

Google Gemini Nano Banana 图像的创意应用案例

📅 2025年9月6日 ⌚ 4 分钟阅读

#image #Gemini #Nano #Banana

本文介绍Google Gemini Nano Banana模型的创意应用案例。

Google Gemini Nano Banana（正式名称为Gemini 2.5 Flash Image）是Google DeepMind发布的最先进图像生成和编辑模型，其创意应用案例令人惊叹，已经在多个领域展现出革命性的潜力。[1][2][3]

专业图像编辑与创意制作

虚拟试衣体验

Nano Banana在虚拟试衣领域表现卓越。用户只需上传自己的照片和想要尝试的服装图像，AI就能生成逼真的试衣效果，保持面部特征和身体比例完全一致。这项技术已经被应用到Chrome扩展程序中，让用户在在线购物网站上直接进行虚拟试衣，大大降低了在线购物的退货率。[4][5][6]

产品摄影革命

在商业摄影领域，Nano Banana彻底改变了传统产品拍摄模式。品牌可以将简单的产品图像转化为专业的广告级视觉效果，无需昂贵的摄影工作室。例如，用户可以将一个普通的香水瓶图像转换为“放置在大理石台面上，柔和聚光灯照射的工作室级产品照片”。这种能力使得小企业也能制作出高质量的产品营销材料。[7][8]

室内设计与建筑可视化

Nano Banana在建筑和室内设计领域展现出强大的应用潜力。建筑师可以快速将3D模型视图转换为逼真的渲染图，或者通过简单的文本指令修改建筑材料、照明条件和环境设置。例如，上传一张客厅照片后，可以通过指令“将蓝色沙发换成棕色皮质沙发，保持其他所有元素不变”来快速预览设计效果。[9][10][11]

创意内容生成与品牌营销

多图像融合与风格迁移

该模型具备强大的多图像融合能力。用户可以将不同照片中的元素无缝合成到一张图像中，比如将一个人物、一部手机和一个地下通道的图像融合成“女性在地下通道中通话”的自然场景。这种功能在创意设计和广告制作中极为有用。[12]

目录

文章信息

字数

阅读时间

发布时间

更新时间

标签

#image #Gemini

角色一致性维护

Nano Banana的一大突破是能够在不同场景中保持角色一致性。无论是将同一个人放在不同的职业场景中（棋手、赛车手、射箭运动员），还是创建IP角色的系列形象，都能确保人物特征保持一致。这对于品牌营销和内容创作具有重要价值。^[12]

创意广告制作

营销专业人士可以利用Nano Banana创建一致性的品牌资产和营销材料。例如，通过上传产品图像，可以生成各种风格的广告场景，从简约的电商展示到奢华的品牌推广，所有这些都能在几分钟内完成。^[1]^[7]

教育与专业应用

医学教育支持

研究表明，Nano Banana在医学教育领域也展现出卓越性能。在解剖学多选题测试中，该模型达到了95.9%的准确率，表现优于其他AI平台。这表明Nano Banana不仅在图像生成方面出色，在知识理解和应用方面也具有很高的价值。^[13]

教育内容创建

教育工作者可以使用Nano Banana创建视觉教学材料、信息图表和互动内容。该模型能够将手绘图表转换为专业的教学视觉，帮助学生更好地理解复杂概念。^[14]

技术特点与创新优势

自然语言编辑

用户无需专业的图像编辑技能，只需用自然语言描述想要的修改即可。例如“给这张照片中的猫戴上一顶小巫师帽”，AI就能准确理解并执行编辑任务。^[15]^[1]

多轮对话编辑

Nano Banana支持渐进式编辑，用户可以在一个对话中持续完善图像。每一轮反馈都基于之前的修改，让用户能够精确达到想要的效果，这种迭代能力在传统图像编辑中很难实现。^[9]

世界知识融合

与传统图像生成模型不同，Nano Banana融合了Gemini的世界知识，能够理解真实世界的语义信息。这使得它能够处理更复杂的创意任务，如将产品置于时代广场的广告牌上，或创建电影级别的场景。^[2]^[16]

