

GPT Image 2 と Nano Banana 2 の比較調査報告

ChatGPT-5.5

エグゼクティブサマリ

本稿では「GPT Image 2.0」を、gpt-image-2 モデルページと 示される最新世代の画像モデルとして扱い、「Nano Banana 2」を、gemini-3.1-flash-image-preview モデルページ 12search5 と 示される高効率画像モデルとして扱う。結論から言うと、文字・レイアウト・編集の確実性を重視するなら **GPT Image 2**、遅延・大量反復・検索接地を重視するなら **Nano Banana 2** である。盲検のユーザー選好ベースでは GPT Image 2 が新規生成と画像編集の両方で上位にあり、文字タイポグラフィのランキングでも上にいる。一方、Nano Banana 2 は高効率モデルとして位置付けられ、検索接地、長文テキスト、極端アスペクト比、反復の速さで実務上の優位がある。[1]

コストは「どちらが安い」を一言で決めにくい。GPT Image 2 は 1024×1024 の 1 枚あたりで low \$0.006 / medium \$0.053 / high \$0.211 と品質段階が広く、Nano Banana 2 は 1K / 2K / 4K で \$0.067 / \$0.101 / \$0.15 という比較的わかりやすい体系になっている。したがって、大量の下書きでは GPT Image 2 の low が非常に強く、高解像度の量産では Nano Banana 2 の価格見通しのほうが立てやすい。[2]

ガバナンス面では、両者ともクローズドなクラウド提供であり、学習データの詳細な構成、重み、ローカル推論向け GPU 要件は公開されていない。本調査範囲で確認できた一次資料では、どちらも画像モデル自体のファインチューニング / LoRA は公式提供がなく、商用利用は可能だが、前者は出力所有権を比較的明確に示し、後者は所有権を主張しない一方で同一・類似出力の可能性や法令順守責任を強く残している。[3]

比較表

比較軸	GPT Image 2	Nano Banana 2	根拠
正式名称 / API ID	ChatGPT Images 2.0 / gpt-image-2。スナップショットあり	Nano Banana 2 / gemini-3.1-flash-image-preview。Gemini 3系の preview 扱い	[5]
提供チャンネル	ChatGPT[6]、Image API、Responses API、Batch	Gemini[7]、AI Studio、Gemini API、Vertex AI	[8]
位置づけ	「state-of-the-art」の高品質生成・編集モデル	Gemini 3 Pro Image の「high-volume / high-efficiency / lower price-point」相当	[9]
入出力	text in、image in/out。高忠実度編集、参照画像、複数画像生成	text / image / PDF in、image / text out。会話型編集に強い	[10]
解像度・サイズ制御	任意サイズ可。最大 edge 3840px、比率 3:1 まで。2K 超は experimental。品質は low / medium / high / auto	image_size は 0.5K / 2K / 4K、出力縦横比は 1:4、4:1、1:8、8:1 まで。16:9 などネイティブ	[11]
生成品質の方向性	総合選好、編集、複雑な視覚テキスト、タイポグラフィで優位	長文テキスト、dense prompt alignment、検索接地込みの実在物表現で健闘	[12]
速度	quality:"low" が最速。正方形が速く、JPEG は PNG より速い。複雑な prompt は最大 2 分	公式に low latency / high-volume を強調。独立レビューでも「seconds」級と評価	[13]
価格	1024×1024 で low \$0.006 / medium \$0.053 / high \$0.211	1K \$0.067 / 2K \$0.101 / 4K \$0.15	[2]
カスタマイズ	size / quality / format / compression。Responses API で multi-turn 編集。fine-tuning	image_size / aspect_ratio / Search grounding / thought signatures。公式 SDK が thought signatures を自	[14]

比較軸	GPT Image 2	Nano Banana 2	根拠
	非対応。スナップショット固定可	動処理。画像モデルの supervised fine-tuning 記載なし	
安全性	入力 prompt・入力画像・出力画像の多層分類、moderation パラメータ、C2PA と不可視 watermark、公表 system card	広範な filtering / data labeling / red teaming、child safety 評価、SynthID watermark、利用規約上の安全設定義務	[15]
データ利用・プライバシー	API 入力は既定で学習に使わない。abuse logs は通常 30 日以内。承認制で zero data retention あり	unpaid services では入力と出力が製品改善に使われ、人手レビューもありうる。有償 tier は価格表で “Used to improve our products: No”	[16]
商用利用	利用者が output を所有。ただし非一意性あり	Google は generated content の所有権を主張しないが、同様出力の可能性と法令順守・必要に応じた attribution を明示	[17]
学習データ公開状況	未公開 / 不明。能力説明はあるが、データセット一覧・比率・権利処理の詳細は本調査の一次資料では確認できず	未公開 / 不明。knowledge cutoff は示すが、学習データセットの内訳は確認できず	[18]
推論時ハードウェア要件	ベンダー管理 API。ローカル実行用 weight 非公開のため、自前 GPU 要件は未公開 / 実質不要	同左。完全マネージド前提で、ローカル推論向け仕様は未公開	[19]
用途別推奨			
シナリオ	推奨	理由	根拠

