WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal Homepage: www.ijrpr.com ISSN 2582-7421

More Than an Upgrade: A Survey on the Divergent Paradigms of DLSS 4 and DLSS 5

Sparsh Nimbark [a], Jay Narwani [b], and Dr. Mansi Shanishwara [c] (Formerly Published as Mansi Bosamia)

[a] Department of Computer Science and Engineering, Gyanmanjari Innovative University, Survey No.30, Sidsar Road, near Iscon Eleven, Bhavnagar, Gujarat 364060, India

[b] Department of Computer Science and Engineering, Gyanmanjari Innovative University, Survey No.30, Sidsar Road, near Iscon Eleven, Bhavnagar, Gujarat 364060, India

[c] Department of Computer Science and Engineering, Gyanmanjari Innovative University, Survey No.30, Sidsar Road, near Iscon Eleven, Bhavnagar, Gujarat 364060, India

DOI : https://doi.org/10.55248/gengpi.07.1026.34030

A B S T R A C T

Real-time rendering has moved from brute-force pixel shading toward learned reconstruction, and NVIDIA's Deep Learning Super Sampling (DLSS) family is its most widely deployed example. This survey argues that DLSS 4 and DLSS 5, though they share a name, embody divergent paradigms. DLSS 4 is reconstructive and interpolative: vision-transformer Super Resolution and Ray Reconstruction approximate a higher-sample render of the same scene, and a split-network Multi Frame Generation model inserts up to three interpolated frames per rendered frame at about 1 ms each, paced by Blackwell's hardware flip metering. DLSS 5 is generative: a compact one-step pixel-space diffusion transformer, conditioned on the rendered frame, motion vectors, a carried temporal state and artistic-direction values, synthesizes a new final appearance using priors learned from real-world imagery. We trace the lineage from temporal antialiasing through DLSS 1–4.5, dissect both architectures, and compare them on fidelity, determinism, latency, compute, integration and authorship. We review competing technologies (AMD FSR 4/4.1, Intel XeSS 2, Unreal Engine TSR, Apple MetalFX), consolidate published benchmarks with a simple frame-time model that recovers DLSS 5's ≈8 ms per-frame cost at 4K from independent measurements, and survey gaming and non-gaming applications, including digital twins, medical imaging and flood-map super-resolution. We conclude that DLSS 5 extends rather than supersedes DLSS 4, and that its open challenges are latency under frame generation, verifiable faithfulness to authored content, and governance of artistic intent.

Keywords: Deep learning super sampling, diffusion models, frame generation, generative AI, neural rendering, one-step distillation, real-time graphics, super-resolution, temporal antialiasing, vision transformers.

1. Introduction

1.1 Motivation: Rendering at the End of Brute Force

REAL-TIME rendering has always been an exercise in approximation. Every frame is a numerical estimate of the rendering equation [1], an integral over all incident light at every visible surface point, and every generation of graphics hardware has been asked to evaluate that integral more faithfully, at more pixels, more often. Display resolutions have climbed to 3840 × 2160 and beyond, refresh rates now reach 240–480 Hz, and physically based ray and path tracing multiply the cost of every pixel. Meanwhile, the transistor-density gains that historically paid for this growth have slowed. NVIDIA frames its Blackwell generation explicitly as an attempt to keep raising visual quality after the end of Moore's Law, with neural rendering as the proposed answer [2].

Deep Learning Super Sampling (DLSS) is the most widely deployed embodiment of that answer. Since 2019 it has steadily shifted real-time graphics from shading every displayed pixel toward inferring most of them. With DLSS 3, NVIDIA stated that the combination of super resolution and frame generation allows a GPU to reconstruct seven-eighths of displayed pixels with AI [3]; with DLSS 4 Multi Frame Generation (MFG) and a 2× per-axis upscale, the arithmetic falls to one shaded pixel in sixteen (Fig. 1).

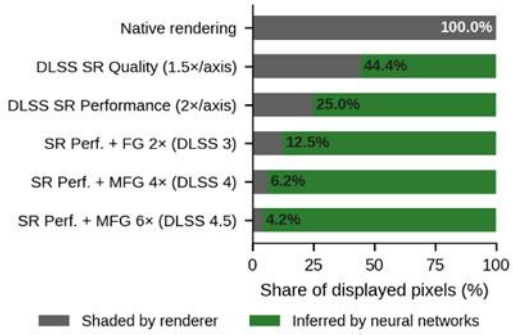


Fig. 1. Share of displayed pixels actually shaded by the renderer under successive DLSS configurations, computed as $\rho = 1/(r^2k)$ from the per-axis upscale factor r and the frame multiplier k (see (4)). The remainder is inferred by neural networks. DLSS 3 corresponds to NVIDIA's "seven-eighths" statement [3].

DLSS 5, launched on September 1, 2026, changes the question being asked [4]. Every earlier DLSS stage answers: what would the renderer have produced with more samples or more frames? DLSS 5 instead answers: what would this scene look like in the real world? NVIDIA's technical report describes it as the first DLSS technology that generates the final displayed appearance rather than reconstructing a higher-cost reference output [5]. It notes that earlier reconstruction methods remain bounded by what the renderer itself can represent and simulate. The title of this survey, *More Than an Upgrade*, refers to precisely this change of target distribution (Fig. 2).

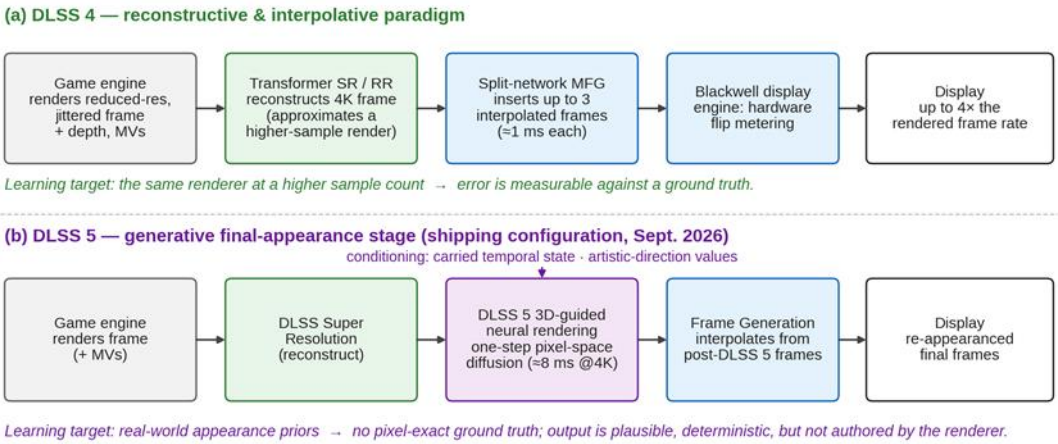


Fig. 2. The two paradigms at a glance. (a) DLSS 4 reconstructs and interpolates: every network is trained to approximate an output the same renderer could produce with a larger budget. (b) DLSS 5 inserts a generative stage that replaces the final appearance using learned real-world priors; in the shipping configuration it sits after DLSS Super Resolution and before Frame Generation, so the two paradigms compose rather than compete [5], [11].

1.2 Scope and Survey Methodology

This survey covers the DLSS family from its 2019 debut to the October 2026 DLSS 5 model update, together with the classical techniques DLSS builds on and the competing vendor technologies it is measured against. Primary sources were prioritized: NVIDIA Applied Deep Learning Research (ADLR) technical reports for DLSS 4 and DLSS 5 [5], [6], peer-reviewed literature on temporal antialiasing (TAA), aggregate G-buffer antialiasing and neural rendering [7]–[9], and engine documentation [10]. Quantitative performance claims were cross-checked against independent outlets (TechPowerUp, Digital Foundry, Tom's Hardware, TechSpot, Gamers Nexus) wherever possible. Vendor-only figures are explicitly labeled as such. Community reverse-engineering and forum discussion are reported where they add insight but are flagged as unverified.

1.3 Contributions

The survey makes the following contributions:

1) a unified taxonomy of real-time upscaling, reconstruction, interpolation, reprojection and generative techniques that separates them by *what they are trained to approximate* (Section III);

2) a verified timeline of the DLSS family from DLSS 1.0 to DLSS 5 (Section IV, Table I);

3) architectural deep-dives into DLSS 4 (transformer super resolution and ray reconstruction, split-network MFG, hardware flip metering, Reflex Frame Warp, Blackwell co-design) and DLSS 5 (3D-guided one-step pixel-space diffusion, conditioning, determinism and artistic control) (Sections V–VI);

4) a multi-axis comparison of the two paradigms and an analysis of how they compose in a single pipeline (Section VII, Table II);

5) a consolidated, source-checked benchmark analysis, including a simple frame-time model that explains DLSS 5's measured cost (Section IX);

6) a survey of gaming and non-gaming applications with a suitability matrix, and a discussion of open technical, ethical and artistic challenges (Sections X–XIII).

1.4 Organization

Section II reviews the mathematical and algorithmic foundations. Section III presents the taxonomy, and Section IV traces the evolution of DLSS. Sections V and VI dissect DLSS 4 and DLSS 5, and Section VII compares them. Section VIII reviews competing techniques, Section IX consolidates performance evidence, and Sections X and XI cover applications. Section XII discusses challenges, Section XIII outlines future directions, and Section XIV concludes. Appendix A lists nomenclature and Appendix B proposes a benchmarking checklist.

2. Background and Foundations

2.1 The Rendering Equation and the Cost of a Pixel

The outgoing radiance L_o at surface point x in direction ω_o is given by Kajiya's rendering equation [1]:

$$L_o(x,\omega_o) = L_e(x,\omega_o) + \int_\Omega f_r(x,\omega_i,\omega_o)\, L_i(x,\omega_i)\, (\omega_i \cdot n)\, d\omega_i \qquad (1)$$

where L_e is emitted radiance, f_r the bidirectional reflectance distribution function (BRDF), L_i incident radiance and n the surface normal. Rasterization approximates (1) with local shading models and a handful of light samples; path tracing estimates it stochastically with Monte Carlo sampling, producing noise that falls only as the inverse square root of the sample count. Both strategies therefore face the same trade-off: quality scales with samples, and samples scale with time. Every technique surveyed here is a way of obtaining more *effective* samples per unit of GPU time, either by reusing samples across space and time (reconstruction), synthesizing whole frames between rendered ones (interpolation), or, in DLSS 5, importing prior knowledge of appearance learned from data (generation).

2.2 Sampling, Aliasing and Multisampling

A rasterizer point-samples continuous signals (geometric coverage, shading, texture) at pixel centers, and any content above the Nyquist limit aliases into jagged edges, crawling silhouettes, shimmering foliage and specular flicker. Supersampling (SSAA) removes aliasing by brute force. Multisample antialiasing (MSAA) decouples visibility from shading, evaluating coverage at several sub-pixel positions but shading once per primitive per pixel. Yang *et al.* note that MSAA was the dominant technique until deferred shading broke the visibility-to-shading mapping, turning MSAA into costly supersampling and multiplying G-buffer storage and bandwidth [7].

