

基于生成式人工智能的火山喷发科学插画设计

——三维可视化科学图像建构与优化

陈弘正^{1,2*}, 陈泰一^{2,3}, 付玲珊⁴, 陈中文⁵

¹黄冈师范学院机电与智能制造学院, 湖北 黄冈

²京都大学防灾研究所, 日本 京都

³河南理工大学资源环境学院, 河南 焦作

⁴荆州开发区高级中学, 湖北 荆州

⁵黄冈师范学院教育学院, 湖北 黄冈

收稿日期: 2025年11月27日; 录用日期: 2026年1月6日; 发布日期: 2026年1月16日

摘要

火山喷发是一种剧烈且高度结构化的地球内部物质快速释放的现象, 其多层次的动力过程与快速演化特性, 对科学图像的准确再现提出了严峻挑战。传统火山演化模型图多依赖手绘或静态示意, 难以兼顾三维结构的物理一致性与视觉表达力。本研究提出一种由语言驱动的图像生成策略, 结合GPT-4o的语意理解与Midjourney的跨模态图像生成能力, 设计出具备语意链推导与多轮优化的提示词 workflow, 逐步强化火山喷发图像的结构真实性与动态逻辑。研究透过五组图像版本的时空演化比对, 对应提示词语意的层层建构, 提出一套可追踪、可控且具教育潜力的科学图像生成流程。结果显示, 当语意设计越贴近地质术语与物理因果链时, 生成图像的解释力与教学适用性大幅提升。本研究展示了大语言模型在地球科学可视化中的知识转译潜力, 并为未来AI辅助科普图像建构开启新的可能性。

关键词

火山喷发三维剖面, 生成式人工智能, 提示词优化, 语意链设计, 地球科学可视化

Generative AI-Based Scientific Illustration Design for Volcanic Eruptions

—Construction and Optimization of 3D Visualized Scientific Imagery

Hung-Cheng Chen^{1,2*}, Taiyi Chen^{2,3}, Lingshan Fu⁴, Zhongwen Chen⁵

¹School of Mechatronics and Intelligent Manufacturing, Huanggang Normal University, Huanggang Hubei

²Disaster Prevention Research Institute (DPRI), Kyoto University, Kyoto, Japan

³School of Resources and Environment, Henan Polytechnic University, Jiaozuo Henan

*通讯作者。

文章引用: 陈弘正, 陈泰一, 付玲珊, 陈中文. 基于生成式人工智能的火山喷发科学插画设计[J]. 创新教育研究, 2026, 14(1): 100-108. DOI: 10.12677/ces.2026.141014

⁴Jingzhou Development Zone Senior High School, Jingzhou Hubei⁵School of Education, Huanggang Normal University, Huanggang Hubei

Received: November 27, 2025; accepted: January 6, 2026; published: January 16, 2026

Abstract

Volcanic eruptions are characterized by the violent and highly structured rapid release of Earth's internal matter. The multi-level dynamic processes and rapid evolutionary nature of these events pose a significant challenge to accurate scientific visualization. Traditional cross-section diagrams often fall short in capturing the three-dimensional complexity and causal dynamics required for educational use. This study introduces a language-to-visualization workflow that integrates the semantic reasoning of GPT-4o with the cross-modal image synthesis capabilities of Midjourney. We designed a prompt workflow incorporating semantic chain deduction and iterative optimization to progressively reinforce the structural authenticity and dynamic logic of volcanic eruption imagery. By comparing the spatiotemporal evolution of five sets of image versions, corresponding to the layered construction of prompt semantics, we establish a traceable, controllable, and pedagogically valuable framework for scientific image generation. Results indicate that when prompt semantics are closely aligned with geological terminology and physical causal chains, the explanatory power and educational applicability of the generated images are significantly enhanced. This study demonstrates the potential of Large Language Models (LLMs) in knowledge translation for geoscience visualization, opening new avenues for future AI-assisted construction of popular science imagery.

