Comparing GAN, Diffusion, and Diffusion-GAN for Single-Image Deraining of UAV Imagery

Salsabilah Aulia Rahman[1*], Laksmita Rahadianti[2]

[1-2]*Faculty of Computer Science, Universitas Indonesia, Depok, Indonesia*

*corr-author: salsabilah.aulia@ui.ac.id

Abstract - Single-image deraining for Unmanned Aerial Vehicle (UAV) imagery remains challenging due to non-uniform rain patterns, motion blur, and real-time processing requirements. Existing generative paradigms, including Generative Adversarial Networks (GAN), Diffusion, and Diffusion–GAN, each face inherent trade-offs among restoration quality, stability, and efficiency. To address the lack of unified and fair benchmarking across these paradigms, this study presents a systematic and controlled comparative evaluation of three representative models, including TBGAN, WeatherDiff, and SupResDiffGAN, to assess their relative performance in UAV deraining tasks. The models are evaluated on the UAV-Rain1K and Rain100L datasets using PSNR, SSIM, and inference efficiency metrics to support informed selection of paradigms for UAV applications. Experimental results show that WeatherDiff achieves the highest fidelity with 19.99 dB PSNR, 0.8375 SSIM on UAV-Rain1K and 29.51 dB PSNR, 0.9093 SSIM on Rain100L. TBGAN yields sharper details but lower structural consistency, whereas SupResDiffGAN offers balanced performance with 19.03 dB PSNR and 0.7053 SSIM on UAV-Rain1K and 28.51 dB PSNR and 0.8681 SSIM on Rain100L, with faster inference. These findings highlight the practical trade-offs among the three paradigms and demonstrate that diffusion–GAN frameworks provide the most practical solution for UAV deraining, combining diffusion stability with adversarial sharpness for real-time restoration.

Keywords: Single-image deraining; UAV imagery; generative adversarial networks; diffusion models; diffusion–GAN.

I. INTRODUCTION

Image restoration in adverse weather conditions remains a fundamental challenge in computer vision, especially for real-time outdoor applications such as aerial surveillance, autonomous navigation, and traffic monitoring [1, 2]. Rain-induced degradation not only reduces image clarity but also interferes with downstream tasks, including object detection, tracking, mapping, and scene understanding [3]. When rain streaks or water droplets partially or completely obscure critical features, high-level perception systems may misidentify objects or lose essential contextual information for decision-making [4].

This restoration problem, commonly referred to as single-image deraining, aims to recover fine textures and enhance visual quality using only a single degraded input. The challenge becomes significantly more complex in images captured by Unmanned Aerial Vehicles (UAVs), where rapid motion, changing viewpoints, and inconsistent rainfall intensity lead to highly non-uniform rain patterns [5-7]. These factors contribute to motion blur, atmospheric scattering, and spatially varying degradation, making UAV-based deraining more difficult than its ground-level counterpart [4].

Earlier deraining approaches relied heavily on handcrafted priors such as sparsity, low-rank decomposition, and frequency-based filtering [8, 9]. Although effective under controlled conditions, these priors tend to generalize poorly to diverse or real-world rain patterns. The emergence of deep learning marked a substantial improvement, beginning with Convolutional Neural Network (CNN) based architectures that introduced multi-scale feature extraction and layered decomposition strategies for rain removal. Representative CNN models, such as SPDNet [6] and ECNet [1], demonstrate improved structural consistency and effective streak suppression; however, they frequently produce oversmoothed textures and remain sensitive to severe or highly irregular rain patterns. Recent comparative studies have further emphasized the importance of selecting appropriate deep learning architectures for achieving optimal performance in image processing tasks [10].

To address these limitations, this study groups recent generative restoration models into three representative paradigms for comparative analysis, namely Generative Adversarial Networks (GANs) [11], Diffusion Probabilistic Models (DPMs) [12], and Diffusion-GAN architectures [13]. This categorization is used in this work to reflect the dominant design principles found in recent deraining and image restoration literature. GAN-

based methods, such as SARain-GAN [7] and TBGAN [2], produce sharp and realistic textures through adversarial learning but suffer from instability and artifacts [14]. Diffusion models, including DDPM-based restoration [15] and WeatherDiff [9], improve reconstruction fidelity and structural coherence but exhibit slow inference due to multi-step sampling [12, 16]. Diffusion-GAN approaches, such as SupResDiffGAN [17], aim to combine the stability of diffusion with the sharpness of GANs; however, their effectiveness for UAV-based deraining remains underexplored [17].