2.3 Aggregate G-Buffer Antialiasing (AGAA)

Crassin *et al.* quantify the geometric-detail problem that later learned methods inherit [8]. A 1920 × 1080 G-buffer with 16 samples per pixel at a typical 20-byte layout exceeds 600 MB. AGAA instead uses the rasterizer to build a compact, prefiltered per-pixel representation: a small number c of statistical aggregates of depth, normals and material attributes, which are then shaded at a fixed rate independent of geometric complexity (Fig. 3). Normal-distribution width is tracked with Toksvig-style variance so that specular lobes are correctly broadened. With two aggregates per pixel, AGAA approaches the quality of 32× MSAA while using 54% less memory and running up to 2.6× faster; relative to 8× MSAA it saves 30% memory at 1.7× the speed [8]. AGAA matters to this survey because it makes the principle *shade less, reconstruct more* explicit, a principle DLSS later learns end-to-end.

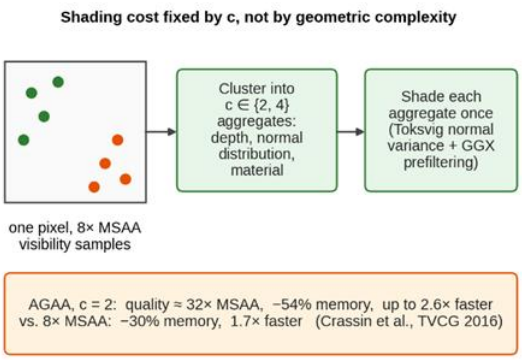


Fig. 3. Aggregate G-buffer antialiasing. Dense visibility samples are clustered into $c \in \{2, 4\}$ prefiltered aggregates per pixel, each shaded once, so shading cost is set by c rather than by geometric complexity. Figures reported by Crassin *et al.* [8].

2.4 Temporal Antialiasing (TAA)

TAA, described by Yang *et al.* as temporally amortized supersampling and the most widely used antialiasing technique in modern engines [7], spreads the cost of supersampling across frames. Each frame renders one sample per pixel at a sub-pixel jittered position (typically a Halton sequence). Using engine motion vectors v, the accumulated history is reprojected from the previous frame, validated, and blended with the new sample (Fig. 4):

$$C_t(p) = \alpha\, x_t(p) + (1 - \alpha)\, \mathrm{clip}(C_{t-1}(p - v_t(p))) \quad (2)$$

where x_t is the current sample, C_{t-1} the history buffer, α a (fixed or adaptive) blend weight, and clip(·) a history-rectification operator such as neighborhood clamping. Yang *et al.* decompose TAA into *sample accumulation* and *history validation* and observe that validation is fundamentally ill-posed, because the only reliable evidence available to judge stale history is the current, sparsely sampled frame [7]. Over-aggressive rejection causes shimmering; under-aggressive rejection causes ghosting and blur. This heuristic, hand-tuned validation step is exactly what DLSS 2 replaced with a learned network [12].

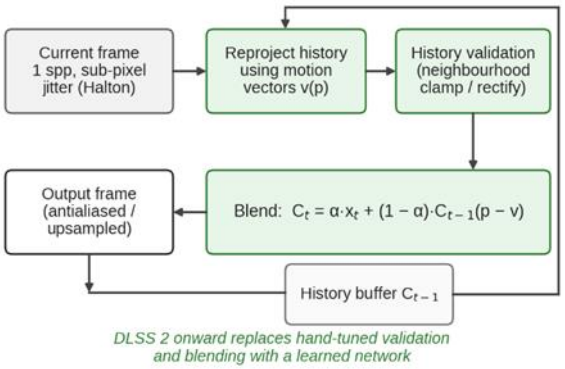


Fig. 4. Canonical TAA loop: jittered sampling, motion-vector reprojection, history validation and exponential blending per (2), after Yang *et al.* [7]. Learned reconstruction (DLSS 2 onward) replaces the validation and blending stages with a neural network.

2.5 Temporal Upsampling and Temporal Super Resolution

When the input sample density drops below one sample per output pixel, TAA becomes *temporal upsampling*: the history is accumulated at output resolution while each new frame contributes only a fraction of the needed samples. Yang *et al.* caution that this must not be confused with frame interpolation, which increases frame *rate* rather than spatial resolution [7]. The distinction maps directly onto DLSS Super Resolution (SR) versus DLSS Frame Generation (FG).

Epic Games' documentation for Unreal Engine Temporal Super Resolution (TSR) quantifies the trade-off [10]. In its example, a native-4K frame costs 57.50 ms versus 33.37 ms for a TSR frame rendered at 1080p, roughly a 2× speed-up. The documentation also gives the time needed for a newly disoccluded region to converge to one sample per pixel. With screen percentage s (internal-to-output resolution ratio per axis) and frame rate f:

$$t_{1\mathrm{spp}} = 1000 / (s^2 \cdot f) \ \ [\mathrm{ms}] \qquad (3)$$

At $s = 0.5$ and 60 Hz, (3) gives 66.7 ms, i.e., four frames (Fig. 5). Epic warns that temporal upscalers are therefore *frame-rate dependent*: a CPU-bound frame rate slows convergence, and low internal resolutions lengthen the period during which disocclusions look soft [10]. Equation (3) also explains why learned reconstruction is so valuable: a network that can infer plausible detail before history has converged hides exactly this transient.

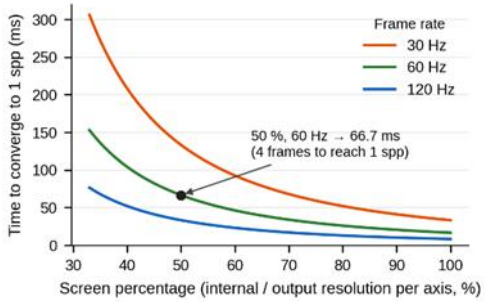


Fig. 5. Time for disoccluded regions to converge to one sample per pixel under temporal upsampling, from (3) [10]. Lower screen percentages and lower frame rates both lengthen convergence, which is why temporal upscalers degrade when frame rates are CPU-bound.

Combining spatial upscaling with frame multiplication, the fraction of displayed pixels that are actually shaded is:

$$\rho = 1 / (r^2 \cdot k) \qquad (4)$$

where r is the per-axis upscale factor and k the number of displayed frames per rendered frame. DLSS 4 in Performance mode ($r = 2$) with 4× MFG ($k = 4$) yields $\rho = 1/16$ (Fig. 1).

2.6 Neural Rendering

Neural rendering spans learned scene representations, learned denoisers and learned image formation. Neural radiance fields (NeRF) [13] entangle geometry, materials and lighting inside a single volumetric function, which makes relighting difficult. Knodt *et al.*'s *Neural Ray-Tracing* [9] restores the decomposability of light transport by combining a signed-distance-function surface model with a disentangled neural BRDF inside a physically based ray-tracing pipeline. This enables both relighting and view synthesis, and the paper simulates multi-bounce reflection at $O(n \log n)$ rather than $O(n^2)$ cost. The work illustrates a spectrum that frames the rest of this survey. At one end, neural components sit *inside* a physically based pipeline and remain interpretable (Neural Ray-Tracing, DLSS Ray Reconstruction). At the other end, a network supplies appearance knowledge the renderer never simulated (DLSS 5).

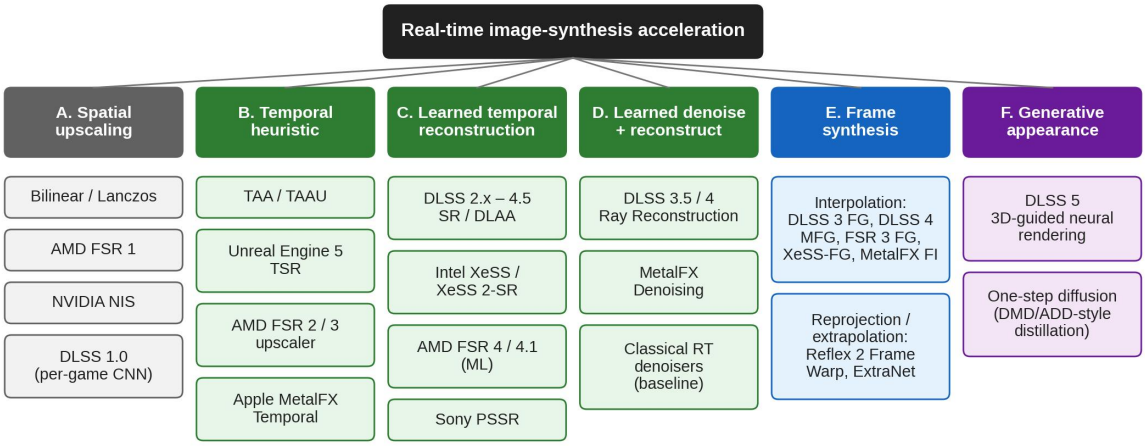


Fig. 6. Taxonomy of real-time image-synthesis acceleration, organized by target distribution. Branches B–E approximate what the renderer would produce with more samples or more frames; branch F (DLSS 5) synthesizes appearance the renderer never simulated.

2.7 Transformers, Diffusion Models and One-Step Distillation

Self-attention [14] lets every token weigh evidence from every other token, and vision transformers (ViT) [14] and windowed variants such as Swin [15] adapt the idea to images. For temporal reconstruction, attention is attractive because the most useful history sample for a pixel may lie far away in screen space after motion.

Diffusion models [16] generate images by iteratively denoising random noise. Their quality is unmatched, but they require many network evaluations per image, which is far beyond a 16.7 ms frame budget. Distillation compresses them. Distribution Matching Distillation (DMD) turns a diffusion model into a one-step generator and reaches an FID of 2.62 on ImageNet 64 × 64 [17]; DMD2 improves this to 1.28 [18]. Adversarial Diffusion Distillation (ADD) distills SDXL into a 1–4-step model using a discriminator built on pretrained vision features [19], and Latent ADD extends this to high-resolution synthesis [19]. Later work shows that diversity and control can survive aggressive distillation. Graphics-specific work such as RenderFlow recasts G-buffer-to-image rendering as single-step conditional flow matching [20]. DLSS 5 belongs to this lineage: a one-step diffusion transformer distilled from far larger foundation models [21].

3. A Taxonomy of Real-Time Image-Synthesis Acceleration

Existing classifications usually group upscalers by vendor or by whether they use machine learning. We argue that a more useful axis is the *target distribution*: what is the method trained or designed to approximate? Fig. 6 organizes the field into six branches. Spatial upscalers (A) interpolate a single frame. Temporal heuristic methods (B) and learned temporal reconstruction (C) accumulate jittered samples over time and target a higher-sample render of the same scene. Learned denoise-and-reconstruct methods (D) extend that target to noisy ray-traced signals. Frame-synthesis methods (E)

either interpolate between rendered frames or reproject the latest frame to a newer camera pose. Only the generative branch (F) targets a distribution the renderer never produces: real-world appearance.

This lens makes the central claim of the survey precise. Branches B–E differ in mechanism but share a ground truth, namely the renderer at a higher budget, so their errors (ghosting, smearing, disocclusion artifacts) can be measured against a reference. Branch F has no such reference, so its outputs can be *different* without being *wrong*. Most of the technical, evaluative and ethical consequences discussed in Sections VII and XII follow from this single distinction.

4. Evolution of DLSS

4.1 DLSS 1.0 (2019): Per-Game Spatial Upscaling

The first DLSS shipped in February 2019 in *Battlefield V* and *Metro Exodus*. It was a largely spatial convolutional upscaler trained separately for each game. Reception was poor: TechSpot found that the *Battlefield V* implementation looked and performed worse than a simple resolution upscale [22], and NVIDIA acknowledged that the network required a fixed GPU time per frame and that training focused on 4K [22].