Keywords

3D Volcanic Eruption Cross-Section, Generative AI, Prompt Optimization, Semantic Chaining Design, Geoscience Visualization

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>

Open Access

1. 引言

1.1. 研究背景

火山喷发是地球动力系统中最剧烈、最具视觉冲击力的自然现象之一，常在地球科学教学与公众展示中被用作关键案例[1]。然而，其复杂的地质结构、能量释放机制与快速动态过程，使得传统静态图像难以全面呈现其三维特征与物理逻辑。手绘或剖面图虽常见于教材，但在科学准确性与视觉传达之间仍难以平衡。

随着生成式人工智能的发展，诸如 Midjourney 与 GPT-4o 等以自然语言为输入的图像生成工具，为非专业用户提供了构建科学图像的新途径。然而，这些生成图像是否具有足够的科学合理性？提示词如何设计才能引导模型生成具有物理一致性与解释力的教学图像？目前尚缺乏系统研究与实践验证。

本研究以火山喷发三维演化模型图为例，提出融合 GPT-4o 语义能力与 Midjourney 图像生成机制的优化流程。通过语义演化的提示词设计与多轮图像比对分析，本研究探索了一种兼顾视觉表现与科学解

释力的 AI 辅助图像生成范式，为地学科普教学与跨领域应用提供了方法参考与理论基础。

1.2. 火山喷发的视觉挑战与 AI 图像生成的突破潜能

火山喷发是地球内部能量释放的集中体现，涉及岩浆活动、地层破裂、气体逸散与热羽上升等一系列复杂的地球物理过程。从板块构造的角度来看，火山多分布于俯冲带、裂谷带或热点区域，是板块动力学、地幔对流与地壳应力共同作用的结果[2]。然而，这一自然现象虽然具有高度的科学价值和社会关注度，其三维结构与动态演化过程却常常难以被有效传达，尤其是在面向公众的教学与展示中，存在“难以想象”“难以还原”的长期问题。

传统的火山演化模型图多为二维简化模型，主要依靠人工绘图或静态渲染，虽然具备一定的示意性与教育功能，但在表现岩浆通道系统、气体储层分布、喷发动能路径等空间要素时存在局限[3][4]。这类图像往往过于理想化或艺术化，缺乏物理结构的准确性与过程机制的因果表达，难以满足新一代地球科学教育中“可视化-理解-推理”一体化的学习目标。

近年，生成式人工智能(Generative AI)技术快速发展，尤其是以自然语言提示词驱动图像生成的工具，如 Midjourney、DALL·E 与 Stable Diffusion 等，为科学图像的表达打开了新的可能。用户无需掌握建模软件与绘图技术，仅需输入语义明确的描述性语言，即可生成具备艺术张力与视觉复杂度的图像。这类系统具备强大的语义理解能力与图像生成算法支撑，在科学教育、博物馆展览、教材设计等领域展现出极大潜力[5]。但与此同时，如何确保 AI 生成图像的科学合理性、结构清晰度与可解释性，也成为当前跨领域应用中的关键挑战。

本研究以“火山喷发三维演化模型图”为核心案例，探讨如何通过提示词优化与多轮 AI 生成过程，使生成图像逐步贴近科学结构、符合物理逻辑，并具有教学可用性。研究结合 OpenAI 的 GPT-4o 模型进行语义链分析，并使用 Midjourney 进行图像生成，以“语意-结构-动能-标签”四大要素为导向，提出一套具备迭代性、可控性与教育潜能的 AI 图像生成策略。

本研究的目标，不仅是提升 AI 图像的视觉效果，更重要的是探索一种具备科学内涵的图像建构路径，使生成式 AI 成为地球科学教育中的知识转译工具。通过构建跨模态的“语意驱动图像生成链”，本研究期望推动生成式 AI 从“美术辅助”走向“科学建模”的教育价值转型，也为未来在气象灾害、地质构造、环境变化等多类型地学主题中的图像生成应用奠定基础。

2. AI 图像优化流程构建：多轮迭代与结构强化

本研究构建的图像优化流程围绕四个核心目标展开：结构可视化、动能强化、物理一致性、科学标注。通过引入 GPT-4o 模型的图像生成与解析能力，配合逐轮提示词的语意设计，我们系统性地提升了火山喷发三维演化模型图的科学性与教学适用性。

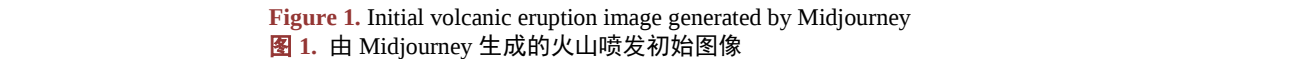
从图 1~5 的演进，不仅是画面元素的添加，更是一次次物理概念的重建。每轮图像生成，皆以新的提示词对物理现象进行语意拆解与重构，借由 GPT-4o 的图像识别能力回馈结构完整性，推动下一阶段的迭代优化。