Motivated by these challenges, this study provides a unified comparative analysis across three representative generative paradigms, TBGAN (GAN) [2], WeatherDiff (Diffusion) [9], and SupResDiffGAN (Diffusion-GAN) [18] for the task of single-image deraining. The objective is to evaluate how each paradigm influences restoration quality, structural fidelity, perceptual sharpness, and computational efficiency under fair, controlled experimental settings. Through this comparison, the study highlights the trade-offs among realism, stability, and efficiency across generative methods. Finally, we provide practical insights into the strengths and limitations of each when applied to diverse and complex rainy conditions in UAV imagery.

II. METHOD

This section describes the methodological framework used to conduct a fair and controlled comparison across three generative paradigms for single-image deraining: TBGAN (GAN) [2], WeatherDiff (Diffusion) [9], and SupResDiffGAN (Diffusion-GAN) [18]. The study is structured into four main stages: dataset preprocessing and normalization, model training under unified hyperparameter settings, standardized inference, and quantitative and qualitative evaluation. This design ensures that performance differences arise from the generative mechanisms themselves rather than inconsistencies in data preparation or training configuration.

A. Research Design

The overall workflow follows a comparative experimental approach in which each generative model processes a rainy UAV image to produce a restored output. To ensure fairness, all models are trained and evaluated under identical configurations, including dataset splits, image resolution, batch size, optimizer type, and evaluation metrics. This unified setup ensures that performance differences arise solely from the generative mechanisms rather than inconsistencies in training or preprocessing. The comparative pipeline adopted in this study aligns with the standard evaluation practices established [2, 9, 18], supporting a standardized and model-agnostic restoration procedure. A general comparative pipeline is shown in Fig.1.

B. Datasets

Two widely used deraining datasets are employed to assess performance under both complex UAV-specific scenarios and simpler static conditions.

1) UAV-Rain1K: UAV-Rain1K [19] contains simulated UAV images with highly non-uniform degradations arising from motion blur, changing viewpoints, and fluctuating rainfall intensity. These characteristics make it well-suited to evaluating model robustness and generalization in dynamic, realistic environments.

2) Rain100L: Rain100L consists of images with uniform light rain streaks [1]. It serves as a baseline for assessing restoration quality under simple rainfall conditions free from motion-induced distortions.

C. Evaluation Metrics

Two standard full-reference image quality metrics are used to assess restoration fidelity.

1) Peak Signal-to-Noise Ratio (PSNR): PSNR [20] measures pixel-level reconstruction accuracy as the mean squared error between the restored image and the ground truth. Higher PSNR values indicate fewer reconstruction errors and better fidelity.

2) Structural Similarity Index (SSIM): SSIM [21] measures structural similarity across luminance, contrast, and texture. It is widely regarded as more aligned with human perception than PSNR, with values closer to 1.0 indicating better structural recovery.

D. Model Configuration and Training Strategy

This subsection outlines the preprocessing steps, data augmentation strategy, and training configuration applied uniformly across all generative models to ensure a fair and consistent comparison (Fig. 1).

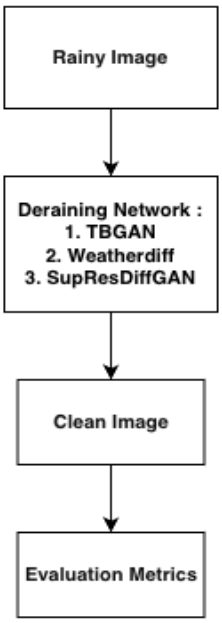

Fig. 1 General comparative pipeline for UAV image deraining across TBGAN, WeatherDiff, and SupResDiffGAN

1) Preprocessing and Augmentation: During training, 128×128-pixel patches are randomly extracted from full-resolution images using RandomResizedCrop and normalized to [-1, 1]. This patch-based training approach follows the standard methodology widely adopted in the image restoration literature [9], [17] and enables a fair comparison across generative paradigms under controlled computational constraints. The selected patch size balances GPU memory efficiency with sufficient spatial context for learning rain streak patterns, the framework supports alternative resolutions such as 256×256 at the cost of proportionally increased computational overhead. During inference, full-resolution test images are processed directly or, where applicable, via sliding-window patch reconstruction. Paired transformations are jointly applied to rainy–clean image pairs to ensure spatial consistency. Augmentation details are provided in Table I.

2) General Training Configuration: To ensure fairness, all models use identical optimization settings. The summary of the general configuration is presented in Table II.