4.2 DLSS 2.0 (2020): The Reconstructive Paradigm Begins

DLSS 2.0 replaced per-game training with a single generalized convolutional autoencoder that consumes the same inputs as TAA: jittered color, motion vectors and depth. Edward Liu's GTC 2020 presentation lists a 4× pixel upscaling ratio (e.g., 1080p to 4K), one model for all games, and a cost of about 1.5 ms at 4K on an RTX 2080 Ti [12]. Because the network learns the history-validation and accumulation steps of (2) rather than hand-coding them, any TAA-capable engine could integrate it [12]. DLSS 2.0 is therefore the true beginning of the reconstructive paradigm.

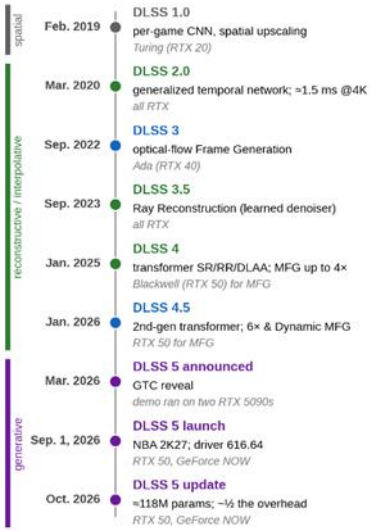


Fig. 7. DLSS timeline, 2019–2026, with GPU generations. Colors follow the taxonomy of Fig. 6: grey = spatial, green = reconstructive, blue = interpolative, purple = generative [3], [4], [12], [22]–[26].

4.3 DLSS 3 (2022): Frame Generation

Announced on September 20, 2022 for the Ada Lovelace generation, DLSS 3 added frame generation [3]. The Optical Flow Accelerator computes a dense flow field between two rendered frames, and a convolutional network combines the two frames, the flow field, engine motion vectors and depth to synthesize an intermediate frame. Combined with SR, NVIDIA claimed up to 4× higher frame rates than native rendering, with Reflex integrated to offset the added latency [3].

4.4 DLSS 3.5 (2023): Ray Reconstruction

DLSS 3.5 introduced Ray Reconstruction (RR), replacing the hand-tuned denoisers of ray-traced effects with a unified network trained on substantially more data than DLSS 3, so that denoising and upscaling happen jointly at output resolution [6].

4.5 DLSS 4 (2025) and DLSS 4.5 (2026)

DLSS 4, launched in January 2025, brought the first real-time transformer models for SR, RR and DLAA (available on all RTX GPUs) and Multi Frame Generation of up to 4× (exclusive to the RTX 50 series) [23], [27]. Seventy-five games and applications supported MFG at launch, rising to

more than 250 by January 2026 [29]. DLSS 4.5 (January 2026) added a second-generation SR transformer, a 6× MFG mode and *Dynamic* MFG, which adjusts the frame multiplier in real time [24].

4.6 DLSS 5 (2026): The Generative Paradigm

DLSS 5 was announced at GTC in March 2026 [25] and launched on September 1, 2026 in *NBA 2K27* with Game Ready Driver 616.64, running on all RTX 50 series GPUs and on GeForce NOW [4]. An October 2026 model update roughly halved its overhead, and *Hogwarts Legacy* support was scheduled for November 5, 2026 [26]. Fig. 7 and Table I summarize the lineage.

Table 1 – Evolution of the DLSS Family

Version	Date	Core model and new capability	Hardware	Paradigm
DLSS 1.0	Feb. 2019	Per-game CNN; spatial AI upscaling	RTX 20+	Spatial
DLSS 2.0	Mar. 2020	Generalized convolutional autoencoder; temporal SR, ≈1.5 ms at 4K (RTX 2080 Ti)	All RTX	Recon.
DLSS 3	Sep. 2022	CNN + hardware optical flow; Frame Generation (one frame)	FG: RTX 40+	Interp.
DLSS 3.5	Sep. 2023	Unified denoiser + SR; Ray Reconstruction	All RTX	Recon.
DLSS 4	Jan. 2025	Vision transformer SR/RR/DLAA; split-network MFG up to 4×; flip metering	ViT: all RTX; MFG: RTX 50	Recon. + interp.
DLSS 4.5	Jan. 2026	2nd-gen SR transformer; 6× and Dynamic MFG	MFG: RTX 50	Recon. + interp.
DLSS 5	Mar. 2026 (ann.); Sep. 1, 2026	One-step pixel-space diffusion transformer; 3D-guided neural rendering	RTX 50; GeForce NOW	Generative

Recon. = reconstructive; interp. = interpolative.

5. DLSS 4: The Reconstructive and Interpolative Paradigm

NVIDIA's technical overview frames DLSS 4 around three pillars: image quality, smooth frame rates and minimal input latency, with ease of integration as a fourth design constraint [6]. Fig. 8 shows the complete data flow.

5.1 Transformer-Based Ray Reconstruction

In a conventional pipeline, ray-traced signals are denoised at input resolution and then upscaled, so shading detail never exceeds the input resolution. Hand-tuned denoisers also tend to produce images that are plausible rather than faithful to ground truth, and they need per-title tuning [6]. RR merges denoising and upscaling into one network. NVIDIA reports that convolutional backbones reached their limits as training data grew, showing ghosting, painterly smearing and weak temporal coherence. The replacement is a custom transformer whose attention can aggregate individual path-traced samples even when they are far apart in space and time [6]. The model must generalize across very different sampling patterns: white noise, blue noise, quasi-Monte Carlo and spatiotemporally correlated ReSTIR samples. Ground-truth references used tens of thousands of samples per pixel [6]. This is the clearest statement of the reconstructive objective: the network is trained to reproduce what the *same* path tracer would produce at a vastly higher budget.

5.2 Transformer-Based Super Resolution and DLAA

Relative to the DLSS 3 CNN, the DLSS 4 SR transformer uses four times the computation and twice the parameters within a similar frame budget [6]. This was made possible by co-designing the network with custom CUDA kernels and running training and inference in FP8 on Ada and Blackwell tensor cores. NVIDIA's comparisons show the transformer in Performance mode (2× per axis) matching or exceeding the CNN in Quality mode (about 1.5× per axis) [6]. That is the practical reason DLSS 4 lets users drop a quality tier for the same perceived image, raising frame rate at no visual cost. DLAA applies the same network at native resolution purely for antialiasing.

5.3 Multi Frame Generation: The Split Network

DLSS 3 generated one 4K frame in about 3.25 ms on an RTX 4090. DLSS 4 generates each of up to three frames in about 1 ms on average on an RTX 5090 [6]. The key change is architectural. One portion of the network runs once per *pair* of rendered frames, and its features are reused; a much smaller portion runs once per *generated* frame [6]. For a multiplier k (one rendered plus $k-1$ generated frames), the per-rendered-frame generation cost is therefore:

$$T_{\text{gen}}(k) = T_{\text{heavy}} + (k-1)\cdot T_{\text{light}} \qquad (5)$$

so the marginal cost of each additional generated frame is only T_{light}. This is what makes 4× (and, in DLSS 4.5, 6×) multipliers economical. The learned flow-estimation component also replaces the dedicated Optical Flow Accelerator used by DLSS 3, which is a defining DLSS 4 change.

NVIDIA also explains why game interpolation differs from video interpolation [6]. Standard benchmarks contain no user interface, and engine motion vectors are wrong for specular highlights, reflections and UI motion. The network must therefore learn when to trust sparse engine inputs and when to override them. NVIDIA compares against video interpolators such as FILM on large-motion cases. Fig. 9 illustrates the timing consequence: interpolated frames between rendered frames F0 and F1 cannot be displayed until F1 exists, so interpolation adds roughly one rendered frame of latency, which Reflex must then claw back.

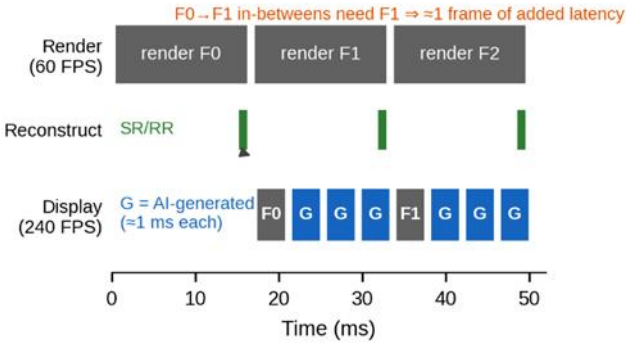


Fig. 9. Schematic MFG timing at a 60 FPS render rate and 4× output (240 FPS). In-between frames require the *next* rendered frame, so the display trails rendering by roughly one rendered frame; flip metering spaces the outputs evenly. Illustrative; generation cost ≈1 ms per frame on RTX 5090 [6].

5.4 Frame Pacing and Hardware Flip Metering

At 240–360+ displayed frames per second, the limiting factor becomes the coordination between neural generation, GPU scheduling and display presentation [6]. DLSS 4 therefore uses hardware flip metering and moves frame-pacing logic into the Blackwell display engine. TechPowerUp reports that this reduces frame-display variability by 5–10× [27], and the display engine supports up to 480 Hz at 4K [27]. NVIDIA cites flip metering and the improved tensor cores as the reasons MFG is RTX 50-exclusive [27]. Gamers Nexus points out that this also changes how performance must be measured: metered output is captured correctly by PresentMon's display-change interval rather than its present interval [30].

5.5 Reflex Low Latency and Reflex 2 Frame Warp

Reflex reduces latency by keeping the CPU render queue short. Reflex 2 Frame Warp goes further: just before display, it samples the newest camera position from the CPU, warps the just-rendered frame to that pose, and inpaints edges revealed by the warp [28]. This is a reprojection or extrapolation technique in the spirit of ExtraNet [31], not an interpolation technique, so it reduces latency instead of adding to it. NVIDIA's demonstration in *THE FINALS* on an RTX 5070 at 4K reported 56 ms without Reflex, 27 ms with Reflex Low Latency and 14 ms with Frame Warp, a 75% total reduction. In *VALORANT* on an RTX 5090 at more than 800 FPS, average latency was under 3 ms [28]. Caution is warranted: as of mid-2026, Frame Warp had still not shipped publicly, so these remain vendor demonstration figures [32].

5.6 Blackwell Tensor-Core Co-Design

The RTX 5090 pairs 21,760 CUDA cores and 32 GB of GDDR7 with fifth-generation tensor cores rated at 3,352 AI TOPS [2]. Blackwell adds FP4 and FP6 tensor formats, a second-generation FP8 Transformer Engine, and an AI Management Processor that schedules concurrent AI models alongside graphics work [2]. The DLSS 4 frame-generation model is 40% faster and uses 30% less VRAM than its predecessor; in *Warhammer 40,000: Darktide* it saves about 400 MB at 4K [27]. NVIDIA notes that a single rendered frame can involve several AI models across SR, RR and MFG, all executed within a few milliseconds [23]. Fig. 10 summarizes the reported gains.

