-to-run variance

H2 *Flat-region hallucination* is reduced because small m bounds residual amplitude there.

H3 *Domain-shift behavior* improves when the noise prior matches the measured domain spectrum.

H4 *Data efficiency* improves, with graceful degradation as the paired fraction shrinks, through the unpaired spectrum and tagged simulated pairs.

H5 *Average sampling cost* falls, because flat tiles need fewer DDIM steps than dense tiles.

H6 *Training is stable* without an adversarial discriminator, given the plain noise-prediction loss.

Each hypothesis has a pre-set decision rule. If plain residual diffusion matches Module 1, the normalizer claim is dropped. If adaptive budgets do not beat uniform sampling at equal average steps, Module 3 is dropped. If the tag does not reduce augmentation-induced shift, it is reported as unsupported. If Exp. 0 shows no low-frequency problem in plain diffusion, the motivation is reframed.

IX. Ablation Study

Each configuration adds one element to the previous one:

A0 *Plain conditional DDPM*: tests diffusion alone.

A1 + *Residual anchor*: tests the known residual device.

A2 + *Module 1 (spectrum + detail map)*: tests domain-statistics normalization.

A3 + *Module 2 (tag, incl. simulated pairs)*: tests leakage handling.

A4 + *Module 3 (adaptive sampling)*: tests compute allocation at equal average steps.

A5 *Full DSN-RD*: tests all components together.

Additional studies: γ *sensitivity* (strength of spectral shaping); *oracle detail map* (an upper bound on the effect of m_φ errors, using s instead of the prediction); *no projection* (contribution of consistency); *tag on/off* with a spectrum-altering augmentation; *number of tile buckets B*; and *paired vs. unpaired spectrum estimation*. No ablation results are available.

X. Complexity, Limitations, and Discussion

Complexity (qualitative). *Parameters*: the U-Net plus the additional detail-map predictor m_φ . *Training*: stage 0 is a one-time estimate, stage 1 adds a separate short training, stage 2 costs about that of a conditional diffusion model with an extra conditioning channel. *Inference*: the cost scales with the number of U-Net evaluations per tile. With bucket fractions π_b, the average step count is $\bar{N} = \sum_b \pi_b N_b$, which is the quantity Module 3 aims to reduce relative to a uniform N; overlapping tiles add overhead and tile processing introduces blending cost. Module 3 targets the average, not the worst case. Numerical costs are to be measured.

Limitations.

- *Normalization equivalence*: Eqs. (3) and (4) show that Module 1 is standard DDPM on a conditionally normalized target. Any benefit is empirical, not theoretical.
- *Single point of failure*: m_φ serves as prior, guard, and allocator, so errors in it degrade all three. Clamping, freezing, and the oracle ablation mitigate but do not remove this.
- *Degradation operator*: anchoring and projection assume a known or well-estimated D; misregistered real data weakens both.
- *Tile seams*: different step counts across neighbors can produce boundary artifacts; overlap and blending are used and seam error must be measured.
- *Sensor/domain mismatch*: a spectrum from a different sensor is a wrong prior.
- *Mispredicted detail*: a dense tile judged flat receives too few steps; conservative thresholds and a floor N_1 are proposed.
- *Remaining latency*: diffusion still costs more than a single-pass GAN.
- *Unproven tag*: no guarantee against augmentation leakage is claimed.
- *Overlap with existing work*: residual and null-space-projected diffusion are known outside the reviewed corpus, and related ideas likely exist; the novelty, if any, lies in the choice and interaction of parts. Papers [4, 6, 9, 10] were read at abstract level only, and only ten papers were reviewed.

A broader literature search is required before any formal novelty claim, covering residual or colored-noise diffusion for SR, remote-sensing diffusion SR, augmentation-conditioned diffusion, null-space-projected diffusion SR, and spatially adaptive step allocation.

XI. Conclusion

Diffusion-based SR in frequency-shifted domains faces scarce paired data, domain-shifted statistics, augmentation-induced bias, adversarial instability, and sampling cost, and the reviewed literature does not combine a diffusion generator with domainstatistics normalization and statistic-driven compute allocation. DSN-RD estimates a residual spectrum and a detail-scale map once and uses them as the noise prior, a flat-region guard, and a tile-level step allocator, together with a shared tag for augmentation and simulated pairs. Expected benefits are hypotheses only: better low-band consistency, less flat-region hallucination, better data efficiency, lower average sampling cost, and stable training without a discriminator. Their validation through the pre-specified experiments and ablations is the next step, and no superiority is claimed before it.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

REFERENCES

[1] F.-A. Croitoru, V. Hondru, R. T. Ionescu, and M. Shah, "Diffusion models in vision: A survey," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 9, 2023.

[2] A. Creswell, T. White, V. Dumoulin, G. Arulkumaran, B. Sengupta, and A. A. Bharath, "Generative adversarial networks: An overview," *IEEE Signal Process. Mag.*, vol. 35, no. 1, pp. 53–65, 2018.

[3] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and improving the image quality of StyleGAN," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2020.

[4] C. Ebert and P. Louridas, "Generative AI for software practitioners," *IEEE Software*, vol. 40, no. 4, 2023.

[5] N.-T. Tran, V.-H. Tran, N.-B. Nguyen, T.-K. Nguyen, and N.-M. Cheung, "On data augmentation for GAN training," *IEEE Trans. Image Process.*, vol. 30, 2021.

[6] S. Lei, Z. Shi, and Z. Zou, "Coupled adversarial training for remote sensing image super-resolution," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 5, 2020.

[7] J. Li, K. A. Skinner, R. M. Eustice, and M. Johnson-Roberson, "WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images," *IEEE Robot. Autom. Lett.*, vol. 3, no. 1, 2018.

[8] H. Cao, C. Tan, Z. Gao, Y. Xu, G. Chen, P.-A. Heng, and S. Z. Li, "A survey on generative diffusion models," *IEEE Trans. Knowl. Data Eng.*, 2024.

[9] W. Meng, X. Xu, X. Wang, T. Sun, W. Xia, J. Han, and H. Zhang, "Physical-layer authentication based on hierarchical variational autoencoder for industrial Internet of Things," *IEEE Internet Things J.*, vol. 10, no. 3, 2023.

[10] R. Zhang, Z. Xiong, H. Du, D. Niyato, J. Kang, X. Shen, and H. V. Poor, "Generative AI-enabled vehicular networks: Fundamentals, framework, and case study," *IEEE Network*, vol. 38, no. 4, 2024.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)