3) Training Strategy and Convergence: Although the three paradigms differ in optimization behavior, the total training duration is kept comparable to maintain fairness. TBGAN [2], WeatherDiff [9], and SupResDiffGAN [18] are trained under matched hyperparameters, with differences limited to the update frequency dictated by each paradigm's learning dynamics. A summary of the strategy is presented in Table III.

To maintain consistent optimization across models, each paradigm adopts a loss formulation aligned with its generative mechanism. For TBGAN, the generator minimizes a combination of reconstruction and adversarial losses. The reconstruction term is computed using the L_1 distance between the clean reference image x_clean and the generated output $\hat{x}_{clean}$. The adversarial term L_{adv} promotes perceptual sharpness, and the weighting factor controls its strength λ_{adv}. The complete objective is expressed in (1).

$$L_{GAN} = L_1(x_{clean}, \hat{x}_{clean}) + \lambda_{adv}.L_{adv} \quad (1)$$

WeatherDiff follows the standard DDPM noise-prediction objective shown in (2). By minimizing the difference between the true and predicted noise at each time step, the model learns a stable, consistent denoising trajectory.

$$L_{diff} = \mathbb{E}_{(x,t,\varepsilon)}[\,||\,\varepsilon - \varepsilon_\theta(x_t, t, y)||_{2^2}] \quad (2)$$

SupResDiffGAN integrates diffusion modeling and adversarial refinement through a weighted objective, as shown in (3). This formulation balances diffusion reconstruction, adversarial feedback, and pixel-level consistency. The coefficients λ_1, λ_2, and λ_3 regulate the contribution of each term to ensure stable and convergent training.

$$L_{total} = \lambda_1 L_{diff} + \lambda_2 L_{GAN} + \lambda_3 L_{rec} \quad (3)$$

This unified formulation ensures that each paradigm converges within comparable training durations. GANs typically converge faster due to adversarial feedback, diffusion models require more update steps to model noise transitions, and Diffusion-GAN architectures benefit from the stability of diffusion while leveraging adversarial refinement for accelerated training and improved perceptual sharpness.

This configuration ensures that each paradigm operates under conditions that support its optimal performance, while maintaining fairness in computational effort and total training duration. Diffusion-based models inherently require more sampling steps, whereas adversarial models converge faster. Diffusion-GAN models, such as SupResDiffGAN, leverage both advantages, making them well-suited for balanced quality-efficiency evaluation.

III. RESULT AND DISCUSSION

This section presents the experimental results of TBGAN, WeatherDiff, and SupResDiffGAN on UAV-Rain1K and Rain100L. Both quantitative metrics and qualitative visual inspection are used to assess restoration fidelity and analyze the trade-offs among generative paradigms.

A. Quantitative Results and Analysis

Table IV summarizes the PSNR and SSIM scores obtained by the three models across both datasets. These metrics reflect pixel-level fidelity and structural similarity between the restored images and the corresponding ground-truth clean images.

As shown in Table IV, WeatherDiff consistently achieves the highest numerical scores across both datasets. On UAV-Rain1K, a challenging and dynamic environment, WeatherDiff achieves a PSNR of 19.99 dB and an SSIM of 0.8375, outperforming both TBGAN and SupResDiffGAN. This superior performance can be attributed to the iterative noise-modeling mechanism of diffusion models, which gradually refines image structure and enables strong preservation of global contextual information. The multi-step denoising process reduces structural artifacts, explaining the consistently higher SSIM values observed in diffusion-based restoration.

Although SupResDiffGAN achieves competitive PSNR, its lower SSIM indicates minor inconsistencies in fine-texture reconstruction. This behavior aligns with the architectural design of diffusion-GAN models: while the latent diffusion pathway stabilizes the global structure, the adversarial decoding branch tends to emphasize high-frequency details, which can occasionally lead to slight deviations in texture alignment relative to the ground truth. Nevertheless, SupResDiffGAN demonstrates a favorable balance between diffusion-level fidelity and GAN-level sharpness. TBGAN demonstrates adequate PSNR performance but yields lower SSIM, indicating weaker structural preservation. This outcome is consistent with the characteristics of adversarial training, where GAN generators prioritize perceptual sharpness and texture realism rather than strict pixel-level accuracy. As a result, GAN-based models often produce visually appealing outputs but may introduce subtle distortions that reduce structural similarity metrics. To further analyze computational efficiency across generative paradigms, the inference-time comparison of all evaluated models on both datasets is summarized in Table V.