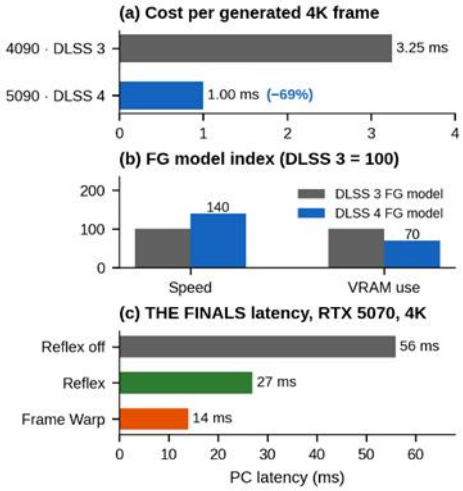


Fig. 10. DLSS 4 hardware co-design, vendor-reported: (a) cost per generated 4K frame [6]; (b) frame-generation model speed and memory relative to DLSS 3 [27]; (c) PC latency in THE FINALS with Reflex and Reflex 2 Frame Warp on an RTX 5070 at 4K [28].

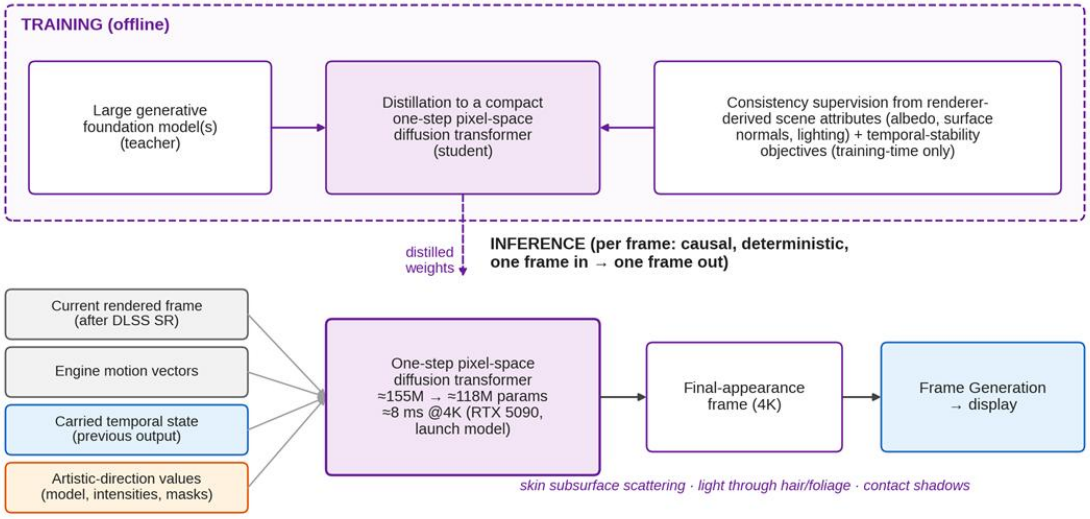


Fig. 11. DLSS 5 architecture. Training (dashed) distills large generative teachers into a compact one-step pixel-space diffusion transformer, with consistency supervision from albedo, normals and lighting. Inference is causal and deterministic, conditioned only on the post-SR frame, motion vectors, the carried temporal state and artistic-direction values [5], [11], [21], [26].

6. DLSS 5: The Generative Paradigm

6.1 Problem Statement

NVIDIA's ADLR report identifies two barriers to using generative foundation models in games [5]. First, such models do not reliably preserve developer-authored scene content and structure, even when their outputs look plausible. Second, their inference cost far exceeds interactive frame budgets, particularly at 4K. A third barrier is consistency: a standard diffusion model prompted ten times produces ten different images, which is unacceptable when a character's face cannot be allowed to change from one frame to the next [4].

6.2 3D-Guided Neural Rendering

DLSS 5 addresses these barriers with a renderer-grounded one-step pixel-space diffusion model [5]. At inference it is conditioned on the current rendered frame, engine motion vectors, a carried temporal state and artistic-direction values. During *training only*, consistency supervision from renderer-derived scene attributes grounds generation in the 3D scene [5]. At SIGGRAPH 2026, NVIDIA identified those attributes as albedo, surface normals and lighting, and described the model as a one-step diffusion transformer operating directly in pixel space and distilled from substantially larger foundation models [21]. Inference is causal and deterministic and processes one frame in and one frame out, unlike offline video diffusion, which operates on multi-frame chunks [21].

The runtime inputs are deliberately minimal. TechPowerUp summarizes them as the image the player would see with DLSS 5 disabled, plus motion vectors [11]. This is the architectural core of the phrase *3D-guided*: the 3D scene constrains the model during training, while inference needs only 2D image and motion data. Fig. 11 separates the two phases.

6.4 A Formal View

Abstracting from the published descriptions, the per-frame operator can be written as:

$$O_t = G_\theta(I_t, V_t, S_{t-1}, a), \quad S_t = O_t \qquad (6)$$

where I_t is the rendered (post-SR) frame, V_t the motion vectors, S_{t-1} the carried temporal state, a the artistic-direction vector, and G_θ a single network evaluation. Two contrasts with (2) are instructive. First, TAA's temporal state is an explicit accumulation buffer with a convergence guarantee, whereas DLSS 5's state is a learned feedback signal, so temporal stability must be *trained* rather than *derived*. Second, (2) is a convex combination of observed samples and cannot invent radiance, whereas G_θ can output values that appear in neither I_t nor S_{t-1}. That capacity is the point of the method and also the source of its risks.

6.5 Model Size, Placement and Cost

Pipeline position. DLSS 5 runs after SR upscaling and immediately before frame generation [11], [33]. MFG therefore interpolates from the post-DLSS-5 base frame rate.

Per-frame cost. Tom's Hardware reports that NVIDIA's paper gives an execution time of 8 ms per 4K frame on an RTX 5090 [34]. Digital Foundry's measurements imply about 8.3 ms [35].

Size and efficiency. NVIDIA's October 13, 2026 update shrinks the model by 24% to roughly 118 million parameters, cuts 4K VRAM use by 20%, and halves GPU overhead relative to the September model [26]; this implies a launch model of about 155 million parameters [26]. A Digital Foundry discussion summary cites a peak of about 731 MB of GPU memory for the launch model [35], a secondhand figure. The March 2026 reveal ran on two RTX 5090s, one rendering and one running the model [36]; by October, NVIDIA claimed the model was about 10× faster overall than at announcement [26]. Fig. 12 places these costs against frame budgets: the launch model alone consumed nearly the entire 120 FPS budget.

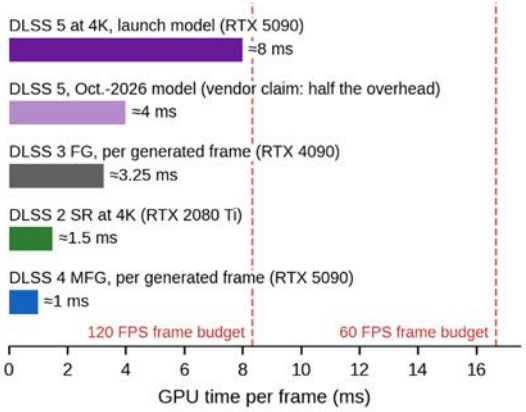


Fig. 12. Per-frame GPU cost of DLSS stages against 120 FPS and 60 FPS frame budgets. DLSS 5's generative stage is roughly an order of magnitude more expensive than a DLSS 4 generated frame [6], [12], [26], [34]. The October-2026 value is derived from NVIDIA's "half the overhead" claim.

6.6 Artistic Control Surface

Because DLSS 5 alters the authored look by design, NVIDIA exposes controls to developers [21], [25] (Fig. 13). These are: selection among several trained models (three were described at SIGGRAPH), mixable across scenes, gameplay and cut-scenes; *Structure Intensity* (high-frequency detail) and *Tone Intensity* (broad lighting and color response); semantic AI masks; and engine-level masks for asset groups such as glassware, water droplets and foliage. Input quality matters: richer source data such as path-traced lighting yields markedly more accurate results [25]. In *NBA 2K27*, Visual Concepts uses a per-pixel uplift mask to fine-tune character detail while respecting athlete likeness [4]. NVIDIA states that DLSS 5 preserves scanned facial geometry while enhancing skin subsurface scattering, light transmission through hair and ears, and contact shadows [25].

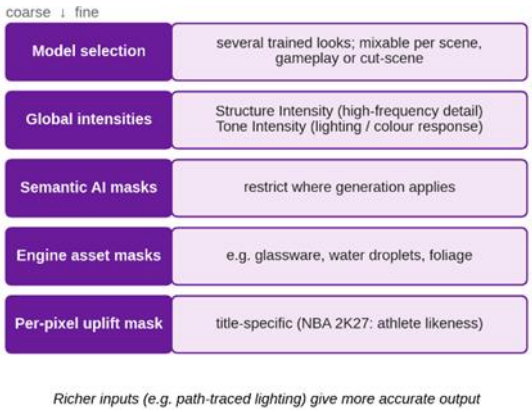


Fig. 13. DLSS 5 artistic-direction controls, from coarse (model choice) to fine (per-pixel masks) [4], [21], [25].

6.7 Community Reverse Engineering (Unverified)

The open-source OpenDLSS-NR project claims a bit-exact Vulkan reimplementation of the shipped network [37]. It describes a shifted-window transformer U-Net (in the spirit of Swin [15]) with a global ViT bottleneck, FP8 activations, roughly 141 MiB of weights, and inputs that include Gaussian-noise channels and the reprojected previous output. The last two details are consistent with a distilled one-step diffusion model and with the carried state in (6). None of this is confirmed by NVIDIA, and it should be treated as indicative only.

7. Comparative Analysis of the Divergent Paradigms

Table II compares the two paradigms across twelve design axes. The subsections below discuss the most consequential ones.

Table 2 - DLSS 4 Versus DLSS 5 Across Design Axes

Axis	DLSS 4 (reconstructive / interpolative)	DLSS 5 (generative)
Learning target	Higher-sample render of the same scene	Real-world appearance priors
Core networks	ViT SR / RR / DLAA; split-network MFG	One-step pixel-space diffusion transformer
Runtime inputs	Jittered colour, depth, motion vectors, exposure, ray-traced buffers	Post-SR colour, motion vectors, carried temporal state, artistic-direction values
Training-only signals	Ground-truth renders (tens of thousands of spp for RR)	Albedo, normals, lighting consistency; teacher distillation
Output	Higher-resolution frame plus up to 3 (DLSS 4.5: 5) interpolated frames	Re-appearanced final frame (one in, one out)
Cost (RTX 5090, 4K)	≈1 ms per generated frame	≈8 ms per frame at launch; ≈2× faster after Oct. 2026
Effect on frame rate	Raises rendered FPS (SR) and displayed FPS (MFG)	Lowers base FPS (≈51% in NBA 2K27 on RTX 5090)
Determinism	Deterministic; temporal accumulation	Deterministic; learned causal stability
Typical failure mode	Ghosting, disocclusion smear, UI / particle errors	Stylistic drift, identity or lighting alteration, invented detail
Hardware	SR / RR all RTX; MFG RTX 50 only	RTX 50 and GeForce NOW (RTX 40 announced as planned)
Integration	Streamline drop-in; driver-level model overrides	Art-directed: model choice, intensities, masks
Authorship	Preserves the authored look	Alters the authored look by design, under developer control

7.1 Target Distribution: Reconstruction Versus Generation

DLSS 4 SR and RR are trained toward references produced by the same renderer at higher sample counts [6]. Their error is, in principle, measurable against a ground truth. DLSS 5's target is real-world appearance, which the renderer never produces, so no pixel-exact ground truth exists for its outputs [5]. This is the deepest divergence. DLSS 4 can be *wrong* (ghosting, smearing); DLSS 5 can be *different* (plausible yet unauthored). Fig. 14 maps the family on two axes, departure from the renderer's ground truth and the proportion of frames synthesized in time, and shows that the shipping September 2026 configuration combines both extremes.

