

A Comprehensive Survey of Deep Learning Super-Sampling and Neural Rendering Techniques for Real-Time AI-Powered Image Upscaling

Goldi Soni

Assistant Professor

Amity University Chhattisgarh

gsoni@rpr.amity.edu

Anukriti Soni

B.Tech (CSE)

Amity University Chhattisgarh

anukriti.soni@s.amity.edu

Sarwang Sinha

B.Tech(CSE)

Amity University

Chhattisgarh

sarwang.sinha@s.amity.edu

Abstract— The high level of realism required in modern video games and interactive applications goes beyond the possibilities provided by traditional rasterization pipelines in terms of computational power necessary for achieving both photorealism and decent frame rates. This survey is aimed at providing an overview of more than 30 relevant papers on deep learning approaches such as DLSS, neural super-resolution, temporal anti-aliasing, generative adversarial networks upsampling and real-time denoising. In total, the reviewed studies can be organized into six main themes, namely, convolutional network-based image upscaling, GAN-based perceptual upscaling, temporal feedback and anti-aliasing, real-time image denoising, upscaling in specific domains and virtual reality, and Transformer networks-based image restoration. An analysis of the reviewed literature proves that the evolutionary path leads from purely spatial solutions for image scaling based on SRCNN to the advanced temporal models like DLSS 4 and SwinIR with attention mechanisms. In this regard, it is concluded that hybrid solutions involving denoising and upscaling should be considered among the most interesting areas of future research together with transformer models and optimized hardware implementation.

Keywords—Super-resolution, Deep learning, DLSS, Temporal anti-aliasing, Generative adversarial network, Convolutional neural network, Real-time rendering, Neural denoising, Transformer, Image upscaling.

I. INTRODUCTION

The need for high-resolution high-quality graphics in video games and other real-time applications has increased drastically over the last decade. In modern games, there is a requirement to provide a realistic environment along with dynamic lighting, ray-tracing, reflections, physically-based materials and maintain frame rates of 60 FPS or above. Traditional techniques used in rendering using rasterization pipeline require substantial computing power which makes it impossible to use native resolutions such as 4K or 8K on consumer devices economically and thermally feasible.

NVIDIA has developed Deep Learning Super-Sampling technique which has changed this paradigm completely, by rendering the scene at a lower resolution and then reconstructing the high-resolution image using a neural network [1]. Since its inception in 2018, DLSS has been upgraded in four generations, including developments in the areas of temporal stability, attention mechanisms and frame synthesis [10]. Even after such advancement, the field continues to be split into various different areas and the research work continues in parallel in these areas.

In addition, the practical considerations associated with deployment on console hardware, mobile GPUs, and VR headset applications call for efficiencies not covered by most academic benchmarks [23], [29].

The aim of this survey is, first and foremost, to bring together these disparate strands of literature in one organized, comparative form. Through a review of more than 30 peer-reviewed publications, this paper provides an overview of the technical lineage between early CNN models through to current multi-frame and Transformer-based upscaling work. The paper will also explore cross-disciplinary insights from remote sensing, VR, and real-time graphics production pipelines.

The goals of this paper are threefold: (1) the creation of a taxonomy around super-resolution and upscaling topics that can be clustered into distinct categories of analysis; (2) the examination of the development of DLSS

algorithms, from game-specific spatial upscaling through to multi-frame Transformer-based neural architectures; and (3) the discovery of future opportunities in efficiency, generalization, and cross-platform hardware support.

The key contributions of this survey are as follows:

- A structured thematic taxonomy of 30+ research works spanning CNN-based, GAN-based, temporal, denoising, VR, and Transformer-based super-resolution methods.
- A systematic comparative analysis of key architectural and algorithmic choices, including loss functions, temporal feedback strategies, and dimensionality of input signals.
- A detailed tracing of the DLSS evolutionary lineage from DLSS 1.0 through DLSS 4, contextualizing each iteration within the broader academic landscape.
- Identification of convergence trends toward unified neural rendering pipelines and hardware-aware efficiency optimization for mobile and console platforms.

II. RELATED WORKS

Research in AI-driven image upscaling has progressed through various stages, with each stage being defined by a unique architectural model. This section categorizes the reviewed papers into six groups and highlights the major contributions, challenges, and interrelationships among the contributions in each group.