Table V presents the inference efficiency of each model. The results follow the expected paradigm-based computational hierarchy: TBGAN (GAN) achieves the fastest inference, SupResDiffGAN offers intermediate speed, and WeatherDiff requires the longest processing time due to multi-step sampling. The significantly slower inference on UAV-Rain1K (1.6×–6.8× slower than Rain100L) is attributed to the larger native image resolutions, which require sliding-window patch reconstruction. While WeatherDiff achieves superior restoration quality, its computational cost (1530 ms per image on UAV-Rain1K) limits real-time applicability. SupResDiffGAN provides a practical trade-off, balancing restoration fidelity with moderate inference speed (700 ms), making it suitable for near-real-time UAV scenarios where both quality and efficiency are critical.

Overall, the quantitative results validate fundamental architectural differences across generative paradigms: GANs favor perceptual sharpness with high computational efficiency, diffusion models excel in structural coherence at the cost of inference speed, and diffusion–GAN architectures strike an intermediate balance between detail preservation, global stability, and computational feasibility for UAV deployment.

B. Qualitative Results and Analysis

Qualitative comparisons of deraining performance on UAV-Rain1K and Rain100L are presented in Fig. 2 and Fig. 3, respectively. The visual inspection focuses on texture recovery, structural coherence, and residual artifacts.

TABLE I
AUGMENTATION PARAMETER

Augmentation	Parameter
RandomResizedCrop	size = 128, scale = 0.8–1.0
RandomHorizontalFlip	p = 0.5
ColorJitter	brightness = 0.1, contrast = 0.1, saturation = 0.05

TABLE II
HYPERPARAMETERS

Parameter	Value
Learning Rate	1×10^{-4}
Optimizer	Adam
Batch Size	8

TABLE III
TRAINING STRATEGIES ACROSS MODELS

Aspect	TBGAN	WeatherDiff	SupResDiffGAN
Epoch / Update Count	200 epochs	225,000 updates (≈ 2,500 epochs)	200 epochs
Reason for Difference	Fast convergence via adversarial learning	Requires many steps to model noise transitions	Accelerated training compared to pure diffusion
Optimizer	Adam	Adam	Adam
Main Loss Function	L_1 + adversarial loss	L_{noise} (DDPM Loss)	Diffusion Loss + GAN Loss + Reconstruction Loss

TABLE IV
QUANTITATIVE RESULTS

Dataset	Model	PSNR (↑)	SSIM (↑)
UAV-Rain1K	TBGAN	18.88	0.7661
	WeatherDiff	**19.99**	**0.8375**
	SupResDiffGAN	19.03	0.7053
Rain100L	TBGAN	27.67	0.8885
	WeatherDiff	**29.51**	**0.9093**
	SupResDiffGAN	28.51	0.8681

TABLE V
INFERENCE TIME COMPARISON

Model	UAV-Rain1K (ms)	Rain100L (ms)
TBGAN	450.68	66.45
SupResDiffGAN	700.73	264.43
WeatherDiff	1530.49	983.08