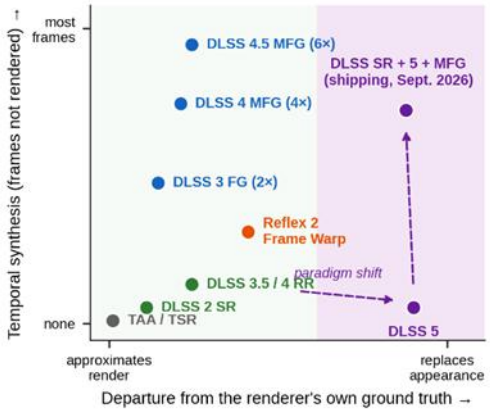


Fig. 14. Paradigm map. The horizontal axis measures departure from the renderer's own ground truth; the vertical axis measures temporal synthesis. DLSS 3–4.5 move upward (more synthesized frames) while staying near the renderer; DLSS 5 moves rightward. The shipping pipeline (SR → DLSS 5 → MFG) occupies the upper-right corner. Qualitative placement by the authors.

7.2 Fidelity Versus Hallucination

DLSS 4 MFG synthesizes frames that never existed *in time*, but each is bounded by two real frames. DLSS 5 synthesizes appearance that never existed *in the scene*, bounded by the input frame and the model's priors. NVIDIA argues that the output remains grounded in the engine's base frame, so final quality scales with input fidelity [4]. Critics counter that grounding constrains *structure* but not *style* [38]. Both positions are consistent with (6): the network is conditioned on I_t, which fixes geometry and layout, but its priors determine how materials and light read.

7.3 Determinism and Temporal Stability

Both paradigms are deterministic at inference. DLSS 4 inherits stability from temporal accumulation. DLSS 5 must learn stability explicitly, through causal conditioning on its carried state and stability-aware training objectives [21], because a one-step generator has no accumulation buffer that guarantees convergence.

7.5 Latency and Frame Budget

DLSS 4 SR *reduces* render time; MFG adds roughly one frame of buffering, which Reflex mitigates. DLSS 5 adds serial GPU work to every rendered frame. Modeling it as a fixed additive cost T_5 gives:

$$FPS' = 1000 / (1000 / FPS + T_5), \quad FPS_{\text{disp}} = k \cdot FPS' \quad (7)$$

With the measured RTX 5090 values in *NBA 2K27* (128.4 FPS without and 62.7 FPS with DLSS 5 [35]), (7) yields $T_5 = 15.95 - 7.79 \approx 8.2$ ms, consistent with the 8 ms reported from NVIDIA's paper [34]. The model makes an important property explicit: the *percentage* loss grows as base frame rate rises, because a fixed cost consumes a larger share of a shorter frame. Because DLSS 5 precedes frame generation, its cost also lowers the base rate from which MFG interpolates, so its latency penalty compounds with frame generation instead of being offset by it.

7.6 Compute, Memory and Power

A DLSS 4 generated frame costs about 1 ms; DLSS 5 costs about 8 ms per frame at 4K on the fastest consumer GPU at launch. Tom's Hardware also found the RTX 5090 Founders Edition consistently at or near its power limit with neural rendering enabled [34].

7.7 Developer Integration and Authorship

DLSS 4 is a drop-in through NVIDIA Streamline, and driver overrides can upgrade the model in existing titles. DLSS 5 requires art-direction work, including model choice, intensities and masks [25]. It behaves less like a performance toggle and more like a look-development tool. Fig. 15 condenses these trade-offs into a qualitative profile, and Fig. 16 illustrates the kind of visual difference at stake.

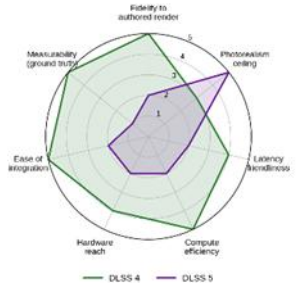


Fig. 15. Qualitative profile of DLSS 4 and DLSS 5 on seven axes (1 = weak, 5 = strong); authors' assessment synthesized from the evidence in Sections V–IX.

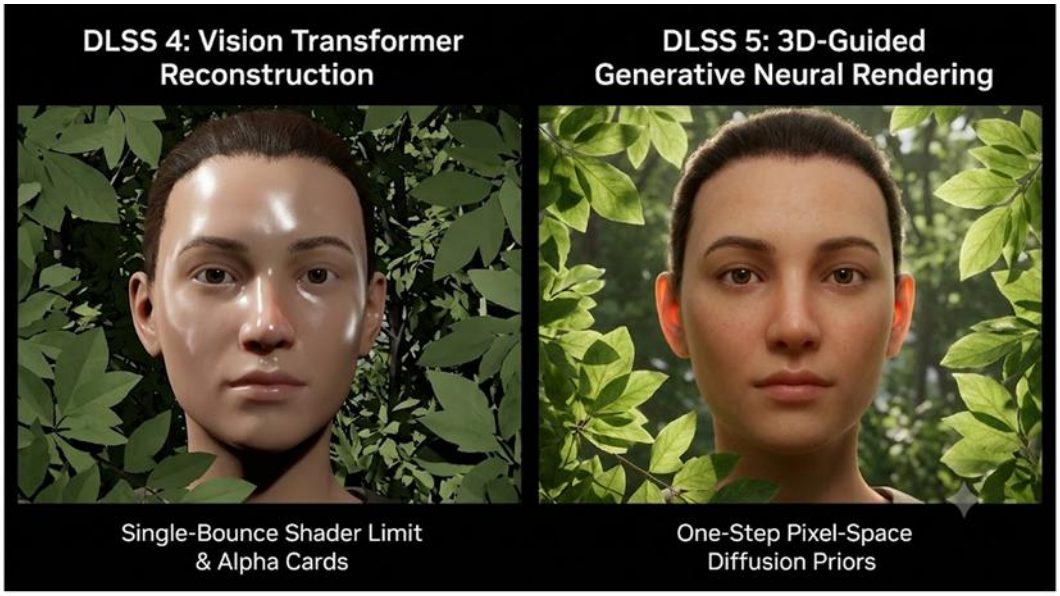


Fig. 16. Schematic illustration of the paradigm difference. Left: reconstruction remains bounded by what the renderer simulates (single-bounce shading, alpha-card foliage, simplified skin response). Right: generative rendering adds subsurface scattering in skin, light transmission through foliage and soft contact shading drawn from learned priors, the effects NVIDIA attributes to DLSS 5 [25].

7.8 Composition Rather Than Replacement

The shipping configuration, DLSS SR followed by DLSS 5 followed by frame generation [11], shows that the paradigms are complementary. Reconstruction supplies a high-resolution, temporally stable base image cheaply; generation re-appearances it; interpolation restores smoothness. The open engineering question is whether the generative stage can become cheap enough that this ordering still yields responsive frame rates on mid-range hardware.

8. Related and Competing Techniques

8.1 AMD FidelityFX Super Resolution (FSR)

FSR 1 was a spatial upscaler; FSR 2 and 3 are heuristic temporal upscalers, with FSR 3 adding analytic frame generation. FSR 4 (2025) is AMD's first machine-learning upscaler, designed around the FP8 capability of RDNA 4 [39]. FSR 4.1 ports the model to INT8 so that it can run on RDNA 3 (from July 2026) and RDNA 2 [39]; RX 6000 support arrived with Adrenalin 26.10.1 on October 20, 2026 [39], and AMD has confirmed FSR 4 for APUs by the end of 2026 [39]. AMD reports more than 300 games using its ML technologies by the second quarter of 2026 [39].

8.2 Intel XeSS

XeSS 2 combines XeSS-SR with XeSS-FG frame generation and XeLL low-latency technology [40]. Intel claimed up to 3.9× higher frame rates in *F1 24* on the Arc B580 [40]. The XeSS 2.1 SDK extends SR and FG to non-Intel GPUs supporting Shader Model 6.4 [40]; XMX-equipped GPUs run the advanced model, and others fall back to a DP4a model [40].

8.3 Unreal Engine Temporal Super Resolution

TSR is a platform-agnostic temporal upscaler designed for Nanite geometry and Lumen lighting and optimized for console GPUs [10]. Epic reports that its runtime was roughly halved, from 3.1 ms to 1.5 ms for 1080p-to-4K on console, to make 60 Hz *Fortnite* feasible, and that asynchronous compute can hide about 0.5 ms of that cost [10].

8.4 Apple MetalFX

MetalFX introduced spatial and temporal upscaling in 2022 [41]. Metal 4 (WWDC 2025) adds MetalFX Frame Interpolation, which generates one intermediate frame for every two input frames, and MetalFX Denoising [41]. Apple recommends a base frame rate of at least 30 FPS before interpolation [41].

8.5 Research Lineage

Frame extrapolation and neural reprojection, exemplified by ExtraNet [31], anticipate Reflex 2 Frame Warp. One-step diffusion distillation [17]–[19] and single-step neural rendering [20] anticipate DLSS 5. Table III summarizes the commercial landscape; notably, as of October 2026 no competitor ships a generative appearance stage.

Table 3 - Vendor Technologies for Upscaling, Frame Synthesis and Generation (October 2026)

Technology	Upscaling / generation	Frame synthesis	Low latency	Hardware path
NVIDIA DLSS 4 / 4.5	Transformer SR, DLAA	MFG 4×, 6×, Dynamic	Reflex; Reflex 2 (ann.)	Tensor cores, FP8
NVIDIA DLSS 5	DLSS SR + generative 3D-guided neural rendering	Feeds MFG	Reflex	RTX 50; GeForce NOW
AMD FSR 4 / 4.1	ML upscaler	FSR frame generation	Anti-Lag 2	FP8 (RDNA 4); INT8 (RDNA 2/3)
Intel XeSS 2	XeSS-SR	XeSS-FG	XeLL	XMX or DP4a fallback
Epic UE5 TSR	Heuristic temporal	—	—	Any modern GPU
Apple MetalFX	Spatial / temporal	Frame Interpolation (Metal 4)	—	Apple GPUs

Only DLSS 5 ships a generative appearance stage as of October 2026.

9. Performance and Quality Evaluation

9.1 Methodological Note

Headline vendor figures, such as “up to 8×” for DLSS 4 or several hundred FPS at 4K for DLSS 5 configurations, bundle SR, MFG and Reflex into a single number. Independent outlets increasingly report *rendered* frame rate and end-to-end PC latency alongside *displayed* frame rate. TechTimes shows that a DLSS 5 headline result built on a 6× multiplier implies a far lower rendered rate; for example, an RTX 5070 rendering at about 44 FPS [42]. Gamers Nexus raises similar concerns about review guidance for MFG [30]. Throughout this section we report all three quantities where available, and Appendix B proposes a reporting checklist.