A. CNN-based Super-Resolution Fundamentals

The first fundamental contribution of deep learning for super-resolution came from the work of Dong et al. [2]. In their seminal work, Dong et al. introduced SRCNN, a three-layer neural network whose goal was to perform low-resolution to high-resolution conversion in an end-to-end manner. By eliminating the need for the multi-step approach involving manual pipeline design and mean square error training, the authors were able to demonstrate that CNN-based super-resolution is superior to all other approaches such as bicubic interpolation and sparse coding.

One significant limitation in the practical implementation of SRCNN was the high computational requirement since it was a high-resolution operation where the input and output were high resolution images. Shi et al. [20] proposed the use of an ESPCN in which the entire process was done in low resolution while using a pixel shuffle technique to arrange channels to form pixels at the output level. This helped in lowering the computational requirements by r^2 when the scale factor is r and became the standard output module for all future real-time upscaling networks. Lim et al. [19] found that eliminating batch normalization in residual networks (EDSR) helped save 40% memory usage, thereby making them deeper with better results.

B. GAN-Based Perceptual Upscaling

Despite achieving outstanding PSNR metrics through CNN approaches, the images yielded exhibited extremely smooth textures due to optimization in the pixel domain. Ledig et al. [18] resolved this issue by presenting SRGAN, which utilized perceptual loss calculated in the features space using the pre-trained VGG model and adversarial training. The discriminator compelled the generator to generate indistinguishable visual textures from reality despite being mathematically different. Studies performed on human raters always preferred images produced by SRGAN; therefore, it was established that perceptual quality is quite different from signal fidelity.

The progress was made further with Wang et al. [3], who introduced ESRGAN, which used RRDBs in place of simple residual blocks and adopted the relativistic discriminator. Since the adversarial loss was computed relative to the realism of the image, the hallucination problems associated with SRGAN vanished, and naturally textured images were obtained. ESRGAN emerged victorious in the PIRM2018-SR Challenge. Zhang et al. [21] built on this idea by introducing Residual Dense Network (RDN). It employed dense skip connections that retained all the intermediate layers, thereby making sure that all the information present in early layers, including fine details, was utilized in the final reconstruction process. Wang et al. [25] made use of this concept to degrade real-world images using Real-ESRGAN.

C. Temporal Anti-Aliasing and Upscaling

The move from spatial to temporal upscaling is undoubtedly the most significant change in real-time rendering technology. Modern TAA was first implemented by Karis [7] in Unreal Engine 4, where jitter offsets were applied to sample sub-pixel coordinates on a frame-by-frame basis, while neighborhood filtering was applied in the color domain to eliminate bad history samples. This pioneering effort showed that temporal integration could achieve 8x supersampling accuracy within less than 1ms per frame, demonstrating that time can be traded for space.

A survey conducted by Yang et al. [4] classified the temporal anti-aliasing approaches according to how they accumulated history and validated it and noted that ghosting and blur were the two main failure types. The analysis revealed that technologies behind TAA, such as the use of motion vectors and history rejection, directly informed DLSS 2.0 and AMD FSR 2.0 development. Marrs et al. [8] came up with Adaptive TAA (ATAA), which combined TAA with some sparse ray-traced samples at failure points; this particular idea was an early indication of neural rendering-based approaches to reconstruction. Epic Games [9] introduced Temporal Super Resolution (TSR) to their Unreal Engine 5; the method consisted of carefully designed shaders capable of upscaling to near-4K resolution from the 1080p feed in real time for all GPU brands without any neural hardware involved.

AMD FidelityFX Super Resolution [11] made a logical step forward towards better accessibility by combining spatial upscaling (FSR 1.0 based on contrast adaptive sharpening) and a full temporal feedback loop (FSR 2.0 relying on reactive masks and depth-luminance consistency). Through the substitution of the neural network used by DLSS with an analytical shader, FSR 2.0 managed to provide equal image quality without using any machine learning hardware, indicating that all key components of temporal scaling can be expressed mathematically.

D. Real-Time Denoising Frameworks

Real-time denoising deals with generating a clean image using Monte Carlo ray tracing sampling 1 sample per pixel. Schied et al. [15] set out the groundwork for real-time denoising using Spatiotemporal Variance-Guided Filtering (SVGF). SVGF utilizes the concept of applying spatial filtering only in areas where variance is low while distinguishing direct lighting and indirect lighting to process each independently. It is now an established fact that filtering should be performed proportional to the signal variance in almost any game ray tracing pipeline.