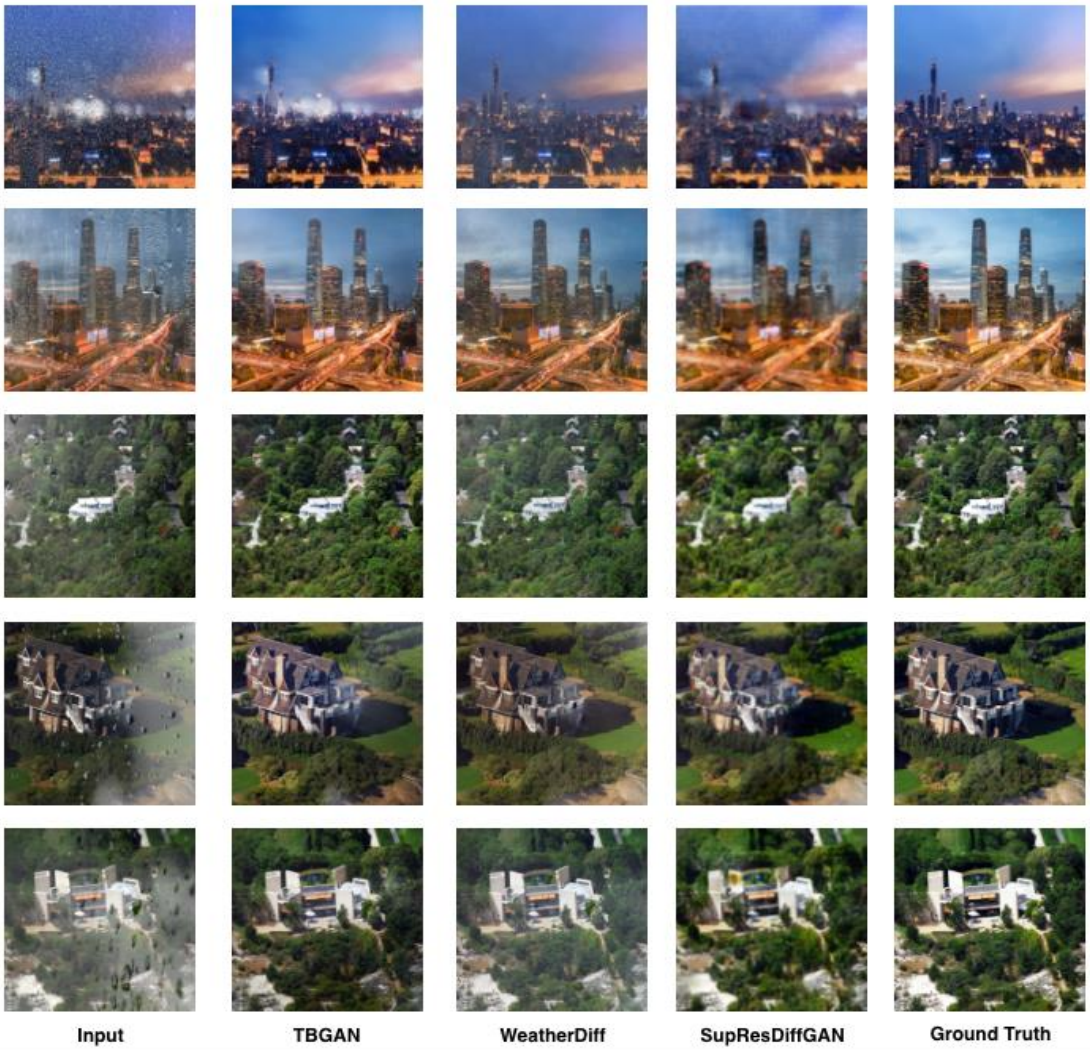

Fig. 2 Qualitative comparison of deraining results on UAV-Rain1K.
From left to right: Input, TBGAN, WeatherDiff, SupResDiffGAN, and Ground Truth.

