# Mô hình Alibaba Cloud (Qwen và Z-Image)

* **Qwen-Image (20B)**: Mô hình tạo ảnh từ văn bản và chỉnh sửa ảnh do Alibaba phát triển, cho chất lượng hình sinh động, đặc biệt trong việc hiển thị văn bản phức tạp trên ảnh[[1]](https://www.alibabacloud.com/blog/introducing-qwen-image-novel-model-in-image-generation-and-editing_602447#:~:text=Alibaba%20released%20Qwen,rendering%20and%20precise%20image%20editing)[[2]](https://qwenlm.github.io/blog/qwen-image/#:~:text=%2A%20Superior%20Text%20Rendering%3A%20Qwen,visual%20realism%20during%20editing%20operations). Qwen-Image-Edit (phiên bản chỉnh sửa ảnh) hỗ trợ cả *semantic edit* (thêm, xóa, thay đổi đối tượng) và *appearance edit* (thêm bảng hiệu, văn bản, hoán đổi phong cách) trên ảnh đầu vào[[3]](https://huggingface.co/Qwen/Qwen-Image-Edit#:~:text=,addition%2C%20deletion%2C%20and%20modification%20of). Ví dụ bên dưới là kết quả Qwen-Image-Edit thêm các phụ kiện cho chú chuột Capybara gốc, thể hiện khả năng thêm đối tượng và style mới trên ảnh đầu vào[[3]](https://huggingface.co/Qwen/Qwen-Image-Edit#:~:text=,addition%2C%20deletion%2C%20and%20modification%20of):  
    
  *Hình: Ảnh minh hoạ Qwen-Image-Edit thêm phụ kiện (guitar, mũ, váy) cho nhân vật. Qwen-Image-Edit duy trì nét nhân vật gốc trong khi thêm các đối tượng mới*[*[3]*](https://huggingface.co/Qwen/Qwen-Image-Edit#:~:text=,addition%2C%20deletion%2C%20and%20modification%20of)*.*
* **Z-Image Turbo (6B)**: Mô hình text-to-image hiệu năng cao của Alibaba-Pai, chỉ cần ~8 bước phân phối (*NFEs*) cho ra ảnh chất lượng cao và hỗ trợ tiếng Anh/Trung song ngữ[[4]](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=parameters,variants). Z-Image Turbo nổi bật về khả năng tạo ra ảnh **siêu thực (photorealistic)** nhanh chóng. Ví dụ bên dưới là ảnh phong cảnh được Z-Image Turbo sinh ra cho prompt ngắn, thể hiện độ chi tiết và màu sắc chân thực:  
    
  *Hình: Ảnh phong cảnh phong cách thực tế do Z-Image Turbo tạo, cho thấy chi tiết sắc nét và màu sắc sống động*[*[4]*](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=parameters,variants)*.*  
  Mô hình cũng xử lý tốt **văn bản trên ảnh** ở cả tiếng Anh và Trung. Ví dụ dưới là áp phích phim do Z-Image Turbo sinh, với tiêu đề và tagline rõ ràng bằng cả hai ngôn ngữ:  
    
  *Hình: Minh hoạ khả năng hiển thị văn bản song ngữ (Anh & Trung) phức tạp trên ảnh của Z-Image Turbo*[*[4]*](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=parameters,variants)*.*  
  Ngoài ra, Alibaba chuẩn bị phát hành **Z-Image-Edit** (fine-tune từ Z-Image Turbo) chuyên cho nhiệm vụ chỉnh sửa ảnh. Mặc dù còn đang hoàn thiện, tài liệu của Z-Image cho biết bản Edit sẽ hỗ trợ “biến hình ảnh theo chỉ dẫn” hiệu quả[[5]](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=%2A%20Z,tuning%20and%20custom%20development). Ví dụ minh hoạ (dưới) cho thấy Z-Image-Edit có thể xoay đổi phong cách, xoay đối tượng và thêm vật thể mới trên ảnh đầu vào:  
    
  *Hình: Ảnh minh hoạ Z-Image-Edit áp dụng các lệnh chỉnh sửa phức tạp (thêm bảng hiệu, chuyển đổi phong cách nghệ thuật) với thông tin ngôn ngữ song ngữ, đồng thời giữ nguyên tính thống nhất của đối tượng*[*[6]*](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=match%20at%20L123%20Creative%20Image,imaginative%20and%20flexible%20image%20transformations)*.*