9.2 DLSS 4 Multi Frame Generation in Practice

TechSpot’s *Alan Wake 2* measurements illustrate the latency–smoothness trade-off [43] (Fig. 17). Native rendering ran at 96 FPS with 32 ms latency; 2× frame generation displayed 165 FPS but raised latency to 39 ms because the underlying render rate fell by about 14%. Under 4× MFG, displayed

frames rose by about 198% while the render rate fell by about 25% and latency rose from 34 ms to 44 ms [43]. The conclusion is consistent across outlets: MFG improves *motion smoothness*, not *responsiveness*, and it needs a healthy base frame rate.

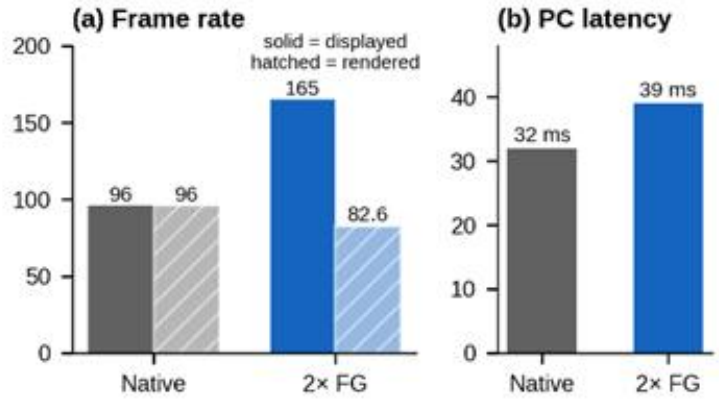


Fig. 17. Alan Wake 2 with and without 2× frame generation (TechSpot [43]). Displayed FPS rises by 72%, but rendered FPS (derived from the reported 14% drop) and therefore input latency move the other way.

9.4 C. Upscaling Gains and the CPU-Bound Regime: Ghost of Tsushima

TechPowerUp's *Ghost of Tsushima* comparison (RTX 4080, Very High preset, DLSS 3.7.0) found that DLSS, XeSS 1.3 and FSR 3 give roughly a 35% performance gain in Quality mode over native rendering at 4K and about 30% at 1440p and 1080p, and that DLSS Quality with frame generation roughly doubles performance at all three resolutions [44]. AMD's frame generation produced slightly more frames in GPU-limited scenes but with worse image quality, DLAA cost about 2% versus TAA, and the game's native TAA was weak enough that even 1080p DLSS Performance (540p internal) looked better [44].

The shrinking upscaling gain at lower resolutions (Fig. 18) is expected. As internal resolution falls, the GPU stops being the bottleneck and frame rate approaches a host-CPU ceiling, which upscaling cannot raise; frame generation, which bypasses the CPU for generated frames, continues to scale. The same CPU-bound regime also slows temporal convergence per (3) [10].

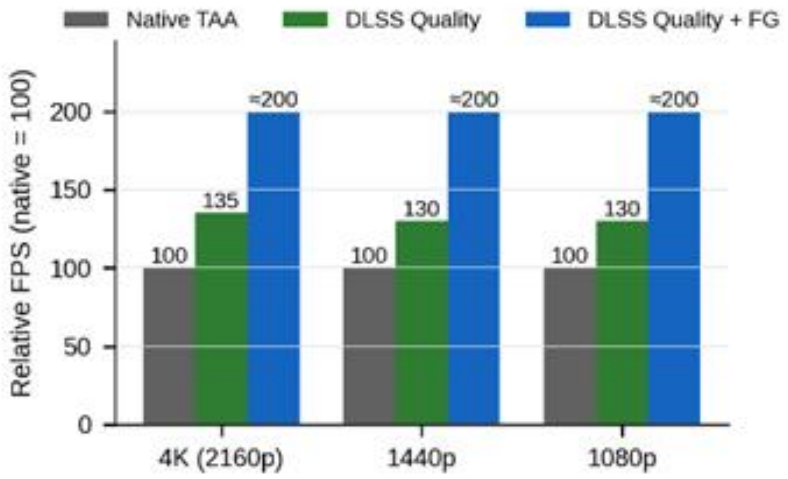


Fig. 18. Relative performance in Ghost of Tsushima (RTX 4080), native TAA = 100. Values are derived from the percentage gains reported in [44].

9.5 DLSS 5 in NBA 2K27

Table IV and Fig. 19 consolidate the first independent measurements of DLSS 5. Digital Foundry measured the RTX 5090 dropping from 128.4 to 62.7 FPS at 4K with ray tracing (−51.2%), with losses of 58%, 55% and 41% on the RTX 5080, 5070 and 5060 respectively; the RTX 5080 landed just above 40 FPS [35]. With 4× MFG and DLSS 5 enabled, Digital Foundry measured roughly 50–60 ms of PC latency on the RTX 5090 and 60–70+ ms on the RTX 5080 and 5060 [33]; another review reported 31.9 ms rising to 79.2 ms with DLSS 5 plus 2× FG [33]. NVIDIA claims its October model raises *NBA 2K27* frame rates by up to 1.4× [26]. Tom's Hardware, testing community-injected builds via a leaked DLL, found 39–49% drops on the

RTX 5090 and noted that an official Streamline integration may cost less [34]. Qualitatively, early hands-on reviews were positive about character lighting and skin [33].

Table 4 - NBA 2K27 Performance With and Without DLSS 5 (4K, Maximum Settings, Ray Tracing On)

GPU	DLSS 5 off	DLSS 5 on	Change
RTX 5090	128.4 FPS	62.7 FPS	−51.2% (≈8.2 ms per frame)
RTX 5080	—	just above 40 FPS	−58%
RTX 5070	—	—	−55%
RTX 5060	—	—	−41%

Source: Digital Foundry measurements as reported by Notebookcheck [35]; per-frame cost computed with (7).

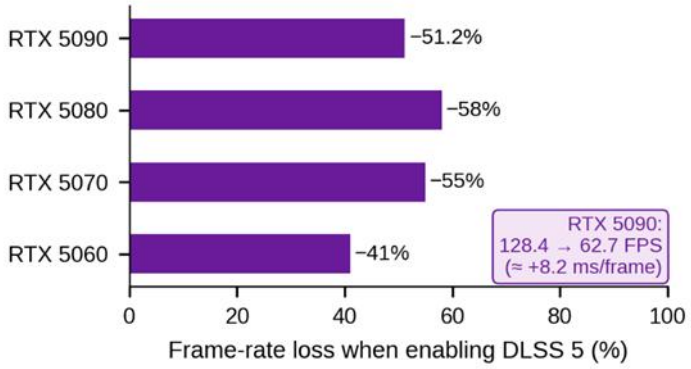


Fig. 19. Frame-rate loss on enabling DLSS 5 in NBA 2K27 by GPU tier [35]. The RTX 5090 inset shows the absolute numbers from which (7) recovers an ≈8.2 ms per-frame cost.

Fig. 20 orders the latency figures reported across this survey. The comparison is illustrative rather than controlled, since games, GPUs and settings differ, but the ordering is instructive. Reprojection (Frame Warp) sits at the bottom, interpolation adds tens of milliseconds, and DLSS 5 combined with frame generation sits at the top.

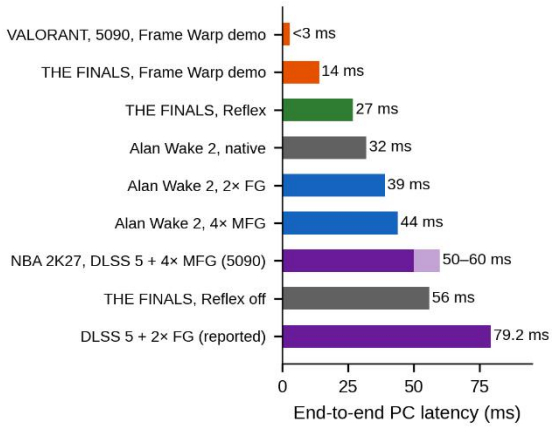


Fig. 20. End-to-end PC latency reported across configurations [28], [33], [43]. Different titles and GPUs: illustrative ordering, not a controlled comparison. Frame Warp figures are vendor demonstrations of a not-yet-shipped feature [32].

10. Gaming Applications

10.1 Path-Traced AAA Titles

The combination of SR, RR and MFG is what makes full path tracing viable at high resolutions. More than 250 games and applications supported DLSS 4 MFG by January 2026 [29]. In this setting, the reconstructive guarantees of DLSS 4 are exactly what developers want: the network approximates the path tracer they authored.

10.2 Competitive Esports

Competitive titles prioritize latency over smoothness. SR combined with Reflex (and, if it ships, Frame Warp) is the appropriate configuration, while interpolation-based MFG and the generative stage are generally counter-productive because both add latency (Section IX).

10.3 Sports and Character-Centric Games

DLSS 5's first deployment targets scanned athlete likenesses in *NBA 2K27*, where skin subsurface scattering, light transmission through hair and ears, and contact shadows are the visible gains [4], [25]. Titles confirmed or demonstrated with DLSS 5 include *Hogwarts Legacy* [26] as well as titles shown at the GTC 2026 reveal [25], [36].

10.4 Cloud Gaming

GeForce NOW streams DLSS 5 from RTX 5080-class servers [26]. This moves the generative cost off client hardware, so even thin clients can see generative output, but it adds network latency to an already latency-heavy pipeline.

10.5 Consoles, Handhelds and Cross-Vendor Reach

The Nintendo Switch 2 is the first console to use DLSS, while AMD's FSR 4.1 is planned for consoles and older Radeon hardware [39]. Learned reconstruction is therefore becoming a cross-platform baseline; generative appearance, for now, remains NVIDIA-specific.

10.6 Remastering and Modding

Community tools that inject DLSS into unsupported titles show strong demand for retrofitting. However, unofficial DLSS 5 builds running through leaked libraries and injection layers [34], [45] operate entirely outside developer control, which undermines the art-direction premise of Section VI.

11. Non-Gaming Applications

The same principles apply far beyond games: render or simulate cheaply, then learn the missing detail. The appropriate paradigm, however, depends on whether the application can tolerate content the underlying simulation did not produce. Fig. 21 summarizes our assessment.

	Super resolution	Ray reconstruction / denoising	Frame generation (interpolation)	Reprojection / frame warp	Generative appearance (DLSS 5-style)
AAA single-player gaming	Recommended	Recommended	Recommended	Conditional	Conditional
Competitive esports	Recommended	Conditional	Avoid	Recommended	Avoid
Cloud gaming / streaming	Recommended	Recommended	Conditional	Conditional	Conditional
Digital twins & industrial simulation	Recommended	Recommended	Conditional	Conditional	Avoid
Architecture, CAD & product visualization	Recommended	Recommended	Conditional	Conditional	Avoid
Film & virtual production (previs)	Recommended	Recommended	Conditional	Conditional	Conditional
AV / robotics synthetic data	Conditional	Conditional	Avoid	Avoid	Avoid
Medical & scientific visualization	Conditional	Conditional	Avoid	Avoid	Avoid
Remote sensing / flood mapping	Recommended	N/A	N/A	N/A	Avoid
VR / XR	Recommended	Recommended	Avoid	Recommended	Avoid

Fig. 21. Application-domain suitability matrix (authors' assessment). Rule of thumb: avoid generative appearance wherever physical, diagnostic or contractual fidelity is required, and avoid interpolation wherever input latency is critical. For remote sensing, "super resolution" denotes domain-specific SR models such as those evaluated in [47], not DLSS itself.

