Google

Gemini 3.1 Flash Image Model Card

Gemini 3.1 Flash Image - Model Card

__Model Cards__ are intended to provide essential information on Gemini models, including known limitations, mitigation approaches, and safety performance. Model cards may be updated from time-to-time; for example, to include updated evaluations as the model is improved or revised. See the [Google DeepMind site](#) for a comprehensive list of model cards.

Published / Model Release: February 2026

Model Information

Description: Gemini 3.1 Flash Image is a member of the Gemini series of models, a suite of highly-capable, natively multimodal reasoning models. Gemini 3.1 Flash Image can comprehend input from different information sources, including text, images, audio and video. Image and text output is generated in the response.

Model dependencies: Gemini 3.1 Flash Image is based on Gemini 3 Flash.

Inputs: Text strings (e.g., a prompt, document(s)) and images, with a token context window of up to 1M.

Outputs: Image, with a 4K token output

Architecture:: Gemini 3.1 Flash Image is based on Gemini 3 Flash. For more information about the model architecture for Gemini 3 Flash, see the Gemini 3 Flash [model card](#).

Model Data

Training Dataset: Gemini 3.1 Flash Image is based on Gemini 3 Flash. For more information about the training dataset for Gemini 3 Flash, see the Gemini 3 Flash [model card](#).

Training Data Processing: For more information about the training data processing for Gemini 3.1 Flash Image, see the Gemini 3 Flash [model card](#).

Implementation and Sustainability

Hardware: Gemini 3.1 Flash Image was trained using [Google's Tensor Processing Units](#) (TPUs). TPUs are specifically designed to handle the massive computations involved in training LLMs and can speed up training considerably compared to CPUs. TPUs often come with large amounts of high-bandwidth memory, allowing for the handling of large models and batch sizes during training, which can lead to better model quality. TPU Pods (large clusters of TPUs) also provide a scalable solution for handling the growing complexity of large foundation models. Training can be distributed across multiple TPU devices for faster and more efficient processing.

The efficiencies gained through the use of TPUs are aligned with Google's [commitment to operate sustainably](#).

Software: Training was done using [JAX](#) and [ML Pathways](#).

Distribution

Gemini 3.1 Flash Image is based on Gemini 3 Flash. For more information about the distribution for Gemini 3 Flash, see the Gemini 3 Flash [model card](#).

Evaluation

The following Evaluation approach and results are for Gemini 3.1 Flash Image. For more information about the evaluation for Gemini 3 Flash, see the Gemini 3 Flash [model card](#).

Approach: Gemini 3.1 Flash Image was evaluated using the methodology below:

- **Human evaluations** of several different quality aspects of image generation were conducted. Prompt sets for human evaluations were in two broad categories:
    - **New capability sets**: diverse Text-to-Image (T2I) and editing evals are curated, covering a wide range of capabilities, e.g. Text, Style, Character, Reasoning, Factuality, Infographics, Ink (doodle) based Editing, Multi-Turn, Multi-Image (Multi-Product Recontextualization, Multi-Character, etc.)
    - **Regression sets**: popular use cases observed on Gemini 2.5 Flash Image and Gemini 3 Pro Image, to ensure Gemini 3.1 Flash Image does not show noticeable regression.
- **Capabilities / Benchmarks:**
    - T2I: General Text-to-Image, Text Rendering, Visual Design.
    - Editing: General Image Editing, Stylization, Character Editing, Object/Environment Editing, Factuality, Infographics, Ink (doodle) based Editing, Multi-Image (Multi-Product Recontextualization, Multi-Character, etc.), Multi-Turn.
    - Public benchmark: a subset of GenExam.
- **Eval Methodology**
    - SxS human eval to get Elo across diverse T2I, Editing, and Multi-Turn.
    - Single sided expert rating on subject-focused tasks on EDU tasks, GenExam.
    - AutoRater on factuality, style diversity of non-natural images.

Results: Results for Gemini 3.1 Flash Image are below.

Capabilities: Text-to-Image

Capability **Benchmark**	**Gemini 3.1 Flash Image** (Thinking + Text Search + Image Search)	**Gemini 3.1 Flash Image**	Gemini 2.5 Flash Image ("Nano Banana")	Gemini 3 Pro Image ("Nano Banana Pro")	GPT-Image 1.5	Seedream 5.0 Lite	Grok Imagine Image Pro
Overall Preference (GenAI-Bench)	1079.0 ± 7.0	1073.0 ± 5.0	942.0 ± 6.0	1021.0 ± 5.0	1047.0 ± 5.0	928.0 ± 8.0	906.0 ± 6.0
Visual Quality (GenAI-Bench)	1140.0 ± 6.0	1129.0 ± 6.0	929.0 ± 6.0	1043.0 ± 5.0	975.0 ± 5.0	759.0 ± 10.0	953.0 ± 5.0
Infographics (Factuality)	1114.0 ± 14.0	1074.0 ± 12.0	881.0 ± 13.0	1102.0 ± 13.0	985.0 ± 12.0	890.0 ± 22.0	942.0 ± 21.0