2.1. 第一轮生成：风格主导，科学性不足

图 1 使用提示词 Prompt-1，导入 Midjourney 生成：

“exploded cutaway of a volcano mid-eruption, upper cone shattered with magma fragments and ash ejecta radiating outward, the dense eruption column spiraling upward with embedded lightning, molten lava surging through fissures and branching vents, subsurface layers revealed with magma chamber, dike networks, and gas reservoirs highlighted, set within a

生成图像具艺术张力, 喷发柱有一定立体感, 岩浆发光效果突出。然而, 底层结构如地壳(crust)、岩浆房(magma chamber)、分支岩脉(dikes)等元素模糊或缺失[6], 图像更接近幻想插画而非科学剖面图。



针对图 1 的缺陷，由 ChatGPT-4o 诊断并于第二轮生成中强化底层结构提示，透过以图生图的方式，以 Prompt-2 生成图 2：

此阶段图像(图 2)首次呈现出具分层结构的火山构造,底部岩浆房、岩浆通道、火山灰、岩脉网络与气体储层布局初具雏形,但喷出口上方的红色网络可去除。其物理基础可参考地质文献中关于火山下部岩浆迁移与气体聚集的描述[7][8]。

同上，以 ChatGPT-4o 分析图 2，其虽具结构层次，但火山锥顶结构仍过于完整，缺乏爆炸表现[9]。因此第三轮提示 Prompt-3 聚焦火山喷发动力学：

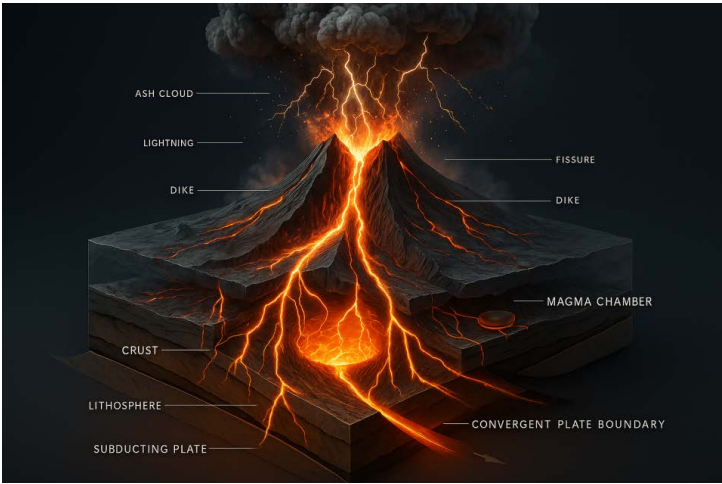


Figure 2. Volcanic eruption image optimized by GPT-4o: Introducing geological stratification and structural elements.
图 2. 由 GPT-4o 优化的火山喷发图像——引入地质分层与构造元件

“exploded cutaway 3D infographic of a volcano mid-eruption, twin volcanic cones violently shattered with magma fragments and ash bursting outward, thick eruption column spiraling upward with embedded lightning bolts, concentric shockwave rings radiating from the blast, glowing lava flowing through fissures and dike networks, underground gas reservoirs visible as translucent pockets, fully exposed crust, lithosphere and subducting plate structure in a convergent tectonic setting, directional light from below highlighting thermal plumes and tectonic stress, photorealistic, labeled scientific diagram in English --v 6 --ar 16:9 --style raw”

图 3 展现明显破碎结构，岩块飞散方向具有动势。震波环形结构亦开始出现，为后续螺旋上升气柱与冲击波动能埋下基础[10][11]；此外，火山弹、火山灰地质元素表达形象，但气体储存模型过于突兀可弱化，红色网络可去除。该阶段图像逐渐体现出火山喷发中“破碎模型”(fragmentation model)与“喷发柱动力学”(eruption column dynamics)交互作用的视觉意象[12]。