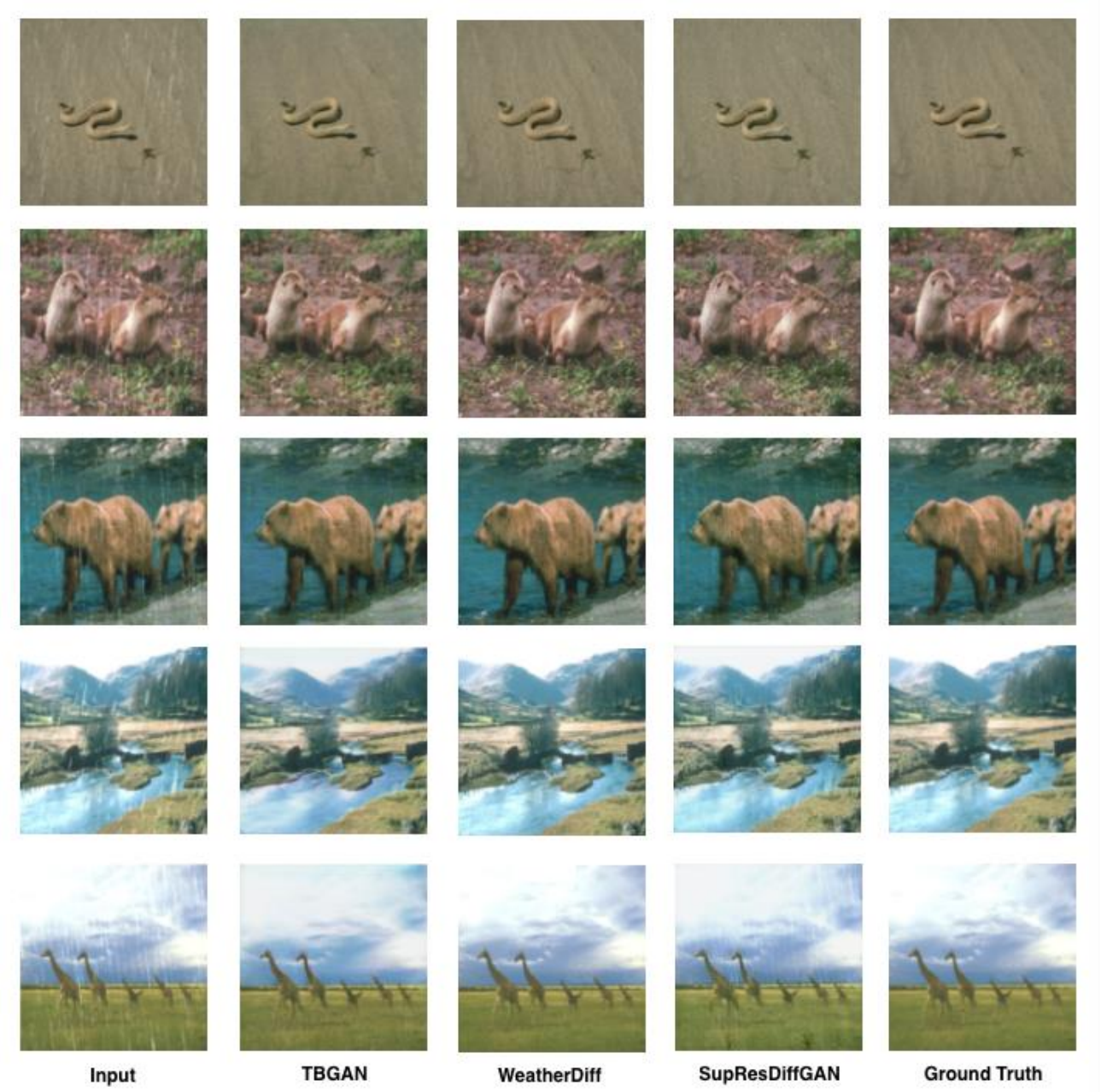

Fig. 3 Qualitative comparison of deraining results on Rain100L
From left to right: Input, TBGAN, WeatherDiff, SupResDiffGAN, and Ground Truth.

The qualitative observations align with the quantitative findings.

1) TBGAN: The outputs exhibit strong perceptual sharpness due to adversarial optimization, which encourages the generator to amplify high-frequency details favored by the discriminator. While this helps suppress rain streaks, it also leads to over-enhancement and occasional artifacts, particularly in textured regions such as foliage. This behavior is characteristic of GAN-based restoration, where visual realism is prioritized over strict structural accuracy.

2) WeatherDiff: The diffusion model produces smooth and globally coherent reconstructions with minimal artifacts. Its multi-step denoising process preserves structural information, which explains the highest SSIM scores across datasets. However, the conservative refinement inherent to diffusion can cause

slight over-smoothing in severely degraded regions, resulting in minor loss of fine-grained textures.

3) SupResDiffGAN: The diffusion-GAN framework delivers the most perceptually balanced results. Diffusion-based stability preserves global structure, while adversarial decoding enhances high-frequency details. This combination reduces artifacts commonly seen in GAN outputs while producing sharper textures than pure diffusion models. Slight softening may still occur due to VAE latent-space compression, which inherently smooths fine structural patterns.

C. Navigating Generative Trade-offs

This subsection discusses the inherent strengths and limitations of the three generative paradigms in the context of UAV-focused single-image deraining.

1) Quality vs. Stability (GAN vs. Diffusion): TBGAN produces the sharpest textures due to its adversarial learning objective, which encourages the generator to emphasize high-frequency components that the discriminator considers realistic. However, this mechanism inherently prioritizes perceptual sharpness rather than pixel-wise accuracy, leading to occasional structural distortions and artifact amplification. This explains the lower SSIM scores obtained by TBGAN, particularly in UAV-Rain1K, where spatial degradation is highly non-uniform. WeatherDiff, in contrast, benefits from the iterative refinement characteristic of diffusion models. By progressively denoising the input through multiple noise-estimation steps, diffusion architectures naturally reinforce global coherence and reduce artifact formation. This behavior is reflected in the consistently higher SSIM scores observed across both datasets. The slight over-smoothing observed in certain regions is consistent with the conservative nature of diffusion, which tends to suppress fine textures under severe degradation.

2) Quality vs. Efficiency (Diffusion vs. Diffusion-GAN): Although WeatherDiff achieves the highest numerical fidelity, its multi-step sampling process results in slow inference (1530 ms per image on UAV-Rain1K), limiting its applicability for real-time UAV scenarios. SupResDiffGAN mitigates this limitation by performing diffusion in the VAE's low-dimensional latent space, thereby significantly reducing the number of denoising operations required. The adversarial decoding branch then recovers sharper textures, enabling the model to achieve a favorable balance between structural fidelity and perceptual detail while maintaining moderate inference speed (700 ms). This explains why SupResDiffGAN maintains competitive PSNR and SSIM values while achieving 2.2× faster inference than diffusion-based methods. Meanwhile, TBGAN achieves the fastest processing (450 ms) at the cost of lower structural consistency, reflecting the computational efficiency of single-pass adversarial generation. The significantly slower inference on UAV-Rain1K compared with Rain100L (1.6×–6.8× across models) stems from the larger native image resolutions, which require sliding-window patch reconstruction.