The breakthrough was made by Vogels et al. [16], who introduced Kernel Prediction Convolutional Networks (KPCN). Instead of predicting final colors, KPCN predicted individual filter kernels for each pixel. This approach makes it impossible to generate colors not present in the input; moreover, the use of a biased loss function favors the presence of dark spots compared to light ones. Meng et al. [14] used the bilateral grid data structure to perform geometry-aware denoising, accounting for depth and normal discontinuity. The technique of combining denoising and super-resolution proposed by Thomas et al. [12] showed 30% faster computation, thanks to being able to discriminate between high-frequency noise and texture in the raw, noisy input.

Firmino et al. [27] considered co-design of renderer and AI denoiser through post-denosing error estimation in a pre-pass followed by assigning rays only to those areas which will not be able to be successfully processed by the AI denoiser. Balint et al. [28] proposed the use of neural partitioning pyramids that routed images to different experts based on a gate-based architecture and obtained better performance at lower cost compared to single-stage models. Research team [29] enabled real-time denoising on mobile GPUs with geometry-based filter prediction, enabling 60 fps processing.

E. Domain-Specific, VR, and Platform-Targeted Upscaling

The paper by Watson [1] compared DLSS against native rendering for ray-tracing games and showed that DLSS can achieve up to 360% better performance than native rendering while the quality of DLSS 2.0 renders was often comparable to or better than native rendering due to built-in anti-aliasing. Researchers at the University of Minnesota Morris [6] offered an easy-to-understand analysis of the differences between DLSS 1.0 and 2.0

architectures with emphasis on the use of a generalized temporal autoencoder instead of the per-game spatial CNN as the primary motivation for adoption.

Kaplanyan et al. [13] aimed to address the problem of high-resolution demands in VR and temporal inconsistencies that cause physiological strain. The authors developed a foveated neural upsampler that used the limited computational resources available to process high resolution only in the image center and apply stereo-consistency loss. This allowed for creating immersive VR experiences using half-resolution rendering on the same mobile platform. Research in remote sensing [24] showed that domain-specific models were essential when applying SR to multispectral satellite images due to the similarity of pansharpening and multi-buffer fusion in DLSS.

[30] In the multi-branch joint denoising and upsampling network proposed by AMD, a noisy low-resolution lighting branch and a clear high-resolution texture branch are allowed to exchange information through a feature interaction bridge, thus leveraging the albedo information to assist the process of reconstructing the lighting. [5] On a parallel track, the idea of separating lighting from materials was independently realized in Neural Super-Resolution with Radiance Demodulation.

F. Transformer-Based and Survey-Level Contributions

According to Liang et al. [26], the self-attention operation of Swin Transformers using shifted window attention for global propagation without the quadratic complexity surpassed all CNN approaches in performance on the set benchmarks for super-resolution. With its capability to correlate far-off pixels, SwinIR can do adaptive filtering beyond the abilities of CNNs due to fixed-size convolutional kernels. For example, NVIDIA [10] applied that approach to the new generation of DLSS 4 which uses not the CNN autoencoder but the Vision Transformer Block with Multi-Frame Generation generating four AI frames for each frame rendered.

The meta-survey of Wang et al. [22] from 2014-2020 considered different approaches to deep learning SR according to their network architecture, losses, and learning strategies. It is worth mentioning that their conclusion regarding PSNR saturation and unresolved issues with perceptual quality correctly anticipated the rise of GAN and Transformer approaches. The NTIRE 2025 Efficient SR Challenge [23] also focused on inference speed with the winning techniques based on NAS and quantization achieving 2018 performance with 1% of computations.

III. SURVEY METHODOLOGY AND TAXONOMIC FRAMEWORK

This section describes the systematic methodology used for paper selection, categorization, and comparative analysis. The survey follows a structured literature review protocol designed to ensure comprehensive coverage, thematic coherence, and reproducible classification.

A. Paper Selection Criteria

The papers chosen were found in well-known academic journals such as IEEE Xplore, ACM Digital Library, arXiv, as well as technical reports in the industry domain. The criteria for selecting the literature included: (1) addressing issues related to super-resolution, temporal upsampling, denoising, or neural rendering in the realm of real-time applications; (2) having a reproducible approach with quantified results; and (3) being published between 2014 and 2025. Technical industry reports from NVIDIA, AMD, and Epic Games were chosen for consideration because of their recognized contributions and benchmarks.