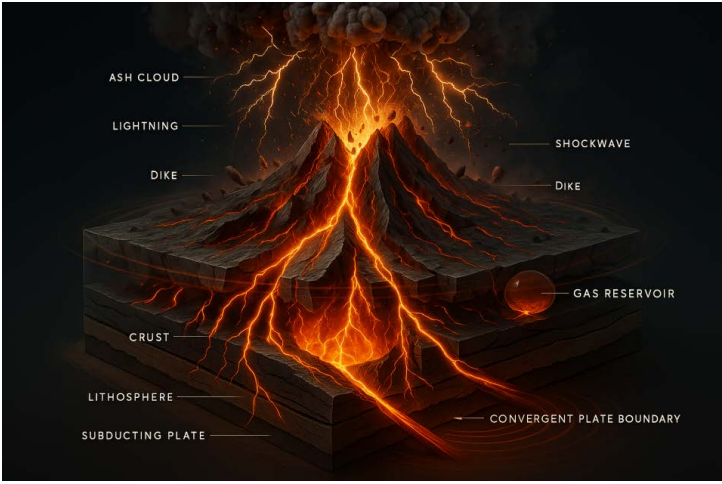


Figure 3. Volcanic eruption image optimized by GPT-4o: Incorporating explosive kinetic energy and eruption dynamics
图 3. 由 GPT-4o 优化的火山喷发图像——导入爆炸动能与喷发动力学

2.4. 第四轮迭代：整合热羽流与震波能量释放

为进一步呈现火山喷发中的热羽流螺旋上升、雷电嵌入与震波辐射三重现象，我们设计 Prompt-4:

“exploded cutaway 3D infographic of a volcano mid-eruption, twin volcanic cones violently shattered with magma fragments and ash rocks flying outward, thick eruption column spiraling upward with multiple lightning bolts, concentric shock-wave rings radiating from the eruption center, lava glowing along branching dikes and fissures, translucent gas reservoirs embedded underground, fully exposed crust-lithosphere-subducting plate layers in a convergent plate boundary setting, plume-like heat flow rising from the magma chamber, photorealistic scientific diagram with labeled elements in English --v 6 --ar 16:9 --style raw”

图 4 为本次实验的关键转折，首次达成动态层次的整合表现。喷发柱如龙卷般螺旋上升，雷电贯穿其间，震波环清晰向外扩张。此外，图 4 展现出岩浆来源，表达专业。但气体储存弱化，形状一般不为圆形，一般是顺着岩体微裂隙储存。上方红色网络删除。此图像基本重建出与热羽流理论相符的结构形态[7][8]，亦直观展现震波传播(shockwave propagation)的视觉路径。

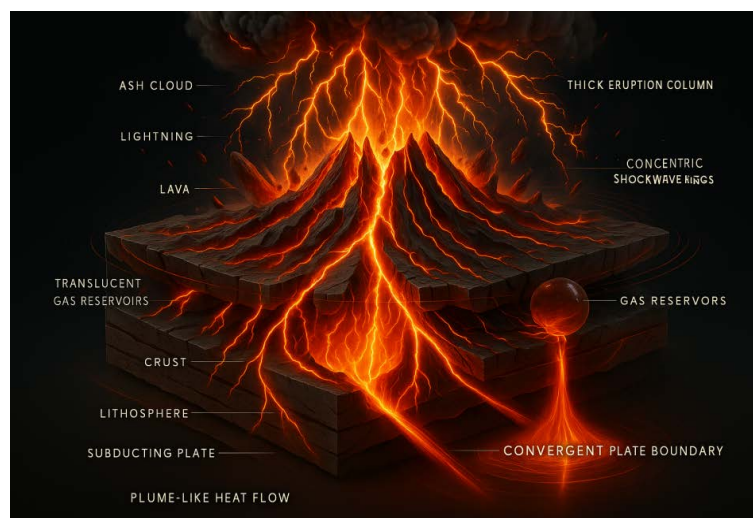


Figure 4. Volcanic eruption image optimized by GPT-4o: Integrating thermal plumes and shockwave energy release

图 4. 由 GPT-4o 优化的火山喷发图像——整合热羽流与震波能量释放

2.5. 最终整合：实现可用于教学的科学剖面图

在多轮迭代的基础上，最终版本(图 5)融合前述所有科学元素，提示词 Prompt-5 综合且精准：

“exploded cutaway 3D infographic of a volcano mid-eruption, twin volcanic cones violently shattered with cracks, missing segments, and magma rock fragments bursting outward, thick ash and fire eruption column spiraling upward like a tornado, embedded with lightning bolts, concentric glowing shockwave rings expanding from the eruption center, lava channels flowing along branching fissures and dike networks, underground translucent gas reservoirs clearly marked, full exposure of crust, lithosphere, and subducting plate in a convergent tectonic zone, rising heat plume with intense bottom lighting, photorealistic scientific infographic labeled in clean English --v 6 --ar 16:9 --style raw”

