Nano Banana 2: Google Embeds 4K Image Generation Directly Into Gemini

Kabui, Charles

2026-03-12

Google released Nano Banana 2 (officially Gemini 3.1 Flash Image), a native image generation model embedded directly in the Gemini architecture. It generates images from 512px to 4K resolution across 14 aspect ratios, at prices starting from $0.045 per image. New to this version: Google Image Search grounding, where the model searches the web for reference images before generating, and support for up to 14 reference images per workflow to maintain

subject consistency across scenes (up to 10 objects and 4 characters). A built-in "thinking" mode lets the model reason through complex compositions, producing interim draft images before the final render. The model handles text-to-image, conversational editing, style transfer, and inpainting (editing a specific region of an image by describing what to change) in a single multi-turn session. All outputs carry SynthID watermarks so AI-generated images can be identified.

At $0.067 per 1K image, Nano Banana 2 costs roughly half the $0.134 charged by its Pro-tier sibling (Gemini 3 Pro Image). A full 4K render runs $0.151. The image search grounding is exclusive to this model: prompt "a detailed painting of a Timareta butterfly" and it pulls real web images as visual context before generating. For designers and developers building at volume, this eliminates the separate reference-gathering step entirely.

Native multimodal generation, where one model handles text, code, and images without routing to separate systems, is becoming the default. Meta FAIR's research showed jointly trained models outperform bolt-on approaches, and Google is now shipping that principle at production scale.

Sources:

- Gemini 3.1 Flash Image Model Page
- Nano Banana Image Generation Documentation
- Gemini API Pricing
- SynthID by Google DeepMind