3) Advantages of Diffusion-GAN Frameworks: SupResDiffGAN demonstrates that integrating adversarial refinement with diffusion priors can resolve key limitations of both paradigms. Diffusion provides stability and global consistency, preventing the artifacts commonly observed in GAN-based outputs. Meanwhile, adversarial decoding compensates for the intrinsic over-smoothing of diffusion, enabling the recovery of fine details. The moderate softening observed in some outputs is consistent with VAE latent compression, which tends to smooth high-frequency structures.

In summary, the comparative findings show that diffusion-based models offer the highest structural fidelity, GAN-based models excel in perceptual sharpness, and diffusion-GAN architectures provide the most practical trade-off between restoration quality and computational efficiency. For UAV-based deraining tasks, where both clarity and responsiveness are crucial, hybrid generative frameworks are the most balanced and effective approach.

IV. CONCLUSION

This study conducted a comparative evaluation of three generative paradigms, TBGAN, WeatherDiff, and SupResDiffGAN for single-image deraining across the UAV-Rain1K and Rain100L datasets. The results show that the diffusion-based WeatherDiff achieves the highest restoration fidelity on both datasets, scoring 19.99 dB PSNR and 0.8375 SSIM on UAV-Rain1K, and 29.51 dB PSNR and 0.9093 SSIM on Rain100L, confirming the strong structural stability and denoising capability of diffusion models despite their slower inference time. TBGAN produces the sharpest textures and fast inference but tends to introduce structural distortions and perceptual artifacts in complex UAV scenarios. SupResDiffGAN consistently provides a balanced trade-off, combining diffusion-driven stability with adversarial refinement to achieve competitive accuracy and substantially faster inference than pure diffusion models. Overall, the findings indicate that no single paradigm universally excels; GANs favor perceptual sharpness, diffusion models offer superior

stability, and diffusion-GAN architectures provide the most practical compromise between quality and efficiency. For real-time UAV applications requiring both clarity and responsiveness, diffusion-GAN frameworks such as SupResDiffGAN represent a promising direction. Several limitations should be acknowledged. This study does not include statistical significance testing or robustness analysis across multiple experimental runs, which limits the assessment of performance stability and reproducibility. Additionally, although test images are evaluated at their native resolution, the patch-based training methodology (128×128 crops) and evaluation on only two benchmark datasets may not fully reflect the model's scalability to extremely high-resolution UAV imagery (e.g., >1024×1024) or to diverse real-world deployment conditions. Future work may explore improved sampling strategies, lightweight hybrid architectures, and multi-weather generalization to enable deployment in real-world aerial imaging systems. Further research incorporating statistical validation, cross-dataset evaluation, higher-resolution training strategies, and real-world UAV field testing would strengthen the generalizability and practical applicability of these findings.

REFERENCES

[1] Y. Li, Y. Monno, and M. Okutomi, "Single Image Deraining Network with Rain Embedding Consistency and Layered LSTM," in *Proc. IEEE/CVF Winter Conf. on Applications of Computer Vision (WACV)*, 2022, pp. 3957–3966. doi: 10.1109/WACV51458.2022.00401.

[2] L. Zhao, J. Long, and T. Zhong, "A Deep Learning-Based Two-Branch Generative Adversarial Network for Image De-Raining," *Sensors*, vol. 24, no. 20, p. 6724, 2024, doi: 10.3390/s24206724.

[3] P. Li, J. Jin, G. Jin, L. Fan, X. Gao, T. Song, and X. Chen, "Deep Scale-space Mining Network for Single Image Deraining," in *Proc. IEEE/CVF Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2022, pp. 4275–4284. doi: 10.1109/CVPRW56347.2022.00473.

[4] Y. Tomida, T. Katayama, T. Song, and T. Shimamoto, "Efficient Deraining Model Using Transformer and Kernel Basis Attention for UAVs," in *Proc. 2024 Int. Tech. Conf. on Circuits/Systems, Computers, and Communications (ITC-CSCC)*, 2024, pp. 1–5. doi: 10.1109/ITC-CSCC62988.2024.10628256.