B. Taxonomic Classification

The chosen papers have been divided into six thematic groups according to the specific type of technical contribution: (T1) spatial super-resolution based on CNNs, (T2) perceptual upscaling based on GANs, (T3) temporal anti-aliasing and upscaling, (T4) neural denoising in real time, (T5) upscaling domain-specific and platform-oriented, and (T6) Transformer-based and survey contributions.

C. Evaluation Dimensions

The criteria used to evaluate each paper include the following evaluation categories: primary architectural paradigm, temporal perception (spatial only or spatial and temporal feedback), loss function category (pixel-based, perceptual, or adversarial), hardware platform (high-end graphics processor, console, or mobile

devices), and innovation. The aforementioned multi-category evaluation helps to construct comparative tables shown in Section IV.

Algorithmically, the survey methodology can be described as follows:

Step 1: Obtain papers that meet the selection criteria from all possible sources.

Step 2: Categorize each paper by assigning it a primary theme cluster (T1–T6).

Step 3: Determine evaluation categories: architecture, temporal perception, loss function category, hardware platform.

Step 4: Create comparative tables within and across clusters.

Step 5: Detect convergence points, problems, and future work directions.

IV. RESULTS AND COMPARATIVE ANALYSIS

This section presents a comparative analysis of the surveyed methods across classification metrics and evolutionary markers. Table 1 summarizes the primary characteristics of representative works from each thematic cluster. Table 2 presents a comparative performance overview for key methods where quantitative benchmarks are available.

Table 1: Thematic classification of surveyed papers showing primary cluster, architectural type, temporal awareness, and loss function.

Paper	Title (Abbreviated)	Cluster	Architecture	Temporal	Loss Type
[2] Dong 2014	SRCNN	T1 - CNN SR	3-Layer CNN	No	MSE
[17] Dong 2015	SRCNN Extended Study	T1 - CNN SR	Deep CNN	No	MSE
[20] Shi 2016	ESPCN (Sub-Pixel CNN)	T1 - CNN SR	Pixel Shuffle	No	MSE
[19] Lim 2017	EDSR	T1 - CNN SR	Deep ResNet	No	L1
[18] Ledig 2017	SRGAN	T2 - GAN SR	ResNet + GAN	No	Perceptual+GAN
[3] Wang 2018	ESRGAN (RRDB)	T2 - GAN SR	RRDB + Rel-GAN	No	Perceptual+GAN
[21] Zhang 2018	RDN	T2 - GAN SR	Dense ResNet	No	L1
[25] Wang 2021	Real-ESRGAN	T2 - GAN SR	RRDB + GAN	No	Perceptual+GAN
[7] Karis 2014	TAA (Unreal Engine 4)	T3 - Temporal	Analytic Shader	Yes	Empirical
[4] Yang 2020	TAA Survey	T3 - Temporal	Survey	Yes	N/A
[11] AMD 2021	FSR 1.0 / FSR 2.0	T3 - Temporal	Analytic + Shader	FSR2: Yes	Analytical
[9] Epic 2022	TSR (Unreal Engine 5)	T3 - Temporal	43-tap Shader	Yes	Analytical

[15] Schied 2017	SVGF	T4 - Denoising	Variance Filter	Yes	L2
[16] Vogels 2018	KPCN (Kernel Prediction)	T4 - Denoising	CNN + Kernel	No	Asymmetric L2
[12] Thomas 2022	Joint Denoise+SR	T4 - Denoising	Multi-Head CNN	Yes	Temporal L2
[28] Balint 2023	Neural Partitioning Pyramids	T4 - Denoising	MoE Network	Yes	Semantic L2
[1] Watson 2020	DLSS Benchmarking	T5 - Domain	CNN AutoEncoder	Yes	MSE
[13] Kaplanyan 2020	Neural SR for VR	T5 - Domain	Foveated CNN	Yes	Stereo L2
[5] Anon 2024	Radiance Demodulation SR	T5 - Domain	Decomposition Net	Yes	L2
[30] AMD 2025	Joint Denoise+Upscale (AMD)	T5 - Domain	Multi-Branch Net	Yes	Perceptual
[26] Liang 2021	SwinIR	T6 - Transformer	Swin Transformer	No	L1
[10] NVIDIA 2025	DLSS 4 (ViT + MFG)	T6 - Transformer	Vision Transformer	Yes	Perceptual+GAN
[22] Wang 2020	DL for SR: A Survey	T6 - Survey	Survey	N/A	N/A
[23] NTIRE 2025	Efficient SR Challenge	T6 - Survey	NAS/Quantized	No	L1

Table 2: Comparative performance of key methods across PSNR improvement, temporal stability, hardware requirement, and scalability.