商业应用前景

电商自动化

已有开发者创建了自动化工作流程，能够为WooCommerce商店批量生成AI试衣图像。这种自动化大大降低了电商企业的视觉内容制作成本，提高了运营效率。[17]

创意产业革命

在创意产业中，Nano Banana被认为是"Photoshop的终结者"，因为它能够完成许多传统图像编辑软件无法实现的任务。从故事板制作到漫画创作，从社交媒体视觉到YouTube缩略图，该模型为创意工作者提供了前所未有的工具。[18] [19]

未来发展方向

随着Nano Banana技术的不断成熟，预期将在更多领域产生深远影响。从建筑可视化到虚拟现实内容创建，从个性化广告到沉浸式教育体验，这项技术正在重新定义我们与视觉内容交互的方式。

然而，正如专家指出的，这种强大的编辑能力也带来了新的挑战，特别是在图像真实性和信任度方面。因此，在享受技术便利的同时，我们也需要建立相应的识别和验证机制，确保这项技术的负责任使用。[20]

Demo

Nano Banana Hackathon

[Kaggle上的Banana参赛作品](#) 由Google Gemini团队发起，发起的挑战是：不要仅仅停留在简单的演示阶段，而是利用这项独特的技术打造一款真正令人惊叹的产品。开发一款能够彻底改变我们与视觉内容互动方式的应用程序，无论是用于动态叙事、彻底革新电子商务，还是构建下一代创意工具。最终48小时内就收到了800多份投稿，现已开始评审流程。预计将在接下来的几周内讨论区公布获奖者名单。

我从[获奖demo](#)里面整理了一些我感兴趣的罗列在这里

Banana Kaggle Winners

1. Simplifying Group Photos with Post Compose 📸

把几个人一起放到一个景点里。==》

<https://www.kaggle.com/competitions/banana/writeups/simplifying-group-photos-with-post-compose>

我的idea, 把人物(全身)和风景(几张图)发给程序, 程序给出建议如何取景, 如何摆post, 如何拍照布局(直接给出一个图片例子效果图, 比如, 背景+人物post, 全身/半身, 光线, 外饰(墨镜, 帽子, 围巾, 耳环)等。)

2. SingleShot3D: Digital twin from a single image

一种智能工具, 它能把产品照片转换为 3D 模型, 其动力源于一种新颖的单元调用生成流程。

注: 提供完整代码:

<https://aistudio.google.com/apps/drive/1kZEUxzHDrwyun4vAqTpZ0yg2CdbqBAXe?showAssistant=true&showCode=true>

3. ForgeOne

<https://www.kaggle.com/competitions/banana/writeups/forgeone-pro-mode-revolution>

ForgeOne 是一个面向专业创作的闭环AI生产系统, 把单一指令自动拆解为项目级任务(编排), 用行业化增强提示驱动专业生成(执行), 再以多模态量化评分自检并通过差分提示自动修复(验证与纠正), 依托 Nano Banana 的连续性能力在多轮修改中保持风格、角色与构图一致, 最终以更低修改成本、更高可复现性与交付确定性, 将“创意到客户验收”的全过程自动化。

没有代码和demo

4. Nano Studio

<https://www.kaggle.com/competitions/banana/writeups/nanostudio>

电商产品图生成与编辑工具: 它通过参考图一致性、身份保真型提示工程与自然语言编辑(Magic Edit), 在保持产品形状、材质、配色与logo不变的前提下, 快速生成多场景、高一致性的营销级 mockup; 配套模板系统(15+专业预设)、动态场景与批处理策略, 实现多比例输出、版本回溯、配额监控与高质量导出, 前端基于 Next.js + React + Tailwind, 图像与元数据存于 Turso, 部署于 Vercel, 显著降低拍摄成本、缩短制作周期并保障品牌一致性, 适用于电商、营销团队与代理商的大规模素材生产。脑洞扩展: 后续可加“模板市场 + 类目感知场景推荐 + 电商平台直连 + 目录级批量生成”, 甚至引入RL优化的“视觉一致性奖励”与小样本产品冷启, 以实现测试时扩展与更强的品牌守护。