11.1 Digital Twins and Industrial Simulation

NVIDIA Omniverse's RTX Real-Time 2.0 mode is a physically based path-tracing renderer built on the DLSS suite [46]. The Omniverse documentation notes that, without ray reconstruction, real-time-mode images appear significantly noisy [46]. NVIDIA's computer-aided engineering (CAE) digital-twin blueprint pairs accelerated physics with real-time RTX visualization for aerospace, automotive and energy design [46]. Here reconstruction is ideal: it accelerates *faithful* visualization of simulated data.

11.2 Architecture, CAD and Product Visualization

Real-time path-traced walkthroughs inherit DLSS 4's reconstructive guarantees. Generative appearance is risky wherever material accuracy is contractual, for example in a client approving a finish or a manufacturer validating a coating.

11.3 Film and Virtual Production

LED-volume and previsualization pipelines benefit directly from RR and SR. A DLSS 5-style generative stage may be acceptable in previs, where the goal is believable look development rather than final pixels, but it must be disabled for anything feeding a physically accurate final render.

11.4 Autonomous-Vehicle and Robotics Simulation

Synthetic sensor data must be physically faithful. Reconstructive denoising is acceptable when validated, but generative appearance priors could introduce systematic domain bias into training data, for example by "beautifying" wet road surfaces. It should be disabled, or at least audited, for data generation.

11.5 Medical and Scientific Imaging

Deep super-resolution of magnetic resonance imaging (MRI) is an established field. Lyu *et al.* demonstrated two-fold resolution enhancement of MR images [48], and a recent comprehensive survey covers convolutional, transformer and diffusion-based MRI super-resolution [48]. The fidelity-versus-hallucination tension of Section VII applies here with clinical stakes: invented detail in a diagnostic image is a safety hazard, so reconstructive methods with quantified uncertainty are appropriate and generative appearance is not.

11.6 Remote Sensing and Flood Mapping

Choi *et al.* compare ResUNet, EDSR and RCAN for 4× and 8× super-resolution of urban flood maps in downtown Portland [47]. They evaluate not only PSNR and SSIM but also flood-specific metrics: critical success index, depth RMSE and flood-severity classes. RCAN achieved the highest recall for non-flood (98.06%) and severe-flood (96.48%) cells, while EDSR led for mild flooding (97.95%). Errors concentrated at wet–dry boundaries and in complex urban geometry [47]. Related work shows deep learning enabling super-resolution hydrodynamic flood modeling under spatiotemporally varying rainfall [47]. These studies are a direct non-gaming analogue of DLSS SR: a cheap coarse simulation followed by learned detail, judged against a high-resolution reference. They also show why task-specific metrics, not just image metrics, are needed, a lesson that carries back to evaluating DLSS 5.

11.7 VR and XR

Head-mounted displays need high refresh rates and low motion-to-photon latency. SR and reprojection, the conceptual ancestor of Frame Warp, are well suited; interpolation-based frame generation and costly generative stages conflict with comfort requirements.

12. Challenges, Limitations and Ethical Concerns

12.1 "Fake Frames" and Measurement Integrity

MFG frames add smoothness without sampling new input. Gamers Nexus criticized vendor review guidance and metric choices for MFG [30], and TechTimes showed how large frame multipliers can hide low rendered rates [42]. The field needs standard reporting of rendered FPS, displayed FPS and PC latency (Appendix B).

12.2 Latency Compounding

Because DLSS 5 precedes frame generation, it lowers the base rate that frame generation relies on (7). The measured 50–80 ms PC latencies of DLSS 5 plus frame-generation configurations [33] are unsuitable for competitive play.

12.3 Artistic Intent and Aesthetic Homogenization

The March 2026 reveal drew wide backlash across Reddit [49], other social platforms and the press [36], [38]. Commentators coined the term "yassified graphics" for faces that appeared smoothed and beautified, and criticized the technology for overriding established art styles [36]. NVIDIA CEO Jensen Huang called critics "completely wrong," arguing that DLSS 5 fuses the controllability of authored geometry and textures with generative AI [50]. Our assessment is that the backlash was partly a demo artifact (maximum intensities, a dual-GPU preview), but the underlying concern is structural. A generative stage trained on real-world priors pulls output toward a learned notion of realism, so stylized art direction must *actively resist* it through the masks and intensity controls of Section VI [38].

12.4 Likeness and Identity

Sports and licensed titles render real people. Visual Concepts' use of a per-pixel uplift mask to respect athlete likeness [4] shows that identity preservation is now an engineering requirement with legal and contractual implications, not merely an aesthetic preference.

12.5 Hallucination and Verifiability

With no ground truth for DLSS 5 output, standard full-reference metrics such as PSNR and SSIM do not apply. Evaluation will require identity-preservation metrics for faces, semantic-consistency checks against G-buffer attributes (albedo, normals, lighting), temporal-stability metrics, and controlled perceptual user studies, mirroring the task-specific metrics used in flood-mapping SR [47].

12.6 Hardware Exclusivity, Cost and Power

MFG and DLSS 5 require RTX 50 hardware; RTX 40 support for DLSS 5 has been described as planned but not delivered [26]. Power draw near board limits on the RTX 5090 Founders Edition [34] suggests the generative stage is constrained by power as well as compute.

12.7 Ecosystem Governance and Vaporware Risk

Leaked libraries and injection mods place DLSS 5 into games whose developers never tuned it [34], [45], which defeats the developer-control premise. Separately, Reflex 2 Frame Warp, announced in January 2025, remained unshipped roughly 18 months later [32], so vendor latency figures for unreleased features should be treated as provisional. Table V maps each challenge to candidate mitigations.

Table 5 - Open Challenges and Candidate Mitigations

Challenge	Candidate mitigation
Measurement integrity (MFG)	Report rendered FPS, displayed FPS and PC latency together; disclose multiplier
Latency compounding (DLSS 5 + FG)	Cheaper generative stage; run before upscaling; pair with reprojection instead of interpolation
Aesthetic homogenization	Per-title model choice; intensity limits; semantic and engine masks; style-conditioned training
Likeness / identity	Per-pixel masks; identity-preservation metrics in QA; licensing review
No ground truth	G-buffer consistency scores, identity and temporal metrics, perceptual studies
Hardware reach and power	Smaller distilled models, FP4 quantization, cloud offload
Ecosystem governance	Signed integrations; developer opt-in; visible "generative rendering on" disclosure

13. Future Directions

1) *Cheaper generative stages.* Running generation *before* upscaling, at render resolution, as community "neural upstream" experiments attempt [45]; FP4 quantization [2]; and further distillation [18], [19].

2) *Unified models.* Merging SR, RR and generative appearance into one network that accepts G-buffer attributes at inference, trading DLSS 5's minimal-input design for stronger, verifiable grounding.

3) *Extrapolation over interpolation.* Frame warping and learned extrapolation [31] to remove interpolation's latency floor.

4) *Faithfulness metrics and standards.* Shared benchmarks for G-buffer consistency, identity preservation and temporal stability, plus user-facing disclosure when generative rendering is active.

5) *Neural materials and shaders.* Tensor-core access from shaders points toward neural materials and texture compression *inside* the renderer [2], which may reduce the need for post-hoc generation.

6) *Cross-domain transfer.* Carrying DLSS-style temporal reconstruction into simulation super-resolution (flood and climate downscaling, medical imaging) with domain-specific fidelity constraints [47], [48].

14. Conclusion

DLSS 4 completes the reconstructive and interpolative program that began with DLSS 2 in 2020. It uses a transformer that approximates higher-sample renders, frame synthesis at about 1 ms per generated frame, and display-engine pacing. DLSS 5 is not its successor in kind. It is a generative final stage that supplies real-world appearance the renderer cannot simulate, at about 8 ms per 4K frame at launch, and in practice it composes with DLSS 4 (SR → DLSS 5 → MFG). The divergence is best understood through the target distribution. DLSS 4 is judged against a ground truth it can approach; DLSS 5 is judged against a reality the renderer never computed. Its success will depend less on raw image quality than on three unresolved issues: latency under frame generation, verifiable faithfulness to authored content, and governance of artistic intent. For practitioners, our recommendation is to use DLSS 4 wherever fidelity to the renderer matters, and to use DLSS 5 selectively, under deliberate art direction, where photoreal character and material appearance is the goal and the latency budget allows.

Appendix A: Nomenclature

Table 6 - Abbreviations and Symbols

Term	Meaning
DLSS	Deep Learning Super Sampling
SR / RR	Super Resolution / Ray Reconstruction
DLAA	Deep Learning Anti-Aliasing (native-resolution DLSS)
FG / MFG	Frame Generation / Multi Frame Generation
TAA / TAAU	Temporal antialiasing / temporal antialiasing upsampling
TSR	Temporal Super Resolution (Unreal Engine)
AGAA	Aggregate G-buffer antialiasing
MSAA / SSAA	Multisample / supersample antialiasing
spp	Samples per pixel
ViT	Vision transformer
DMD / ADD	Distribution matching / adversarial diffusion distillation
s, r, k	Screen percentage; per-axis upscale factor; frame multiplier
ρ	Fraction of displayed pixels actually shaded, (4)
T_5	Per-frame cost of the DLSS 5 stage, (7)
G_θ	DLSS 5 one-step generator, (6)

Appendix B: Recommended Reporting Checklist

Based on the evidence reviewed in Sections IX and XII, we recommend that studies of neural rendering report at least the following:

1) rendered (base) frame rate, displayed frame rate and the frame multiplier *k*;

2) end-to-end PC latency, the measurement tool and metric (e.g., display-change interval for flip-metered output [30]);

3) internal resolution or screen percentage *s*, output resolution and preset;

4) GPU, driver, model version (e.g., DLSS 5 September versus October 2026 model) and integration path (official SDK versus injection);

5) for generative stages, all artistic-direction settings (model, intensities, masks) and at least one faithfulness metric beyond full-reference image metrics.

References

[1] J. T. Kajiya, "The rendering equation," in *Proc. 13th Annu. Conf. Comput. Graph. Interact. Techn. (SIGGRAPH)*, 1986, pp. 143–150.