シナリオ	推奨	理由	根拠
商用 広告	第一候補は GPT Image 2。 ただし、実在物の情報を検索で確認しながら図解・旅行・地域訴求を作るなら Nano Banana 2 かハイブリッド	GPT Image 2 は Text-to-Image / Editing / Text & Typography で強く、複雑な視覚テキスト生成でも優勢。Nano Banana 2 は web search を使った grounded generation と 4K/横長アセット生成が強い	[20]
アー ト制 作	世界観探索は Nano Banana 2、最終ポスターや誌面化は GPT Image 2	独立レビューでは Nano Banana 2 が背景情報量や“絵としての密度”で勝つ場面があり、OpenAI 系は論理整合や文字・誌面化で強い。つまり「着想は Nano、仕上げは GPT」が合理的	[21]
研究 / プ ロト タイ ピン グ	再現性と A/B 比較重視なら GPT Image 2、スループットと Google stack 連携重視なら Nano Banana 2	GPT Image 2 は snapshots と Image API / Responses API の二枚看板で実験設計しやすい。Nano Banana 2 は high-volume で、OpenAI compatibility と search grounding が導入摩擦を下げる	[22]
大量 バリ エー ショ ン生 成	Nano Banana 2 優勢、ただし超低コスト下書きなら GPT Image 2 low も有力	Nano Banana 2 は高効率・低遅延を前提に設計され、独立レビューでも秒単位の反復性が評価される。一方、GPT Image 2 low は単価だけ見ると非常に安い	[23]

実務で最も現実的なのは、**Nano Banana 2** で素早く構図・情報・バリエーションを出し、**GPT Image 2** でタイポ・レイアウト・編集の完成度を上げる二段構えである。どちらか一方に統一するなら、クライアント向けの最終成果物が主目的なら

GPT、内製の高速反復や検索接地型の図解・広告試作が主目的なら Nano が向く。
[24]

技術的詳細

生成品質

品質を一つの数字で決めるのは危険だが、総合傾向はかなり見える。盲検のユーザー選好では、GPT Image 2 が Text-to-Image で Elo 1338、Nano Banana 2 が 1262、Image Editing では GPT Image 2 が 1252、Nano Banana 2 が 1230 で、2026-05 時点では GPT Image 2 が両カテゴリで上にいる。Text & Typography の主題別リーダーボードでも GPT Image 2 は 2 位、Nano Banana 2 は 6 位で、文字組みを含むグラフィックワークでは GPT 側が一段強い。[25]

xychart-beta

```
title "盲検評価ベースの相対品質"  
x-axis ["Text-to-Image Elo", "Image Editing Elo"]  
y-axis "Elo" 1200 --> 1360  
bar "GPT Image 2" [1338,1252]  
bar "Nano Banana 2" [1262,1230]
```

上図は blind preference ベースの Elo であり、**Text-to-Image** と **Editing** は別アリーナなので、カテゴリ内の相対比較として読むべきである。それでも「新規生成でも編集でも GPT Image 2 のほうが好まれやすい」という方向性は明瞭だ。[26]

より細かい公開ベンチマークでは、別のニュアンスが出る。HiDream-O1-Image の公開表では、GenEval は GPT Image 2 が 0.89、Nano Banana 2 が 0.83、HPSv3 は 10.21 対 10.01、CVTG-2K の complex visual text generation の平均は 0.9003 対 0.7875、NED は 0.9515 対 0.8945、CLIP Score は 0.7798 対 0.7212 で、**複雑な視覚テキスト、広告、UI、ポスター、図表**では GPT Image 2 の明確な強みが見える。
[27]

ただし、Nano Banana 2 が弱いわけではない。同じ公開表では DPG-Bench の overall が Nano Banana 2 86.90、GPT Image 2 85.98 で、dense prompt alignment では Nano 側がわずかに上回る。LongText-Bench でも EN/ZH で Nano Banana 2 が 0.980 / 0.965、GPT Image 2 が 0.960 / 0.961 で、**長い文字列の連続レンダリング**や

多言語長文では Nano 側がやや有利だ。要するに、GPT は「誌面・タイポ・整った成果物」、Nano は「長文・条件密度・情報量の多い返し」が得意、という整理が最も実務に近い。[28]

独立レビューもこの傾向と整合する。およびは、Nano Banana 2 が背景・前景の情報量や“物語性”で勝つ場面を指摘する一方、ChatGPT Images 2.0 系の比較では「Google の realism に対し、ChatGPT の logic が勝った」というフレーミングも出ている。したがって、絵としての濃さと指示の論理整合は別軸で見るべきだ。[29]