该图标注明明确，构造逻辑清晰，火山锥缺角、飞岩、熔岩分支流动与气柱螺旋上升具备一致的动力

逻辑，图像质量已可支持作为中学地科教材或博物馆展示。图中所展现的剖面结构与地质构造，亦符合近年来有关火山喷发动力学与岩浆迁移模型之研究成果[13] [14]。

本研究邀请一位地质科学家针对图 5 进行盲评，专家反馈图 5 十分形象及美观，且图中的地质元素呈现专业。专家也建议 1) 若能把喷发物距离火山口邻近的改为火山弹，而随着远离火山口其直径逐渐变小，及至最远的成为火山灰等，2) 山体表面可否增加熔岩？这些建议将纳入下一轮的实验中改善。

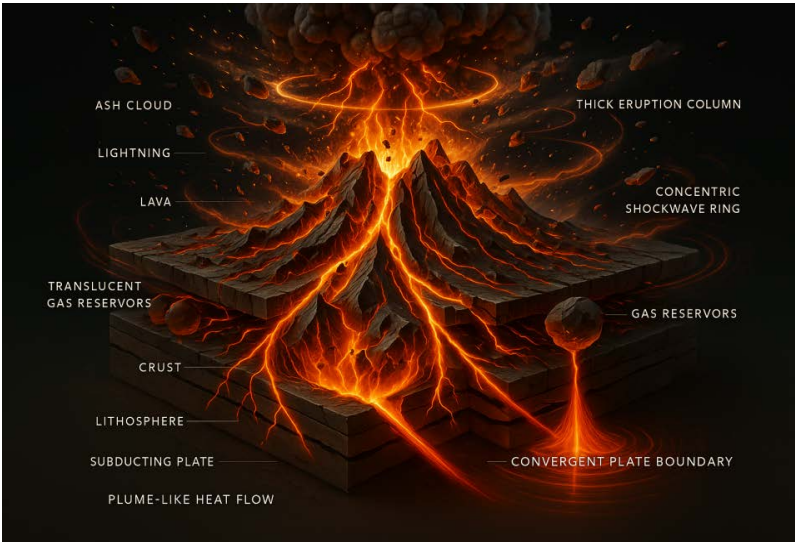


Figure 5. Volcanic eruption image optimized by GPT-4o: Final integrated version
图 5. GPT-4o 优化的火山喷发图像——最终整合版本

3. 多轮语意驱动的视觉生成与物理结构重构

火山喷发图像的生成优化，不仅是图像构成要素的逐步补全，更是一场语言与物理概念之间的编码与解码过程。在本节中，我们结合 GPT-4o 的语意理解与 Midjourney 的图像生成能力，透过多轮提示词 (prompt) 设计与实验，逐步建立一套「语意链驱动的视觉耦合逻辑模型」(semantic-driven visual coupling model)，使火山喷发的复杂结构得以科学且美学兼备地呈现。

3.1. 提示词演化与图像变异的耦合观察

本研究从初始的图像生成(图 1)开始，逐轮调整提示词语意结构与密度，并透过 GPT-4o 进行图像语意检核与结构建议。为系统化观察语意变化与图像生成之间的关联，我们汇整并分析图 1~5 之提示词与关键变异特征(表 1)。

Table 1. Comparison of semantic evolution in visual prompts for volcanic eruptions
表 1. 火山喷发视觉提示词语意演化对照表

图像	关键提示词语句(英文)	显著视觉变化
图 1	“exploded cutaway infographic of a volcano mid-eruption, photorealistic”	仅有装饰性火山形象与粗略剖面，缺乏分层与物理一致性
图 2	“subsurface layers revealed with magma chamber, dike networks, and gas reservoirs highlighted”	出现地下结构与气体储层，部分地层边界可见

续表

图 3	“twin volcanic cones violently shattered with magma fragments and ash rocks flying outward”	火山破裂、飞岩动势出现，动能释放概念初步成形
图 4	“spiraling upward fire column, explosive concentric shockwave rings”	热羽柱与震波环完整形成，开始体现羽流动力与压力释放模型
图 5	“tornado-like fire column, translucent gas reservoirs, labeled English diagram, 3:2 aspect ratio”	各构造整合呈现，图像达到物理一致性与教学展示的完整性