Method	Year	PSNR Gain	Temporal Stability	HW Requirement	Scale Factor
SRCNN [2]	2014	+1.5 dB (Set5)	N/A	Any GPU	2x, 3x, 4x
ESPCN [20]	2016	+0.9 dB vs SRCNN	N/A	Any GPU	2x–4x
EDSR [19]	2017	+0.5 dB vs SRCNN	N/A	High-End GPU	2x–4x
ESRGAN [3]	2018	Lower PSNR, Higher MOS	N/A	High-End GPU	4x
SVGF [15]	2017	1spp denoising	Moderate	RTX GPU	Denoising

DLSS 2.0 [1]	2020	+360% FPS	High	RTX Tensor Core	1.5x–4x
FSR 2.0 [11]	2022	Near-DLSS Quality	Moderate-High	Any GPU	1.5x–4x
SwinIR [26]	2021	SOTA on all benchmarks	N/A	High-End GPU	2x–4x
TSR [9]	2022	~50% GPU time saved	High	Any GPU	1.5x–4x
DLSS 4 [10]	2025	4:1 Frame Generation	Very High	RTX 40+ Series	Multi-Frame

The results presented in Table 1 and Table 2 confirm several important evolutionary patterns. Across the T1 and T2 clusters, there is a consistent trend from MSE-based pixel loss toward perceptual and adversarial objectives, driven by the observation that mathematical fidelity does not correlate with perceived image quality [18], [22]. The introduction of the pixel shuffle output layer [20] and the removal of Batch Normalization [19] collectively became universal design principles.

Within the temporal cluster (T3), the progression from Karis’ TAA [7] to DLSS 4 [10] illustrates how the same core insight of accumulating information across time has been progressively refined: from hand-tuned shaders, through learned convolutional autoencoders, to full Vision Transformers with multi-frame generation. Vendor-agnostic solutions [9], [11] demonstrate that hardware-specific acceleration is not strictly required, enabling democratization across platforms.

The denoising cluster (T4) consistently shows that joint treatment of related tasks (denoising + upscaling, or renderer-aware sampling + denoising) outperforms sequential separate pipelines. Thomas et al. [12] quantified a 30% efficiency gain for the unified approach. Balint et al. [28] demonstrate that semantic segmentation of image content enables further specialization, pointing toward mixture-of-experts architectures as a scalable path for neural rendering.

Transformer-based methods [10], [26] achieve the highest quality benchmarks within their respective tasks, confirming the superiority of content-adaptive self-attention over fixed convolutional filters. The NTIRE 2025 challenge [23] highlights that efficiency, measured in FLOPs, memory bandwidth, and mobile NPU latency, is now the binding constraint rather than raw model capacity.

V. CONCLUSION

This survey has presented a comprehensive review of over 30 research papers spanning foundational CNN super-resolution, GAN-based perceptual upscaling, temporal anti-aliasing and feedback, real-time neural denoising, domain-specific and VR-targeted solutions, and Transformer-based restoration. A systematic thematic taxonomy was constructed across six clusters, and comparative analysis tables were presented to enable quantitative assessment of architectural choices and performance characteristics.

The central finding of this survey is that real-time AI-powered image upscaling has progressed from isolated spatial super-resolution networks to deeply integrated neural rendering pipelines that operate across both spatial and temporal dimensions, incorporating scene geometry, motion vectors, and multi-frame generation. The convergence of denoising and upscaling into unified architectures, as demonstrated by Thomas et al. [12] and AMD [30], represents the current research frontier.

Certain limitations of this survey should be acknowledged. The comparative performance data for several industry technical reports (DLSS, FSR, TSR) is drawn from vendor-provided benchmarks rather than independently reproducible experiments. Additionally, survey coverage of emerging neuromorphic and hardware co-design approaches is necessarily limited by the availability of published work as of early 2025.

Future research should prioritize three directions: (1) development of hardware-agnostic Transformer upscalers that match the efficiency of shader-based methods on mobile and console platforms; (2) integration of fairness-aware and artifact-diagnostic modules to ensure consistent output quality across diverse scene types, including underrepresented content categories; and (3) standardized open benchmarks specifically designed for real-time temporal upscaling evaluation, moving beyond static image datasets such as Set5 and BSD100. As GPU architectures increasingly dedicate silicon area to AI acceleration, the boundary between traditional rendering and neural synthesis will continue to dissolve, making a unified theoretical framework for this emerging discipline an important long-term research goal.