[2] NVIDIA Corp., "NVIDIA RTX Blackwell GPU architecture," Whitepaper, Santa Clara, CA, USA, 2025. [Online]. Available: https://hardwareand.co/images/dossiers/gpu/blackwell/architecture/blackwell-architecture.pdf

[3] NVIDIA Newsroom, "NVIDIA introduces DLSS 3 with breakthrough AI-powered frame generation for up to 4x performance," Sep. 20, 2022. [Online]. Available: https://nvidianews.nvidia.com/news/nvidia-introduces-dlss-3-with-breakthrough-ai-powered-frame-generation-for-up-to-4x-performance

[4] NVIDIA, "DLSS 5: Neural rendering brings lifelike lighting to NBA 2K27," GeForce News, Sep. 2026. [Online]. Available: https://www.nvidia.com/en-us/geforce/news/dlss-5-3d-guided-neural-rendering/

[5] NVIDIA ADLR, "DLSS 5: Generative neural rendering," project page, 2026. [Online]. Available: https://research.nvidia.com/labs/adlr/DLSS5/

[6] NVIDIA ADLR, "DLSS 4: Transforming real-time graphics with AI," technical overview, 2025. [Online]. Available: https://research.nvidia.com/labs/adlr/DLSS4/

[7] L. Yang, S. Liu, and M. Salvi, "A survey of temporal antialiasing techniques," *Comput. Graph. Forum*, vol. 39, no. 2, pp. 607–621, 2020. [Online]. Available: http://behindthepixels.io/assets/files/TemporalAA.pdf

[8] C. Crassin, M. McGuire, K. Fatahalian, and A. Lefohn, "Aggregate G-buffer anti-aliasing," *IEEE Trans. Vis. Comput. Graphics*, vol. 22, no. 10, pp. 2215–2228, Oct. 2016. [Online]. Available: https://graphics.stanford.edu/~kayvonf/papers/agaa_tvcg2016.pdf

[9] J. Knodt, J. Bartusek, S.-H. Baek, and F. Heide, "Neural ray-tracing: Learning surfaces and reflectance for relighting and view synthesis," arXiv:2104.13562v2, 2021. [Online]. Available: https://arxiv.org/pdf/2104.13562v2

[10] Epic Games, "Temporal Super Resolution in Unreal Engine," Unreal Engine 5 Documentation, 2025. [Online]. Available: https://dev.epicgames.com/documentation/unreal-engine/temporal-super-resolution-in-unreal-engine

[11] TechPowerUp, "NVIDIA DLSS 5 technical preview – Neural rendering is here – How it works," 2026. [Online]. Available: https://www.techpowerup.com/review/nvidia-dlss-5-technical-preview/2.html

[12] E. Liu, "DLSS 2.0 – Image reconstruction for real-time rendering with deep learning," presented at the GPU Technology Conf. (GTC), 2020. [Online]. Available: http://behindthepixels.io/assets/files/DLSS2.0.pdf

[13] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NeRF: Representing scenes as neural radiance fields for view synthesis," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2020, pp. 405–421.

[14] A. Dosovitskiy *et al.*, "An image is worth 16×16 words: Transformers for image recognition at scale," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2021.

[15] Z. Liu *et al.*, "Swin Transformer: Hierarchical vision transformer using shifted windows," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, 2021, pp. 10012–10022.

[16] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2020, pp. 6840–6851.

[17] T. Yin *et al.*, "One-step diffusion with distribution matching distillation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2024. [Online]. Available: https://arxiv.org/abs/2311.18828

[18] T. Yin *et al.*, "Improved distribution matching distillation for fast image synthesis," in *Proc. Adv. Neural Inf. Process. Syst. (NeurIPS)*, 2024.

[19] A. Sauer *et al.*, "Fast high-resolution image synthesis with latent adversarial diffusion distillation," in *Proc. SIGGRAPH Asia*, 2024, doi: 10.1145/3680528.3687625.

[20] "RenderFlow: Single-step neural rendering via flow matching," arXiv:2601.06928, Jan. 2026. [Online]. Available: https://arxiv.org/pdf/2601.06928

[21] PC Games Hardware, "DLSS 5: Three AI models, single-GPU operation and full control," SIGGRAPH 2026 coverage, 2026. [Online]. Available: https://www.pcgameshardware.de/Deep-Learning-Super-Sampling-Software-277618/Specials/dlss-5-three-ai-models-single-gpu-control-siggraph-2026-1548907/

[22] NVIDIA, "NVIDIA DLSS: Your questions, answered," GeForce News, 2019. [Online]. Available: https://nvidia.com/en-us/geforce/news/nvidia-dlss-your-questions-answered

[23] NVIDIA, "NVIDIA DLSS 4 introduces Multi Frame Generation & enhancements for all DLSS technologies," GeForce News, Jan. 2025. [Online]. Available: https://www.nvidia.com/en-us/geforce/news/dlss4-multi-frame-generation-ai-innovations/

[24] NVIDIA Technical Blog, "NVIDIA DLSS 4.5 delivers Super Resolution upgrades and new Dynamic Multi Frame Generation," 2026. [Online]. Available: https://developer.nvidia.com/blog/nvidia-dlss-4-5-delivers-super-resolution-upgrades-and-new-dynamic-multi-frame-generation

[25] NVIDIA Technical Blog, "What's new for game developers: DLSS 5 with 3D-guided neural rendering, NVIDIA ACE updates, and new RTX Kit capabilities," 2026. [Online]. Available: https://developer.nvidia.com/blog/whats-new-for-game-developers-dlss-5-with-3d-guided-neural-rendering-nvidia-ace-updates-and-new-rtx-kit-capabilities/

[26] NVIDIA, "DLSS 5 comes to Hogwarts Legacy & RTX Spark, powered by 2X faster model," GeForce News, Oct. 2026. [Online]. Available: https://www.nvidia.com/en-us/geforce/news/dlss-5-faster-ai-model-hogwarts-legacy-rtx-spark-laptops/

[27] TechPowerUp, "NVIDIA GeForce RTX 50 technical deep dive – DLSS 4," Jan. 2025. [Online]. Available: https://www.techpowerup.com/review/nvidia-geforce-rtx-50-technical-deep-dive/4.html

[28] TechPowerUp, "NVIDIA Reflex 2 with new Frame Warp technology reduces latency in games by up to 75%, coming to THE FINALS and VALORANT," Jan. 2025. [Online]. Available: https://www.techpowerup.com/330822/nvidia-reflex-2-with-new-frame-warp-technology-reduces-latency-in-games-by-up-to-75-coming-to-the-finals-and-valorant

[29] NVIDIA, "DLSS 4 is available in over 250 games, and coming to 007 First Light, Phantom Blade Zero, PRAGMATA with path tracing & many more," GeForce News, Jan. 2026. [Online]. Available: https://www.nvidia.com/en-us/geforce/news/dlss-4-rtx-path-tracing-game-announcements-ces-2026/

[30] Gamers Nexus, "'Fake frames' tested: DLSS 4.0 MFG 4x & NVIDIA's misleading review guide," 2025. [Online]. Available: https://gamersnexus.net/gpus/fake-frames-tested-dlss-40-mfg-4x-nvidias-misleading-review-guide

[31] J. Guo *et al.*, "ExtraNet: Real-time extrapolated rendering for low-latency temporal supersampling," *ACM Trans. Graph.*, vol. 40, no. 6, Art. no. 278, 2021.

[32] Hardware Busters, "NVIDIA Reflex 2 is 18 months old and still hasn't shipped," 2026. [Online]. Available: https://hwbusters.com/news/nvidia-reflex-2-is-18-months-old-and-still-hasnt-shipped/

[33] XDA Developers, "The RTX 5090 loses half its performance to DLSS 5, and Nvidia's favorite trick won't save it this time," 2026. [Online]. Available: https://www.xda-developers.com/nvidias-rtx-5090-loses-half-its-power-to-dlss-5-and-their-usual-trick-wont-save-it/

[34] Tom's Hardware, "We explored early DLSS 5 performance with community mods — and the limits of the 12V-2x6 power connector may hold it back on the RTX 5090," 2026. [Online]. Available: https://www.tomshardware.com/pc-components/gpus/we-explored-early-dlss-5-performance-with-community-mods-and-the-limits-of-the-12v-2x6-power-connector-may-hold-it-back-on-the-rtx-5090

[35] Notebookcheck, "First DLSS 5 game tested: DLSS 5 humbles RTX 5090 with over 50% performance loss," 2026. [Online]. Available: https://www.notebookcheck.net/First-DLSS-5-game-tested-DLSS-5-humbles-RTX-5090-with-over-50-performance-loss.1391978.0.html

[36] Euronews, "'Yassified graphics' and uncanny faces: Gamers turn Nvidia's DLSS 5 reveal into a meme storm," Mar. 19, 2026. [Online]. Available: https://euronews.com/next/2026/03/19/yassified-graphics-and-uncanny-faces-gamers-turn-nvidias-dlss-5-reveal-into-a-meme-storm

[37] maanHimself, "OpenDLSS-NR: A Vulkan reimplementation of NVIDIA's DLSS 5 neural rendering network," GitHub repository, 2026. [Online]. Available: https://github.com/maanHimself/OpenDLSS-NR

[38] TechRadar, "Why do gamers already hate DLSS 5? Here are 3 key reasons – and why history suggests Nvidia will win them over eventually," 2026. [Online]. Available: https://www.techradar.com/computing/gpu/why-do-gamers-already-hate-dlss-5-here-are-3-key-reasons-and-why-history-suggests-nvidia-will-win-them-over-eventually

[39] PC Gamer, "AMD bringing FSR 4.1 ML-powered upscaling to RX 7000 series in July, RX 6000 series in 2027," 2026. [Online]. Available: https://www.pcgamer.com/hardware/graphics-cards/amd-bringing-fsr-4-1-ml-powered-upscaling-to-rx-7000-series-in-july-rx-6000-series-in-2027/

[40] TweakTown, "Intel announces XeSS 2 with new Frame Generation and Low Latency tech," Dec. 2024. [Online]. Available: https://www.tweaktown.com/news/101974/intel-announces-xess-2-with-new-frame-generation-and-low-latency-tech/index.html

[41] Engadget, "Metal 4 for macOS Tahoe 26 introduces new gaming features like frame interpolation," Jun. 2025. [Online]. Available: https://www.engadget.com/gaming/metal-4-for-macos-tahoe-26-introduces-new-gaming-features-like-frame-interpolation-183247623.html

[42] TechTimes, "DLSS 5 benchmarks use 6X frame multiplier: RTX 5070 actually renders at 44 FPS," Sep. 2, 2026. [Online]. Available: https://www.techtimes.com/articles/326364/20260902/dlss-5-benchmarks-use-6x-frame-multiplier-rtx-5070-actually-renders-44-fps.htm

[43] TechSpot, "A review of Nvidia's DLSS 4 Multi Frame Generation," 2025. [Online]. Available: https://www.techspot.com/article/2945-nvidia-dlss-4/

[44] TechPowerUp, "Ghost of Tsushima: DLSS vs. FSR vs. XeSS comparison review," May 2024. [Online]. Available: https://www.techpowerup.com/review/ghost-of-tsushima-dlss-vs-fsr-vs-xess-comparison/

[45] TechPowerUp, "Modders rework DLSS 5's rendering pipeline for a big performance boost," 2026. [Online]. Available: https://www.techpowerup.com/352476/modders-rework-dlss-5s-rendering-pipeline-for-a-big-performance-boost

[46] NVIDIA, "RTX – Real-Time 2.0 mode," Omniverse Materials and Rendering Documentation. [Online]. Available: https://docs.omniverse.nvidia.com/materials-and-rendering/latest/rtx-renderer_rt.html

[47] Choi *et al.*, "Comparative evaluation of deep learning–based super-resolution models for urban flood mapping," *J. Flood Risk Manage.*, vol. 18, no. 4, Art. no. e70144, 2025, doi: 10.1111/jfr3.70144. [Online]. Available: https://onlinelibrary.wiley.com/doi/epdf/10.1111/jfr3.70144

[48] M. Khateri *et al.*, "MRI super-resolution with deep learning: A comprehensive survey," arXiv:2511.16854, 2025.

[49] Reddit r/nvidia, community discussion thread on the DLSS 5 reveal, 2026. [Online]. Available: https://www.reddit.com/r/nvidia/s/tUui5t4aAX

[50] Tom's Guide, "'They're completely wrong': Nvidia CEO Jensen Huang responds to DLSS 5 criticism," 2026. [Online]. Available: https://www.tomsguide.com/gaming/theyre-completely-wrong-nvidia-ceo-jensen-huang-responds-to-dlss-5-criticism