提示词的语意设计，呈现出一种由抽象形容(如 “photorealistic”)迈向具体物理概念(如 “shockwave rings” “gas reservoirs”)的语义进化轨迹。从单点式名词(如 lava、eruption)进展至多层结构与动态相互关联的描述，逐步建立视觉与地球科学知识的耦合机制。

这类语意设计实际上建立了从地球科学理论→自然语言→图像生成的语意链结(semantic chaining)。当语句越具物理意义与空间逻辑，其对应图像的结构一致性与物理解释潜力亦随之提升。

3.2. 语意结构与物理原理的映射实践

在每轮图像生成中，语意词组的设计皆对应特定物理过程。例如图 2 中的“subsurface layers revealed”对应地球物理的岩浆房压力累积与气体扩散理论[15]；图 4 提示词中的“spiraling upward”及“shockwave rings”直观对应火山热羽流[7]与爆炸震波[16]的动态模型。

以下为本研究中所使用的提示词(Prompt-1 至 Prompt-5)之摘要整理：

- Prompt-1: 装饰性剖面结构 photorealistic
- Prompt-2: 地下结构层析强化 magma chamber + gas reservoirs
- Prompt-3: 破裂动能表现 shattered cones + rock fragments
- Prompt-4: 羽流与震波场景建构 spiraling column + shockwave rings
- Prompt-5: 全结构整合与教学图定位 translucent reservoirs + labeled components

每一轮提示词皆逐步导入新的物理词汇与视觉元素，形成「语意叠加－视觉聚合」的耦合过程。这种过程既可回溯(由图像追踪语意)也可前导(由语意推进视觉)，具备双向知识建构潜力。

4. 结语

本研究以火山喷发三维演化模型图为例，提出一套语意引导的生成式 AI 科学图像优化流程，结合 GPT-4o 的提示词构建、ChatGPT-4o 多模态图像解析，及 Midjourney 的图像生成能力，探索“语言－图像－物理结构”之间的转换逻辑。在对比初始生成与多轮优化结果的过程中，我们揭示了提示词细化如何逐步增强图像的科学性、可解释性与教学价值，进而构建出具备物理一致性与视觉清晰度的科学图像。

GPT-4o 在多轮提示词对话中展现出对火山动力结构、地层层次、热羽与震波等复杂语义的高度适应能力。通过链式提示词引导(prompt chaining)与语义层次重构，我们能够将地质学知识系统性地转化为具体可视化要素，提升图像的学术参考价值与传播效果。最终版本不仅呈现岩浆库、火山口、裂隙、热柱与气体层等关键结构，也具备用于科普展览与数字教材的潜力。

此外，我们提出提示词演化分析表(表 1)，追踪语句变化如何影响图像生成表现。结果表明，从抽象描述如“photorealistic”转向具体术语如“concentric shockwave rings”能显著提升图像的物理逻辑与教学适用度。这说明语义精度与图像结构质量之间存在可调控的映射关系。

面向未来，AI 辅助科学图像生成有望在数字科普、在线教育、沉浸式展示等领域发挥重要作用[17]。但当前模型仍存在物理机制认知不足、生成不确定性高等问题。因此，构建跨模态语料、开发物理嵌入

式生成框架、增强人机协同机制，将是推动此技术走向成熟的重要方向。本研究不仅提供一套可操作的优化流程，更揭示了 AI 语言模型如何介入科学图像构建过程，形成从语言推理到空间建模的认知桥梁。这对于促进 AI 在地球科学教育与公众传播中的实践应用，具有积极的启示价值。

基金项目

本论文接受黄冈师范学院博士基金项目“似台风涡旋通过地形模拟与扩增实境科普应用”(项目编号: 2042022027)、黄冈师范学院教学研究项目“教育大语言模型于现代教育技术的应用研究”(项目编号: 0601202435)及黄冈师范学院博士基金项目“似台风涡旋通过地形模拟与大语言模型融合应用”(项目编号: 204202535904)的资助, 特此致谢。