Sozofix, AI DIY assistant

<https://www.kaggle.com/competitions/banana/writeups/sozofix-ai-diy-assistant>

“SozoFix”的AI DIY助手项目, 核心是帮用户把“能修就修、能再用就再用、能改造就改造”变成一件有把握、可执行的事。它把多模态感知、对话、TTS、记忆和实验性AR整合到一个端到端应用里, 面向的是“看着坏物件不知道从哪下手”的普通用户与创客。

BeastBlend

<https://www.kaggle.com/competitions/banana/writeups/beastblend-consistent-animal-mash-ups-infinite-sto>

BeastBlend 是一款面向教育场景的动态故事生成应用, 依托 Gemini 2.5 Flash Image 的三大能力: 角色一致性确保同一混合动物在4-10个分镜中外观细节稳定; 图像融合能够基于两种动物的解剖特征创造生动可信的“混种”角色; 语境化场景生成让角色在森林、海底、教室等多样环境中保持身份同时自适应姿态与表情。它为教师与特殊教育场景提供可定制的、连贯的图文与有声故事, 增强儿童情感连接与理解力, 并支持学生参与创作, 快速产出高质量教学素材。脑洞延展: 引入学习目标驱动的剧情生成与分支交互机制, 将词汇/技能点嵌入情节, 形成长期陪伴式角色成长。

Time Travel Map

<https://www.kaggle.com/competitions/banana/writeups/time-travel-map>

Time Travel Map 是一款“时空打点”应用: 你在地球任意位置落一个点并选择年份, 系统基于 Gemini 2.5 Flash Image 的地理与历史理解生成对应时代的逼真场景图像, 同时用 ElevenLabs Sound Effects 合成匹配的环境音(如1920年代的马车与街市喧闹、17世纪乡野的风声与鸟鸣), 实现图像+声音的沉浸式时空回溯; 其亮点在于地理定位与历史语境的自动推断、跨模态同步生成、面向教育/旅行/本地史探索等多场景易用性, 以及对未来年份的前瞻式可视化想象能力。脑洞延展: 引入多源史料与旧照片

对齐的可证伪模式、年代滑杆的连续插值动画、以及群体协作的“时空注释层”，让用户共同纠错与丰富本地时间脉络。

<https://www.kaggle.com/competitions/banana/writeups/comic-book-ai>

源代码：<https://github.com/CheickDiakite-yikes/comicbook-ai>（使用codeX）

Comic Book AI (Nerrame) 是一款“带故事大脑的漫画版Canva”：通过“Story Bible（角色/设定/语气/画风）”一次性定义世界观，依托 Gemini 2.5 Flash Image 的上下文感知生成，实现跨页、跨分镜的人物一致性与风格连贯；支持按脚本一键整页生成或逐格定制，提供智能上下文编辑、角色资产库复用、公共探索与分享、隐私切换、配额与自动保存等完整产品化能力。其亮点在于把复杂的漫画生产管线抽象为结构化语义驱动的可控生成流程，兼顾专业级一致性（角色/场景/画风）与大众创作门槛的极致降低，并通过动态分镜编排与环境语境生成确保叙事连贯与视觉统一。脑洞延展：加入“镜头语言编排器”（景别/运镜/节奏自动化）、跨页节奏曲线与重复元素回钩提示、以及基于读者反馈的在线A/B分镜优化，形成“剧本-分镜-版面-读者数据”的闭环自适应创作系统。

ArtLens: See What the Painter Saw

<https://www.kaggle.com/competitions/banana/writeups/artlens-see-what-the-painter-saw>

代码：https://github.com/markgreenfield00/artlens-see-what-the-painter-saw/blob/main/artist_image.py #简单的地吗，看prompt和gemini api交互。