Hai mô hình Qwen và Z-Image đều **mã nguồn mở** (Apache-2.0), có sẵn trọng số trên Hugging Face và GitHub[[7]](https://www.alibabacloud.com/blog/introducing-qwen-image-novel-model-in-image-generation-and-editing_602447#:~:text=Alibaba%20released%20Qwen,leading%20model%20in%20the%20field)[[5]](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=%2A%20Z,tuning%20and%20custom%20development). Điều này cho phép chạy offline và tinh chỉnh mô hình (fine-tune). Z-Image-Base (phiên bản nền của Turbo) đã được phát hành để cộng đồng tự huấn luyện thêm[[5]](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=%2A%20Z,tuning%20and%20custom%20development). Cả hai cũng dễ dàng tích hợp với ControlNet hay các công cụ bổ trợ (ví dụ Z-Image-Fun-Controlnet hỗ trợ nhiều điều kiện như Canny, Depth, Pose, v.v. để kiểm soát kết quả tạo ảnh[[8]](https://huggingface.co/alibaba-pai/Z-Image-Turbo-Fun-Controlnet-Union#:~:text=,stronger%20control%20and%20better%20detail)). Về API/web UI, bạn có thể dùng **Qwen Chat (qwen.ai)** để sinh và chỉnh ảnh miễn phí hoặc Hệ sinh thái ModelScope của Alibaba; **Tencent Hunyuan Bot** (dưới) và **DeepSeek** (dưới) cũng cho phép tạo ảnh qua web. Những mô hình này không áp dụng kiểm duyệt riêng, tức **NSFW/R18** có thể xử lý nếu prompt cho phép.

## Tencent Hunyuan-DiT

* **Hunyuan-DiT (Tencent)**: Mô hình text-to-image đa vòng hội thoại của Tencent Hunyuan, tối ưu cho tiếng Trung và Anh[[9]](https://huggingface.co/Tencent-Hunyuan/HunyuanDiT#:~:text=We%20present%20Hunyuan,art%20in). Hunyuan-DiT có kiến trúc diffusion trên latent space, sử dụng VAE và transformer, đồng thời tích hợp bộ mã hoá ngôn ngữ song ngữ (bilingual CLIP + đa ngôn ngữ T5)[[9]](https://huggingface.co/Tencent-Hunyuan/HunyuanDiT#:~:text=We%20present%20Hunyuan,art%20in)[[10]](https://huggingface.co/Tencent-Hunyuan/HunyuanDiT#:~:text=Model%20,0B%20DialogGen). Mô hình này hỗ trợ tạo ảnh qua nhiều lượt tương tác (người dùng có thể chỉnh sửa prompt liên tục để cải thiện ảnh)[[9]](https://huggingface.co/Tencent-Hunyuan/HunyuanDiT#:~:text=We%20present%20Hunyuan,art%20in)[[11]](https://huggingface.co/Tencent-Hunyuan/HunyuanDiT#:~:text=Image), giúp linh hoạt hơn so với mô hình 1 bước truyền thống. Theo Tencent, Hunyuan-DiT đạt **trình độ hàng đầu trong cộng đồng mã nguồn mở Trung Quốc** về độ tương thích text-ảnh và tính thẩm mỹ[[12]](https://huggingface.co/Tencent-Hunyuan/HunyuanDiT#:~:text=evaluate%20data%20for%20iterative%20model,art%20in). Hạ tầng có sẵn checkpoint cho phép tải về và dùng offline (nhưng do là 1 thể loại lớn multi-turn, cần GPU mạnh). Hiện Hunyuan-DiT chưa chú trọng NSFW (không công bố cơ chế lọc), nên tính mở rộng nội dung cao. Để triển khai, có thể sử dụng mô-đun diffusers hoặc TensorRT do Tencent cung cấp, hoặc thử nghiệm trực tiếp qua **Hunyuan Bot** miễn phí trên web.