速度とコスト

速度は、前者が「調整可能な品質段階」、後者が「高効率 Flash 系」という設計思想の違いとして表れている。GPT Image 2 は quality: "low" が最速で、正方形が速く、JPEG は PNG より速い。反面、OpenAI の画像ガイドは複雑な prompt で最大 2 分かかる可能性を明記している。対して Nano Banana 2 は、公式モデル説明でも release notes でも一貫して low latency / high-volume を押し出している。独立レビューでも両者とも「seconds」級だが、高頻度の繰り返しでは Nano のほうが有利という評価が多い。[13]

コストは以下のように読める。GPT Image 2 は 1024×1024 で low \$0.006 / medium \$0.053 / high \$0.211、Nano Banana 2 は 1K \$0.067 / 2K \$0.101 / 4K \$0.15 である。1 万枚の単純試算にすると、GPT Image 2 は \$60 / \$530 / \$2,110、Nano Banana 2 は \$670 / \$1,010 / \$1,500 になる。したがって、草案を大量に出す用途では GPT low が非常に安く、高解像度の production-ready 画像を高頻度で回す用途では Nano Banana 2 の単価体系が扱いやすい。なお、これらは主に画像出力料金であり、GPT 側の入力 text / image token、Nano 側の text input や platform 差異は別途の。[2]

制御性とカスタマイズ性

GPT Image 2 の制御面で優れているのは、size / quality / format / compression を細かく持ちながら、Responses API で multi-turn 編集、file ID 入力、複数参照画像入力まで備えていることだ。さらに tool 呼び出し型では revised prompt が返るので

、なぜその画像が出たかを分析しやすい。一方で、モデル自体の fine-tuning は非対応で、LoRA 的な個別適応も公式には出ていない。[30]

Nano Banana 2 の制御面の特徴は、aspect ratio の広さ、Google Search grounding、会話型編集に必須の thought signatures、そして SDK 自動処理にある。特に gemini-3.1-flash-image-preview は google_search と組み合わせた grounded generation が公式に推されており、事実確認を前提とするインフォグラフィックや旅行・ローカライズ文脈では強い。また、Gemini docs は OpenAI compatibility を明示しており、既存の OpenAI 流儀のクライアントから移行しやすい。反面、画像モデルそのものに対する supervised fine-tuning の記載は、本調査範囲の一次資料では確認できなかった。[31]

導入・運用・ガバナンス

導入のしやすさは、単に SDK があるかよりも、**運用時の癖**で決まる。GPT Image 2 は Image API と Responses API の役割分担が明快で、「一枚もの」は前者、「会話型編集」は後者という整理がしやすい。さらに snapshots により再現性を固定しやすく、研究・評価・回帰テストには向く。ただし、組織によっては API Organization Verification が必要になる。[32]

Nano Banana 2 は、Gemini app、AI Studio、Gemini API、Vertex AI と提供面が広く、企業導入では Vertex AI 側の導線がはっきりしている。さらに、Google Cloud 日本語ブログは、Nano Banana 2 を広告、ストーリーボード、プロダクトモックアップ、ローカライズマーケティング、旅行アプリなどのワークフローに自然統合することを強調している。顧客事例も複数掲載されており、画像編集スイート、デザインツール、ノート/ドキュメント系アプリ、ファッション用途、広告代理店ワークフローなど、“**既存アプリに埋め込む**”文脈が強い。[33]

前者の透明性で特筆すべきは、公開 system card があり、安全スタック、biorisk 対策、C2PA と不可視 watermark の両方をかなり詳しく説明している点だ。後者も安全性の説明はあるが、本調査で見つかった一次資料の範囲では、Nano Banana 2 単

体について同等の詳細度の system card は確認できなかった。透明性重視の審査・稟議では、ここは前者の優位点になる。[34]

データ利用ポリシーは実務上かなり重要だ。OpenAI API は 2023-03-01 以降、明示的 opt-in が無い限り API データを学習・改善に使わず、abuse monitoring logs は通常 30 日までとしている。対照的に Gemini API / AI Studio 側は、unpaid services では入力・出力が製品改善に使われ、人手レビューもありうる一方、価格表では有償 tier について “Used to improve our products: No” が示される。したがって、**機密情報**や未公開クリエイティブを扱う場合、**Nano Banana 2** は有償面に寄せる前提で見たほうが安全だ。[16]

