

NVIDIA Blackwell GeForce RTX 50 Series Opens New World of AI Computer Graphics

Next Generation of GeForce RTX GPUs Deliver Stunning Visual Realism and 2x Performance Increase, Made Possible by AI, Neural Shaders and DLSS 4



Figure 1: NVIDIA Blackwell GeForce RTX 50 Series

NVIDIA today unveiled the most advanced consumer GPUs for gamers, creators and developers —the GeForce RTX™ 50 Series Desktop and Laptop GPUs.

Powered by the NVIDIA Blackwell architecture, fifth-generation Tensor Cores and fourth-generation RT Cores, the GeForce RTX 50 Series delivers breakthroughs in AI-driven rendering, including neural shaders, digital human technologies, geometry and lighting.

“Blackwell, the engine of AI, has arrived for PC gamers, developers and creatives,” said Jensen Huang, founder and CEO of NVIDIA. “Fusing AI-driven neural rendering and ray tracing, Blackwell is the most significant computer graphics innovation since we introduced programmable shading 25 years ago.”

The GeForce RTX 5090 GPU —the fastest GeForce RTX GPU to date —features 92 billion transistors, providing over 3,352 trillion AI operations per second (TOPS) of computing power. Blackwell architecture innovations and DLSS 4 mean the GeForce RTX 5090 GPU outperforms the GeForce RTX 4090 GPU by up to 2x.

GeForce Blackwell comes to laptops with all the features of desktop models, bringing a considerable upgrade to portable computing, including extraordinary graphics capabilities and remarkable efficiency. The Blackwell generation of NVIDIA Max-Q technology extends battery life by up to 40%, and includes thin and light laptops that maintain their sleek design without sacrificing power or performance.

NVIDIA DLSS 4 Boosts Performance by Up to 8x

DLSS 4 debuts Multi Frame Generation to boost frame rates by using AI to generate up to three frames per rendered frame. It works in unison with the suite of DLSS technologies to increase performance by up to 8x

over traditional rendering, while maintaining responsiveness with NVIDIA Reflex technology.

DLSS 4 also introduces the graphics industry's first real-time application of the transformer model architecture. Transformer-based DLSS Ray Reconstruction and Super Resolution models use 2x more parameters and 4x more compute to provide greater stability, reduced ghosting, higher details and enhanced anti-aliasing in game scenes. DLSS 4 will be supported on GeForce RTX 50 Series GPUs in over 75 games and applications the day of launch.

NVIDIA Reflex 2 introduces Frame Warp, an innovative technique to reduce latency in games by updating a rendered frame based on the latest mouse input just before it is sent to the display. Reflex 2 can reduce latency by up to 75%. This gives gamers a competitive edge in multiplayer games and makes single-player titles more responsive.

Blackwell Brings AI to Shaders

Twenty-five years ago, NVIDIA introduced GeForce 3 and programmable shaders, which set the stage for two decades of graphics innovation, from pixel shading to compute shading to real-time ray tracing. Alongside GeForce RTX 50 Series GPUs, NVIDIA is introducing RTX Neural Shaders, which brings small AI networks into programmable shaders, unlocking film-quality materials, lighting and more in real-time games.

Rendering game characters is one of the most challenging tasks in real-time graphics, as people are prone to notice the smallest errors or artifacts in digital humans. RTX Neural Faces takes a simple rasterized face and 3D pose data as input, and uses generative AI to render a temporally stable, high-quality digital face in real time.

RTX Neural Faces is complemented by new RTX technologies for ray-traced hair and skin. Along with the new RTX Mega Geometry, which enables up to 100x more ray-traced triangles in a scene, these advancements are poised to deliver a massive leap in realism for game characters and environments.

The power of neural rendering, DLSS 4 and the new DLSS transformer model is showcased on GeForce RTX 50 Series GPUs with Zorah, a groundbreaking new technology demo from NVIDIA.

Autonomous Game Characters

GeForce RTX 50 Series GPUs bring industry-leading AI TOPS to power autonomous game characters in parallel with game rendering.

NVIDIA is introducing a suite of new NVIDIA ACE technologies that enable game characters to perceive, plan and act like human players. ACE-powered autonomous characters are being integrated into KRAFTON's *PUBG: BATTLEGROUNDS* and *InZOI*, the publisher's upcoming life simulation game, as well as Wemade Next's *MIR5*.

In *PUBG*, companions powered by NVIDIA ACE plan and execute strategic actions, dynamically working with human players to ensure survival. *InZOI* features Smart Zoi characters that autonomously adjust behaviors based on life goals and in-game events. In *MIR5*, large language model (LLM)-driven raid bosses adapt tactics based on player behavior, creating more dynamic, challenging encounters.

AI Foundation Models for RTX AI PCs

Showcasing how RTX enthusiasts and developers can use NVIDIA NIM microservices to build AI agents and assistants, NVIDIA will release a pipeline of NIM microservices and AI Blueprints for RTX AI PCs from top model developers such as Black Forest Labs, Meta, Mistral and Stability AI.

Use cases span LLMs, vision language models, image generation, speech, embedding models for retrieval-augmented generation, PDF extraction and computer vision. The NIM microservices include all the necessary components for running AI on PCs and are optimized for deployment across all NVIDIA GPUs.

To demonstrate how enthusiasts and developers can use NIM to build AI agents and assistants, NVIDIA today previewed Project R2X, a vision-enabled PC avatar that can put information at a user's fingertips, assist with desktop apps and video conference calls, read and summarize documents, and more.

AI-Powered Tools for Creators

The GeForce RTX 50 Series GPUs supercharge creative workflows. RTX 50 Series GPUs are the first consumer GPUs to support FP4 precision, boosting AI image generation performance for models such as FLUX by 2x and enabling generative AI models to run locally in a smaller memory footprint, compared with previous-generation hardware.

The NVIDIA Broadcast app gains two AI-powered beta features for livestreamers: Studio Voice, which upgrades microphone audio, and Virtual Key Light, which relights faces for polished streams. Streamlabs is introducing the Intelligent Streaming Assistant, powered by NVIDIA ACE and Inworld AI, which acts as a cohost, producer and technical assistant to enhance livestreams.

Availability

For desktop users, the GeForce RTX 5090 GPU with 3,352 AI TOPS and the GeForce RTX 5080 GPU with 1,801 AI TOPS will be available on Jan. 30 at \$1,999 and \$999, respectively.

The GeForce RTX 5070 Ti GPU with 1,406 AI TOPS and GeForce RTX 5070 GPU with 988 AI TOPS will be available starting in February at \$749 and \$549, respectively.

The NVIDIA Founders Editions of the GeForce RTX 5090, RTX 5080 and RTX 5070 GPUs will be available directly from nvidia.com and select retailers worldwide.

Stock-clocked and factory-overclocked models will be available from top add-in card providers such as ASUS, Colorful, Gainward, GALAX, GIGABYTE, INNO3D, KFA2, MSI, Palit, PNY and ZOTAC, and in desktops from system builders including Falcon Northwest, Infiniarc, MAINGEAR, Mifcom, ORIGIN PC, PC Specialist and Scan Computers.

Laptops with GeForce RTX 5090, RTX 5080 and RTX 5070 Ti Laptop GPUs will be available starting in March, and RTX 5070 Laptop GPUs will be available starting in April from the world's top manufacturers, including Acer, ASUS, Dell, GIGABYTE, HP, Lenovo, MECHREVO, MSI and Razer.