ArtLens 是一款将名画“反推”为现实场景的交互式应用：用户上传绘画后，系统基于 Gemini 2.5 Flash Image Preview 的多模态能力，结合“图像输入+写实提示”在一次生成中输出多张逼真的候选照片（如3张），把画作中的构图、物体与地平线等线索重建为可能的真实景观，从而直观揭示艺术家如何由现实观察跃迁为艺术表达；其亮点在于以教育为核心（如将梵高、莫奈、达利利用地理与环境锚定），强调跨风格的写实重构而非风格复刻，并通过严格的候选采样与可复现实验流程（Pillow预处理、google-genai SDK、inline_data 解码）确保一致性与可验证性。脑洞延展：加入“历史地理校准”（与博物馆藏品元数据、旧地图对齐）、“视点逆推”（估计画家站位与焦距）、以及“多模型合议”（视觉推理+地理检索）来给出可信度分数与可交互的不确定性可视化。

DiagramBot: Turning Dev Blog Posts into Diagram Concepts

<https://www.kaggle.com/competitions/banana/writeups/diagrambot-turning-dev-blog-posts-into-diagram-con>

DiagramBot 的核心功能是一键输入博客 Markdown，借助 Gemini 2.5 Flash Image Preview 自动生成三张高分辨率、风格化且贴合内容的技术图示草稿，显著缩短制图时间并提升读者停留度。

代码：

<https://aistudio.google.com/apps/drive/1kYakfbiENA3lgtHwkZboJKmGpjViyTXv?showAssistant=true&showCode=true>（核心prompt在geminiService.ts）

...

You are a master technical illustrator with a keen eye for clarity, aesthetics, and impact.

Your task is to transform this provided blank white canvas into a powerful and clear technical diagram.

This diagram must visually explain the core concepts of the following blog post content.

The final image should be professional, easy to understand, and visually engaging, suitable for inclusion in a high-quality technical article.

****Instructions:****

1. Analyze the provided blog post to identify the key entities, relationships, processes, or architecture.
2. Design a diagram (e.g., flowchart, architecture diagram, sequence diagram, mind map) that best represents this information.
3. Use clear labels, icons, and a consistent color scheme. Make it look modern and polished.
4. The entire diagram must be generated on the provided blank image. Do not add any elements outside of it.
5. \${variationPrompt}

****Blog Post Content:****

\${markdown}

...

Nano Infograph

<https://www.kaggle.com/competitions/banana/writeups/nano-infograph>

Nano Infograph把YouTube字幕通过多模态管线（文本摘要+Gemini 2.5 Flash Image生成）即时转换为可定制风格的专业信息图，并支持交互式迭代优化与实时预览。
Live Demo: <https://nano-infograph.vercel.app/>

GitHub Repository: <https://github.com/codebird17/nano-infograph>（关键prompt代码: <https://github.com/codebird17/nano-infograph/blob/main/lib/actions.ts>）

专业组

=====

AIR(Architectural AI Render)

<https://www.kaggle.com/competitions/banana/writeups/airarchitectural-ai-render>

A.I.R (Architectural AI Render) 是面向建筑师的轻量级“对话式渲染器”，基于 Gemini 2.5 Flash Image，将“草图+点/多边形遮罩+简短指令”组合为多模态提示，在不破坏几何结构的前提下进行点级编辑与局部修复，实现材料、开口、光照等精准快速调整；其工作流以建筑思维分栏：Upload（图像/截屏即用）、Edit（关键词建议与局部精修）、Reference（情绪板/参考图一键风格融合，含影响力滑块）、Moodboard（方向留存与复用）、Gallery（版本对比留档）、Ask（与AI对话探索下一步），核心亮点是结构保真、风格融合可控、实时迭代、无需重型渲染链路。脑洞建议：把它接入你现有的设计系统，用Kafka事件流记录每次编辑的“掩码-提示-结果”三元组，微调一个小型控制模型做“风格一致性守护”，再用RL从人类反馈优化“Ask”建议质量，做成团队级“设计回放与复盘”闭环。