未公開事項と評価上の限界

- 両モデルとも、**FID / IS** の公式値は確認できなかった。画像内文字やレイアウトを重視する現行モデルでは、研究側も FID・CLIP・OCR だけでは能力差を十分に捉えにくいと指摘しており、今回は Arena Elo、CVTG、LongText、HPSv3 をより重視した。[35]
- 学習データの実リスト、権利処理、比率、重み、蒸留の詳細は未公開 / 不明である。「未公開」は能力が低いという意味ではなく、透明性の限界を指す。[18]
- HiDream-O1-Image のベンチマーク表は便利だが、**競合モデル**作者側が公開した評価であり、完全に中立な第三者ラボとは言い切れない。独立評価としては Artificial Analysis やレビュー記事と合わせて読むのが妥当だ。[36]
- 価格は surface 差異がある。GPT 側は input text / input image token が別途のり、Nano 側も Gemini API / Vertex AI / standard / batch で見え方が異なる。したがって、**最終判断**は自社の **prompt 長・編集回数・解像度・再試行率**で再計算すべきだ。[2]

主要参考リンクとサンプル

日本語の一次情報

独立レビューとベンチマーク

- [Artificial Analysis Text & Typography Leaderboard](https://artificialanalysis.ai/image/leaderboard/text-to-image)

代表的な公式サンプル画像

実装リポジトリと SDK

- [openai-python](https://github.com/openai/openai-python)
- [openai-node](https://github.com/openai/openai-node)
- [python-genai](https://github.com/googleapis/python-genai)
- [js-genai](https://github.com/googleapis/js-genai)

結論

2026-05 時点での実務判断は、「どちらが絶対に上か」ではなく、「どの失敗を嫌うか」で決まる。文字崩れ、誌面崩れ、編集の不安定さ、クライアント提出前の手戻りを嫌うなら GPT Image 2 が安全な本命である。逆に、短いサイクルで大量に試し、実在情報を検索で補い、4K や極端アスペクト比を素早く回したいなら Nano Banana 2 が強い。盲検選好と編集品質では GPT、長文テキストと高効率運用では Nano、という整理が最も再現性の高い結論だった。[37]

単一選定の推奨を一つだけ挙げるなら、**商用広告・販促・図表入り資料を最終成果物として出す組織には GPT Image 2、高速プロトタイピング・アプリ組み込み・検索接地型の生成を主軸にする組織には Nano Banana 2** を勧める。最終的な費用対効果まで含めた実戦解は、Nano で構図と情報を回し、GPT で仕上げるハイブリッド運用である。[38]

[1] [12] [20] [24] [25] [26] [37] <https://artificialanalysis.ai/image/leaderboard/text-to-image>

<https://artificialanalysis.ai/image/leaderboard/text-to-image>

[2] [6] [7] [11] [13] [14] [19] [30] [32] [38]

<https://developers.openai.com/api/docs/guides/image-generation>

<https://developers.openai.com/api/docs/guides/image-generation>

[3] [5] [9] [10] [18] [22] <https://developers.openai.com/api/docs/models/gpt-image-2>

<https://developers.openai.com/api/docs/models/gpt-image-2>

[4] <https://openai.com/ja-JP/index/introducing-chatgpt-images-2-0/>

<https://openai.com/ja-JP/index/introducing-chatgpt-images-2-0/>

[8] <https://help.openai.com/en/articles/6825453-chatgpt-release-notes>

<https://help.openai.com/en/articles/6825453-chatgpt-release-notes>

[15] [34] <https://deploymentsafety.openai.com/chatgpt-images-2-0>

<https://deploymentsafety.openai.com/chatgpt-images-2-0>

[16] <https://developers.openai.com/api/docs/guides/your-data>

<https://developers.openai.com/api/docs/guides/your-data>

[17] <https://openai.com/policies/row-terms-of-use/>

<https://openai.com/policies/row-terms-of-use/>

[21] [29] <https://www.tomsguide.com/ai/chatgpt-5-2-vs-nano-banana-2-i-tested-both-to-find-the-better-ai-image-generator>

<https://www.tomsguide.com/ai/chatgpt-5-2-vs-nano-banana-2-i-tested-both-to-find-the-better-ai-image-generator>

[23] <https://ai.google.dev/gemini-api/docs/models/gemini-3.1-flash-image-preview>

<https://ai.google.dev/gemini-api/docs/models/gemini-3.1-flash-image-preview>

[27] [28] [36] <https://github.com/HiDream-ai/HiDream-O1-Image>

<https://github.com/HiDream-ai/HiDream-O1-Image>

[31] <https://ai.google.dev/gemini-api/docs/interactions/gemini-3>

<https://ai.google.dev/gemini-api/docs/interactions/gemini-3>

[33] <https://cloud.google.com/blog/ja/products/ai-machine-learning/bringing-nano-banana-2-to-enterprise>

<https://cloud.google.com/blog/ja/products/ai-machine-learning/bringing-nano-banana-2-to-enterprise>

[35] <https://www.mdpi.com/2076-3417/15/5/2274>

<https://www.mdpi.com/2076-3417/15/5/2274>