. REFERENCES

- [1] A. Watson, "Deep Learning Techniques for Super-Resolution in Video Games," arXiv preprint, 2020.
- [2] C. Dong, C. C. Loy, K. He, and X. Tang, "Learning a Deep Convolutional Network for Image Super-Resolution," ECCV 2014.
- [3] X. Wang et al., "ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks," ECCVW 2018.
- [4] L. Yang, S. Liu, and M. Salvi, "A Survey of Temporal Antialiasing Techniques," Computer Graphics Forum, vol. 39, no. 2, pp. 607–621, 2020.
- [5] Anonymous Research Team, "Neural Super-Resolution for Real-time Rendering with Radiance Demodulation," Research Publication, 2024.
- [6] University of Minnesota Morris, "Deep-Learning Realtime Upsampling Techniques in Video Games," Scholarly Horizons, vol. 10, iss. 2, 2023.
- [7] B. Karis, "High Quality Temporal Supersampling," SIGGRAPH 2014 Advances in Real-Time Rendering.
- [8] A. Marrs, J. Spjut, H. Gruen, R. Sathe, and M. McGuire, "Adaptive Temporal Antialiasing," HPG 2018.
- [9] Epic Games, "Temporal Super Resolution in Unreal Engine 5," Technical Documentation, 2022.
- [10] NVIDIA Corporation, "DLSS 4.0: Multi Frame Generation and Transformer-Based Upscaling," Technical Whitepaper, 2025.
- [11] AMD, "FidelityFX Super Resolution Technology," GPUOpen Technical Documentation, 2021–2022.
- [12] M. M. Thomas et al., "Temporally Stable Real-Time Joint Neural Denoising and Supersampling," HPG 2022.
- [13] A. Kaplanyan et al., "Neural Supersampling for Real-time Rendering," ACM Transactions on Graphics, 2020.
- [14] X. Meng, Q. Zheng, A. Varshney, G. Singh, and M. Zwicker, "Real-Time Monte Carlo Denoising with Neural Bilateral Grid," SIGGRAPH Asia 2020.
- [15] C. Schied, A. Kaplanyan, C. Wyman et al., "Spatiotemporal Variance-Guided Filtering," HPG 2017.
- [16] T. Vogels, F. Rousselle, B. McWilliams et al., "Denoising with Kernel Prediction and Asymmetric Loss Functions," ACM TOG, vol. 37, no. 4, 2018.
- [17] C. Dong, C. C. Loy, K. He, and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," IEEE TPAMI, 2015.
- [18] C. Ledig et al., "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," CVPR 2017.
- [19] B. Lim, S. Son, H. Kim, S. Nah, and K. M. Lee, "Enhanced Deep Residual Networks for Single Image Super-Resolution," CVPRW 2017.
- [20] W. Shi, J. Caballero, F. Huszar et al., "Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel CNN," CVPR 2016.
- [21] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual Dense Network for Image Super-Resolution," CVPR 2018.

- [22] Z. Wang, J. Chen, and S. C. H. Hoi, "Deep Learning for Image Super-Resolution: A Survey," IEEE TPAMI, 2020.
- [23] NTIRE Workshop, "The Tenth NTIRE Efficient Super-Resolution Challenge Report," 2025.
- [24] Various Authors, "Advancing Image Super-Resolution Techniques in Remote Sensing: A Comprehensive Survey," arXiv, 2025.
- [25] X. Wang, L. Xie, C. Dong, and Y. Shan, "Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data," ICCVW 2021.
- [26] J. Liang et al., "SwinIR: Image Restoration Using Swin Transformer," ICCVW 2021.
- [27] A. Firmino et al., "Denoising-Aware Adaptive Sampling for Monte Carlo Ray Tracing," SIGGRAPH 2023.
- [28] M. Balint, K. Wolski, K. Myszkowski, H. P. Seidel, and R. Mantiuk, "Neural Partitioning Pyramids for Denoising Monte Carlo Renderings," SIGGRAPH 2023.
- [29] Research Team, "Edge-Aware Denoising Framework for Real-Time Mobile Ray Tracing," Graphics Journals, 2025.
- [30] P. Kazmierczyk, S. Kim, W. Uss et al., "Joint Denoising and Upscaling via Multi-branch and Multi-scale Feature Network," I3D 2025.