## Mô hình viện nghiên cứu và cộng đồng Trung Quốc

* **Taiyi-Diffusion-XL (IDEA-CCNL)**: Mô hình Stable Diffusion nâng cao 3.5B tham số cho văn bản-ảnh song ngữ[[13]](https://huggingface.co/IDEA-CCNL/Taiyi-Stable-Diffusion-XL-3.5B#:~:text=Diffusion). Được huấn luyện đặc biệt tăng khả năng hiểu và tạo ảnh cho prompt tiếng Trung dài, đồng thời giữ tốt hiệu năng tiếng Anh. Mô hình này cũng mở mã nguồn (Apache-2.0)[[13]](https://huggingface.co/IDEA-CCNL/Taiyi-Stable-Diffusion-XL-3.5B#:~:text=Diffusion), có code trên GitHub và tích hợp sẵn trong giao diện Fooocus (ví dụ *Fooocus-Taiyi-XL*[[14]](https://huggingface.co/IDEA-CCNL/Taiyi-Stable-Diffusion-XL-3.5B#:~:text=%2A%20Train%20Code%20%E8%AE%AD%E7%BB%83%E4%BB%A3%E7%A0%81%EF%BC%9Ahttps%3A%2F%2Fgithub.com%2FIDEA,XL)). Taiyi-XL phù hợp khi cần tạo ảnh phong cách đa dạng (nhờ cơ sở SD-XL), hỗ trợ fine-tune thêm và dễ kết hợp với ControlNet/LoRA thông qua thư viện Diffusers.
* **Lumina-Image-2.0 (Alpha-VLLM, Shanghai AI Lab)**: Mô hình tạo ảnh 2B tham số mới, sử dụng khung **Lumina2 (Flow-Based Diffusion Transformer)**[[15]](https://huggingface.co/Alpha-VLLM/Lumina-Image-2.0#:~:text=Image%3A%20Lumina,more%20information%2C%20visit%20our%20GitHub). Lumina2 được công bố là hiệu quả và linh hoạt, có độ phân giải cao, và hoàn toàn mã nguồn mở (Apache-2.0)[[15]](https://huggingface.co/Alpha-VLLM/Lumina-Image-2.0#:~:text=Image%3A%20Lumina,more%20information%2C%20visit%20our%20GitHub). Mô hình có sẵn trên Hugging Face kèm bản demo (Gradio). Nó cũng là nền tảng cho các mô hình chuyên biệt mà cộng đồng phát triển, như **Neta-Lumina** (Neta.art) – một mô hình **phong cách anime/cute** cao cấp. Neta-Lumina dựa trên Lumina2, được huấn luyện với hàng triệu ảnh anime và nhãn đa ngôn ngữ[[16]](https://huggingface.co/api/resolve-cache/models/neta-art/Neta-Lumina/65c078e75781946e9f3423fe41457adae3eba3cd/README-ZH.md?download=true&etag=%22b1c5df428e8a54e4cdc35dcbe8749305a735fb47%22#:~:text=%EF%B8%8F%20%20Your%20browser%20does,%E6%95%B0%E6%8D%AE%E9%9B%86%EF%BC%9A%E7%B2%BE%E9%80%89%E6%95%B0%E5%8D%81%E4%B8%87%E5%BC%A0%E9%AB%98%E8%B4%A8%E9%87%8F%E5%8A%A8%E6%BC%AB%E9%A3%8E%E6%A0%BC%E5%9B%BE%E5%83%8F), phù hợp cho vẽ truyện tranh, poster, thiết kế nhân vật v.v. Cả Lumina và Neta-Lumina đều hỗ trợ ComfyUI (file quy trình có sẵn) và chạy tốt trên GPU ~8GB.[[17]](https://huggingface.co/api/resolve-cache/models/neta-art/Neta-Lumina/65c078e75781946e9f3423fe41457adae3eba3cd/README-ZH.md?download=true&etag=%22b1c5df428e8a54e4cdc35dcbe8749305a735fb47%22#:~:text=%E4%BD%BF%E7%94%A8%E6%8C%87%E5%8D%97%EF%BC%88How%20to%20Use%EF%BC%89%20%E8%AF%A5%E6%A8%A1%E5%9E%8B%E9%87%87%E7%94%A8%E5%9F%BA%E4%BA%8E%20Diffusion,safetensors%EF%BC%88%E7%BA%A6%20300%20MB%EF%BC%89). Ví dụ Neta-Lumina có hai phiên bản (Base và Aesthetic) để tinh chỉnh về mỹ thuật.

Nhìn chung, những mô hình trên đều **miễn phí tải về và chạy offline**, hoặc có API/chats web miễn phí từ nhà phát triển. Chất lượng đầu ra của chúng được đánh giá rất cao: Tencent và Alibaba đều công bố đạt thứ hạng SOTA trong các benchmark (DALL-E 3, SDXL) đối với tạo ảnh tổng quát và chỉnh sửa hình ảnh[[12]](https://huggingface.co/Tencent-Hunyuan/HunyuanDiT#:~:text=evaluate%20data%20for%20iterative%20model,art%20in)[[1]](https://www.alibabacloud.com/blog/introducing-qwen-image-novel-model-in-image-generation-and-editing_602447#:~:text=Alibaba%20released%20Qwen,rendering%20and%20precise%20image%20editing). Cụ thể, Qwen-Image và Lumina2 tỏ ra ưu thế trong việc thể hiện văn bản phức tạp bên trong ảnh, còn Z-Image và Taiyi-XL nổi bật về tốc độ và độ chi tiết ảnh. Các mô hình ổn định này đều hiểu prompt phức tạp khá tốt nhờ encoder tiếng Trung tối ưu. Điểm mạnh của chúng là khả năng thêm độ phong phú nội dung (văn bản, ký tự, chi tiết) mà vẫn giữ tính chân thực cao. Nhược điểm chung là cần GPU hiệu năng lớn để chạy bản đầy đủ; dù vậy Z-Image Turbo (6B) vẫn có thể chạy “nhẹ hơn” (VRAM ~16GB) nhờ thiết kế tối ưu và inference nhanh.