Scientific Illustration AI

<https://www.kaggle.com/competitions/banana/writeups/scientific-illustration-ai>

Scientific Illustration AI 是一个面向科研场景的全自动图形摘要生成与优化工具：研究者只需粘贴论文摘要，系统用多模型协作（Gemini 2.5 Flash 解析与评审 + Imagen 4 初稿生成 + Flash 多模态自我批改 + Flash Image Preview 定向编辑）完成从概念设计、初稿出图到多轮科学性与清晰度改进的闭环流程，最终产出出版级、准确且高可读性的图形摘要，并提供简洁的前端交互与专业元提示工程以保证一致的高质量输出。脑洞建议：把强化学习信号接入自我评审环节，用真实审稿意见或点击/停留时长做反馈，训练一个“审稿人偏好模型”，实现跨期刊风格适配与自动对齐投期要求。

代码：

<https://aistudio.google.com/apps/drive/1KkudiQmKXEPliPhuv9aJ8FG81-rlcAzc?showAssistant=true&showCode=true>

AdAPT: AI-Powered Ad Creative Personalization

<https://www.kaggle.com/competitions/banana/writeups/adapt-ai-powered-ad-creative-personalization>

AdAPT是一款面向营销团队的人工智能创意适配工具，基于谷歌 gemini-2.5-flash-image-preview（nano banana）多模态编辑能力，用户只需上传参考广告、产品图和可选 Logo，并给出文字指令，系统在单次调用中自动分析参考广告的构图、光影与风格，将新产品无缝融入并生成高保真、版式不变的成片（通过将参考广告置于 contents.parts 序列最后确保画面框架与尺寸稳定），把原本需要 4-8 小时的手工修图压缩到约 30 秒，极大降低成本并支持快速批量变体生成与 A/B 测试，从而加速数据驱动的投放迭代与市场响应。脑洞建议：未来可加入“竞品相似度守恒”开关，自动提炼参考广告的“因子化风格向量”（构图/色温/材质/景深等）并可控拷贝或偏移，实现从“高拟合复刻”到“风格迁移创作”的连续滑块式调参。

Neighborly

<https://www.kaggle.com/competitions/banana/writeups/neighborly>

Neighborly 是一款基于 Nano Banana 图像模型的城市可视化工具，支持在 Google 地图的街景与卫星视角中一致地编辑并预览新建筑或改造效果；其核心亮点是“框选编辑（Bounding Box Magic）”可精准指定修改区域且不影响周边元素，以及“极致一致性（Extreme Consistency）”确保前后对比仅变化目标建筑、天空树木等背景保持不变，从而实现跨视角、具备建筑语义与透视理解的连贯变换，帮助社区直观评估城市更新方案。脑洞建议：未来可叠加时序版本对比（白天/夜晚/季节）、交通流量模拟、噪声/日照分析与法规合规检查插件，甚至引入多代理协作让“规划师代理+居民代理+可持续代理”共同生成折衷设计方案。

代码：没有提供代码

Edison Plus: Power to Every Creator
<https://www.kaggle.com/competitions/banana/writeups/popshop-ondemand-goods-shop-for-creators>
Edison Plus 让创作者和IP方以零库存、低门槛开启周边商店，借助“Nano Banana”生成高一致性的虚拟产品预览，缩小设计与实物差距，并通过四步流程快速上线、直接收款、专注粉丝互动。

Visual Witness System
<https://www.kaggle.com/competitions/banana/writeups/visual-witness-system>
Visual Witness System 是一款基于网页的智能取证工具，利用 Google 的 Nano Banana (Gemini 2.5 Flash Image) 多图融合、人物/物体一致性与物理与场景常识推理等能力，将来自不同来源、角度与质量的视觉证据（如监控截图、草图、反射影像、蓝图）自动整合为一幅连贯且逼真的场景重建图，帮助调查人员从零散片段中还原全貌、提升分析效率并产出可行动线索。
代码：
<https://aistudio.google.com/apps/drive/1jcl0UuK75gb6dYHRnFgmOaYVuFW6mTp-?showAssistant=true&showCode=true> （核心prompt在 `geminiService.ts`）