[5] W. Li, C. Li, H. Jiang, Y. Wang, S. Wu, and Z. Wu, "High-Resolution Aerial Image Restoration with Latent Diffusion Models," in *Proc. IEEE Int. Conf. on Unmanned Systems (ICUS)*, 2024, pp. 1874–1878. doi: 10.1109/ICUS61736.2024.10840082.

[6] Q. Yi, J. Li, Q. Dai, F. Fang, G. Zhang, and T. Zeng, "Structure-Preserving Deraining with Residue Channel Prior Guidance," presented at the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 4218–4227.

[7] M. H. Kolekar, S. Bose, and A. Pai, "SARain-GAN: Spatial Attention Residual UNet Based Conditional Generative Adversarial Network for Rain Streak Removal," *IEEE Access*, vol. 12, 2024, doi: 10.1109/ACCESS.2024.3375909.

[8] Y. Wang, Z. Ma, and H. Su, "Research on Hole Feature Matching Based on Image Restoration," in *Proc. Int. Conf. on Computers, Information Processing and Advanced Education (CIPAE)*, 2022, pp. 385–389. doi: 10.1109/CIPAE55637.2022.00087.

[9] O. Özdenizci and R. Legenstein, "Restoring Vision in Adverse Weather Conditions with Patch-Based Denoising Diffusion Models," vol. 45, 2023, doi: 10.1109/TPAMI.2023.3238179.

[10] A. F. Akbar, P. D. W. Ayu, and D. P. Hostiadi, "Performance Analysis of Deep Learning Architectures in Classifying Fake and Real Images," *JUITA J. Inform. 132 167–176*, vol. 13, no. 2, 2025, doi: https://doi.org/10.30595/juita.v13i2.25790.

[11] I. J. Goodfellow, "Generative Adversarial Networks," *Adv. Neural Inf. Process. Syst.*, vol. 3, 2014, doi: 10.1145/3422622.

[12] J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," *Adv Neural Inf Process Sys*, vol. 33, pp. 6840–6851, 2020, doi: https://doi.org/10.48550/arXiv.2006.11239.

[13] Z. Wang, H. Zheng, P. He, W. Chen, and M. Zhou, "Diffusion-GAN: Training GANs with Diffusion," *ArXiv Prepr. ArXiv220602262*, 2022, doi: https://doi.org/10.48550/arXiv.2206.02262.

[14] P. Salehi, A. Chalechale, and M. Taghizadeh, "Generative Adversarial Networks (GANs): An Overview of Theoretical Model, Evaluation Metrics, and Recent Developments," *ArXiv Prepr. ArXiv200513178*, 2020, doi: 10.48550/arXiv.2005.13178.

[15] H. Cho, H.-K. Shin, Y. Jang, S.-J. Ko, and S.-W. Jung, "PD-CR: Patch-Based Diffusion Using Constrained Refinement for Image Restoration," *IEEE Signal Process. Lett.*, vol. 31, 2024, doi: 10.1109/LSP.2024.3381908.

[16] X. Lin, J. He, Z. Chen, Z. Lyu, B. Dai, F. Yu, W. Ouyang, Y. Qiao, and C. Dong, "DiffBIR: Towards Blind Image Restoration with Generative Diffusion Prior," *ArXiv Prepr. ArXiv230815070*, 2024, doi: https://doi.org/10.48550/arXiv.2308.15070.

[17] L. T. Trinh and T. Hamagami, "Latent Denoising Diffusion GAN: Faster Sampling, Higher Image Quality," *IEEE Access*, vol. 12, pp. 78161–78172, 2024, doi: 10.1109/ACCESS.2024.3406535.

[18] D. Kopeć, W. Kozłowski, M. Wizerkaniuk, D. Krutul, J. Kocoń, and M. Zięba, "SupResDiffGAN: A New Approach for the Super-Resolution Task," *ArXiv Prepr. ArXiv250413622*, 2025, doi: https://doi.org/10.48550/arXiv.2504.13622.

[19] W. Chang, H. Chen, X. He, X. Chen, and L. Shen, "UAV-Rain1k: A Benchmark for Raindrop Removal from UAV Aerial Imagery," *ArXiv Prepr. ArXiv240205773*, 2024, doi: https://doi.org/10.48550/arXiv.2402.05773.

[20] W. Simoes and M. D. Sá, "PSNR and SSIM: Evaluation of the Imperceptibility Quality of Images Transmitted over Wireless Networks," *Procedia Comput. Sci.*, vol. 251, pp. 463–470, 2024, doi: 10.1016/j.procs.2024.11.134.

[21] J. Nilsson and T. Akenine-Möller, "Understanding SSIM." 2020. doi: https://doi.org/10.48550/arXiv.2006.13846.