参考文献

- [1] 陈亚恒, 刘会玲, 闫丰. 地质学基础课程教学五育并举的探索与实践[J]. 林区教学, 2025(9): 79-83.
- [2] Dingwell, D.B. (1998) Magma Degassing and Fragmentation: Recent Experimental Advances. In: *From Magma to Tephra—Modelling Physical Processes of Explosive Volcanic Eruptions*, Elsevier, 1-23.
- [3] 孙先达, 王璞珺, 索丽敏, 等. 松辽盆地火山岩储层三维可视化描述[J]. 吉林大学学报(地球科学版), 2007(6): 1272-1278.
- [4] 李百寿, 秦其明, 邢立新, 等. 基于 3DS MAX 的火山地质景观三维建模与动态模拟研究[J]. 北京大学学报(自然科学版), 2008(3): 421-428.
- [5] Chen, H. (2023) Harnessing AI for Scientific Illustration: Exploring Tornado Dynamics through Midjourney. 2023 12th International Conference on Awareness Science and Technology (iCAST), Taichung, 9-11 November 2023, 136-141. <https://doi.org/10.1109/icast57874.2023.10359310>
- [6] Zhang, Z., Deng, Y., Xu, Y. and Li, X. (2024) A Translithospheric Magmatic System Revealed beneath Changbaishan Volcano. *Geology*, **52**, 742-746. <https://doi.org/10.1130/g52242.1>
- [7] Zhu, Y., Bardeen, C.G., Tilmes, S., Mills, M.J., Wang, X., Harvey, V.L., et al. (2022) Perturbations in Stratospheric Aerosol Evolution Due to the Water-Rich Plume of the 2022 Hunga-Tonga Eruption. *Communications Earth & Environment*, **3**, Article No. 248. <https://doi.org/10.1038/s43247-022-00580-w>
- [8] Carr, J.L., Horváth, Á., Wu, D.L. and Friberg, M.D. (2022) Stereo Plume Height and Motion Retrievals for the Record-setting Hunga Tonga-Hunga Ha’apai Eruption of 15 January 2022. *Geophysical Research Letters*, **49**, 1-7. <https://doi.org/10.1029/2022gl098131>
- [9] 冯晶晶, 赵勇伟, 李霓, 等. 琼北晚第四纪火山锥体形貌与喷发机制[J]. 地震地质, 2022, 44(5): 1107-1127.
- [10] 王耀平. “打开的火山教科书”——五大连池[J]. 地理教学, 2024(21): 2, 65.
- [11] 刘嘉麒, 郭正府, 刘强. 火山灾害与监测[J]. 第四纪研究, 1999(5): 414-422.
- [12] Zuccarello, F., Andronico, D., Del Carlo, P., de’ Michieli Vitturi, M., Di Roberto, A., Ganci, G., et al. (2025) Trigger Mechanism and Propagation Dynamics of Pyroclastic Density Currents at Basaltic Volcanoes. *Communications Earth & Environment*, **6**, Article No. 495. <https://doi.org/10.1038/s43247-025-02397-9>
- [13] Arzilli, F., La Spina, G., Burton, M.R., Polacci, M., Le Gall, N., Hartley, M.E., et al. (2019) Magma Fragmentation in Highly Explosive Basaltic Eruptions Induced by Rapid Crystallization. *Nature Geoscience*, **12**, 1023-1028. <https://doi.org/10.1038/s41561-019-0468-6>
- [14] Kubota, T., Saito, T. and Nishida, K. (2022) Global Fast-Travelling Tsunamis Driven by Atmospheric Lamb Waves on the 2022 Tonga Eruption. *Science*, **377**, 91-94. <https://doi.org/10.1126/science.abo4364>
- [15] Zhou, X., Kuiper, K., Wijbrans, J. and Vroon, P. (2022) Parameters Controlling the Eruption Frequency of Long-Lived Felsic Magmatic Systems: An Example from the Milos Volcanic Field (Greece). *Geochemistry, Geophysics, Geosystems*, **23**, 1-18. <https://doi.org/10.1029/2021gc009839>
- [16] Dalal, P., Kundu, B., Panda, J. and Jin, S. (2023) Atmospheric Lamb Wave Pulse and Volcanic Explosivity Index Following the 2022 Hunga Tonga (South Pacific) Eruption. *Frontiers in Earth Science*, **10**, Article ID: 931545. <https://doi.org/10.3389/feart.2022.931545>
- [17] Chen, H. (2025) Harnessing AI for Scientific Illustration: Exploring Tropical Cyclone Dynamics Using ChatGPT and Midjourney. 2025 IEEE International Conference on Computation, Big-Data and Engineering (ICCBE), Penang, 27-29 June 2025, 568-573. <https://doi.org/10.1109/iccbe65177.2025.11255765>