Hephaestus
<https://www.kaggle.com/competitions/banana/writeups/hephaestus>
Hephaestus 是一套基于 Gemini 2.5 (Nano Banana) 的多模态生成 workflow：从手绘或简单2D蓝图出发，先通过“图像+文本编辑”生成可迭代优化的顶视平面图；再用多图合成与风格迁移，在保持结构与风格一致性的前提下批量生成房间多视角渲染；最后以高级合成功能将多视角智能“拼接”为无缝的等距矩形全景，输出可探索的、逼真的沉浸式360°虚拟空间。脑洞建议：把平面图语义（房型/动线/采光）编码成约束图，让模型在迭代编辑时做“可居住性优化”，并引入小样本用户偏好学习，快速自适应不同审美与地域风格。
代码：<https://github.com/coldboxer007/HEPHAESTUS>

Github上的Banano Demos

ZHO的repo - 目前有46种原创玩法: <https://github.com/ZHO-ZHO-ZHO/ZHO-nano-banana-Creation> - **免Prompt的玩法的Hub:** <https://github.com/ZHO-ZHO-ZHO/Nano-Banalary>

参考

[1] [Google Nano Banana Overview](#) ↩

Imagine.art - 全面的Nano Banana模型介绍

[2] [Introducing Gemini 2.5 Flash Image](#) ↩

Google Developers Blog - 官方技术发布说明

[3] [Gemini Image Generation Overview](#) ↩

Gemini官方页面 - 图像生成功能概览

[4] [Nano Banana Tutorial: How to Use Google's AI Image Editing Model](#) ↩

Anangsha.me - 详细使用教程

[5] [Virtual Try-On with Nano Banana](#) ↩

ImaginePro.ai - 虚拟试衣应用案例

[6] [Online Shopping Makes It Hard to Know How New Clothes Will Look ↩](#)

Reddit讨论 - 在线购物体验问题

[7] [How to Use Nano Banana AI to Generate Professional Ad Photos ↩](#)

Flux AI Blog - 专业广告照片生成指南

[8] [Nano Banana Has Already Put Product Photographers Out of Business ↩](#)

Reddit讨论 - 对产品摄影行业的影响

[9] [Gemini 2.5 Flash Image \(Nano Banana\) Guide ↩](#)

Codecademy - 开发者使用指南

[10] [Google AI Nano Banana for Architecture ↩](#)

Archilabs.ai - 建筑领域应用

[11] [Nano Banana for Architecture Design ↩](#)

Fenestra.app - 建筑设计应用

[12] [Is There a New King of AI Image Generation? Gemini 2.5 Flash Image is Here ↩](#)

MasterConcept.ai - 技术对比分析

[13] [Medical Education Performance Study ↩](#)

PubMed - 医学教育性能研究

[14] [Nano Banana Image Editing Tool ↩](#)

Geeky Gadgets - 图像编辑工具介绍

[15] [How to Prompt Gemini 2.5 Flash Image Generation for the Best Results ↩](#)

Google Developers Blog - 提示词优化指南

[16] [Gemini 2.5 Flash Image Technical Analysis ↩](#)

Higgsfield.ai - 技术深度分析

[17] [Auto Generate Virtual AI Try-On Images for WooCommerce with Gemini Nano Banana ↩](#)

n8n.io - WooCommerce自动化 workflow

[18] [Nano Banana Creative Possibilities Video ↩](#)

YouTube - 创意应用视频演示

[19] [Nano Banana Use Cases and Creative Possibilities ↩](#)

Imagine.art - 应用案例和创意可能性

[20] [Google Nano Banana: Can't Trust Photos Online ↩](#)

Fast Company - 图像真实性和信任度讨论

其他可参考部分

[22](#) [23](#) [24](#) [25](#) [26](#) [27](#) [28](#) [29](#) [30](#) [31](#) [32](#) [33](#) [34](#) [35](#) [36](#) [37](#) [38](#) [39](#) [40](#) [41](#) [42](#) [43](#) [44](#) [45](#) [46](#) [47](#)

分享这篇文章