Capabilities: Editing

Capability Benchmark	Gemini 3.1 Flash Image (Thinking + Text Search + Image Search	Gemini 3.1 Flash Image	Gemini 2.5 Flash Image ("Nano Banana")	Gemini 3 Pro Image ("Nano Banana Pro")	GPT-Image 1.5	Seedream 5.0 Lite	Grok Imagine Image Pro
General Editing	1065.0 ± 9.0	1047.0 ± 9.0	913.0 ± 9.0	1051.0 ± 10.0	995.0 ± 8.0	937.0 ± 9.0	989.0 ± 8.0
Character Editing	1056.0 ± 7.0	1049.0 ± 7.0	952.0 ± 7.0	1050.0 ± 8.0	1025.0 ± 7.0	894.0 ± 8.0	972.0 ± 7.0
Creative	1023.0 ± 7.0	1031.0 ± 7.0	976.0 ± 7.0	1004.0 ± 7.0	1017.0 ± 7.0	938.0 ± 7.0	1016.0 ± 7.0
Object/Environment Editing	1029.0 ± 8.0	1018.0 ± 8.0	945.0 ± 8.0	1042.0 ± 10.0	976.0 ± 8.0	946.0 ± 9.0	1022.0 ± 8.0
Multi-Input (1-3)	1037.0 ± 8.0	1016.0 ± 8.0	919.0 ± 9.0	1056.0 ± 12.0	1014.0 ± 9.0	951.0 ± 9.0	N/A
Stylization	1045.0 ± 7.0	1031.0 ± 7.0	862.0 ± 8.0	1045.0 ± 9.0	996.0 ± 7.0	984.0 ± 7.0	1021.0 ± 8.0

Intended Usage and Limitations

Benefit and Intended Usage: Gemini 3.1 Flash Image is capable of using Gemini's real-world knowledge to deliver precise results and reflect the world around you, from complex infographics to historically accurate scenes. It is well-suited for applications that require:

- creation and editing of images with professional levels of precision and control and multiple, quick iterations
- generation of clear text for posters and intricate diagrams
- long context real-world knowledge
- localized text rendering across several languages
- studio-quality control

Known Limitations: Gemini 3.1 Flash Image may exhibit some of the general limitations of foundation models, such as hallucinations. There may also be occasional slowness or timeout issues.

Gemini 3.1 Flash Image still has room for several quality improvements:

- Text rendering: poor in small text (often blurry in 1k model), long paragraphs, page length
- Character consistency is not always perfect between input images and generated output image
- Masked/Doodle based editing: partial instruction following and persistent ink
- When editing images: infrequent copying/pasting from user's input image to generated image
- Occasional confusion around spatial localisation (e.g. left/right etc.)
- Still limited in advanced capabilities with world knowledge, 3D reasoning and factuality

The knowledge cutoff date for Gemini 3.1 Flash Image was January 2025.

Acceptable Usage: For more information about the acceptable usage for Gemini 3.1 Flash Image, see the Gemini 3 Flash model card.

Ethics and Content Safety

Evaluation Approach: Gemini 3.1 Flash Image was developed in partnership with internal safety, and responsibility teams. A range of evaluations and red teaming activities were conducted to help improve the model and inform decision-making. These evaluations and activities align with Google's AI Principles and responsible AI approach, as well as Google's Generative AI policies (e.g. Gen AI Prohibited Use Policy and the Gemini API Additional Terms of Service). As Gemini 3.1 Flash Image is based on Gemini 3 Flash, see the Gemini 3 Flash model card for additional Ethics & Content Safety details.

Evaluation types included but were not limited to:

- **Training/Development Evaluations** including automated and human evaluations carried out continuously throughout and after the model's training, to monitor its progress and performance;
- **Human Red Teaming** conducted by specialist teams across the policies and desiderata, deliberately trying to spot weaknesses and ensure the model adheres to safety policies and desired outcomes;
- **Trust Assurance Evaluations** conducted by evaluators who sit outside of the model development team, used to independently assess responsibility and safety governance decisions;
- **Ethics & Safety Reviews** were conducted ahead of the model's release.

Safety Policies: Gemini's safety policies are based on Google's standard framework, which aim to prevent our Generative AI models from generating harmful content, including:

1. Content related to child sexual abuse material and exploitation
2. Hate speech (e.g. dehumanizing members of protected groups)
3. Dangerous content (e.g. promoting suicide, or instructing in activities that could cause real-world harm)
4. Harassment (e.g. encouraging violence against people)
5. Sexually explicit content
6. Medical advice that runs contrary to scientific or medical consensus

We continue to improve our internal evaluations, including refining automated evaluations to reduce false positives and negatives, as well as update query sets to ensure balance and maintain a high standard of results.

Frontier Safety Assessment: Gemini 3.1 Flash Image is part of the Gemini 3 family of models. We evaluated Gemini 3.1 Pro for frontier safety as it was the most generally capable model as of publication of this model card, and it did not reach any Critical Capability Levels (CCLs) outlined in our Frontier Safety Framework. Our assessments have shown that Gemini 3.1 Flash Image is likely to perform similarly to Gemini 3.1 Pro, therefore based on Gemini 3.1 Pro, we are confident that that Gemini 3.1 Flash Image is also unlikely to reach any CCLs. For more information, read the Gemini 3.1 Pro Model Card.

Risks and Mitigations: Safety and responsibility was built into Gemini 3.1 Flash Image throughout the training and deployment lifecycle, including pre-training, post-training, and product-level mitigations. Mitigations include, but are not limited to:

- dataset filtering;
- conditional pre-training;
- supervised fine-tuning;
- reinforcement learning from human and critic feedback;
- safety policies and desiderata;
- product-level mitigations such as safety filtering.