## Công cụ chỉnh sửa ảnh bổ trợ

* **Qwen-Image-Edit & Z-Image-Edit**: Như đã đề cập, cả hai đều hỗ trợ chỉnh sửa ảnh ngôn ngữ tự nhiên mạnh mẽ[[3]](https://huggingface.co/Qwen/Qwen-Image-Edit#:~:text=,addition%2C%20deletion%2C%20and%20modification%20of)[[6]](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=match%20at%20L123%20Creative%20Image,imaginative%20and%20flexible%20image%20transformations) (thêm/xóa đối tượng, chuyển phong cách, xoay xoay), gần tương đương Grok Edit. Qwen-Edit cho phép chỉnh sửa text trên ảnh rất chính xác (giữ phông chữ, kích thước)[[18]](https://huggingface.co/Qwen/Qwen-Image-Edit#:~:text=style%20transfer%2C%20allowing%20overall%20pixel,addition%2C%20deletion%2C%20and%20modification%20of). Z-Image-Edit đang trong giai đoạn phát hành, dự kiến sẽ tích hợp tốt với hệ sinh thái Alibaba (có phiên bản ControlNet sắp ra).
* **Inpaint Anything (USTC)**: Bộ công cụ mã nguồn mở do Đại học Khoa học & Công nghệ Trung Quốc phát triển, dùng SAM+LaMa để **xóa hoặc thay thế đối tượng bất kỳ** trên ảnh[[19]](https://github.com/geekyutao/Inpaint-Anything#:~:text=TL%3B%20DR%3A%20Users%20can%20select,Replace%20Anything). Người dùng chỉ cần chọn vật thể (click) và Inpaint Anything sẽ tự động tách vùng (SAM) và tô vùng trống bằng nội dung mong muốn (LaMa)[[19]](https://github.com/geekyutao/Inpaint-Anything#:~:text=TL%3B%20DR%3A%20Users%20can%20select,Replace%20Anything). Nó hỗ trợ cả xóa, điền (fill) hoặc thay nền (replace) tùy prompt[[19]](https://github.com/geekyutao/Inpaint-Anything#:~:text=TL%3B%20DR%3A%20Users%20can%20select,Replace%20Anything), giúp triệt để loại bỏ vật thể hoặc đổi nội dung mà vẫn giữ nền hiện hữu. Công cụ này có sẵn code và giao diện web demo[[20]](https://github.com/geekyutao/Inpaint-Anything#:~:text=News), nên có thể dùng local để xóa/hiệu chỉnh ảnh sau khi tạo.
* **ControlNet & Pipeline khác**: Các mô hình trên đều có thể tích hợp ControlNet (hoặc tương đương) nếu dùng nền tảng SD bình thường. Ví dụ, Alibaba cung cấp mô hình **Z-Image-Fun-ControlNet** mở rộng với 6 block control trên ảnh (hỗ trợ Canny/HED/Depth/Pose)[[8]](https://huggingface.co/alibaba-pai/Z-Image-Turbo-Fun-Controlnet-Union#:~:text=,stronger%20control%20and%20better%20detail). Ứng dụng ControlNet hoặc SAM/LoRA khác trong quy trình Diffusers hoặc ComfyUI cho phép người dùng tùy chỉnh chính xác tư thế, phối cảnh, hoặc vẽ đè thô sơ trước khi sinh ảnh. Ngoài ra còn có các thư viện chỉnh sửa khác như LaMa (xóa vật thể) hay BrushNet, phù hợp khi cần tinh chỉnh tương tác.

## Triển khai và tích hợp

Tất cả mô hình trên đều có **trọng số và mã nguồn mở**, nên dễ dàng cài chạy local. Ví dụ, Hugging Face có các pipeline sẵn sàng: như QwenImageEditPipeline trong thư viện diffusers (ví dụ code tại model card)[[21]](https://huggingface.co/Qwen/Qwen-Image-Edit#:~:text=import%20os%20from%20PIL%20import,Image%20import%20torch), Lumina2Pipeline cho Lumina-Image-2.0[[22]](https://huggingface.co/Alpha-VLLM/Lumina-Image-2.0#:~:text=import%20torch%20from%20diffusers%20import,Lumina2Pipeline), hoặc pipeline Z-Image mới nhất (yêu cầu cài diffusers bản dev)[[23]](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=Install%20the%20latest%20version%20of,diffusers%2C%20use%20the%20following%20command). Nhiều mô hình đã có hướng dẫn tích hợp vào UI: Taiyi-XL có không gian *Fooocus-Taiyi-XL* sẵn[[14]](https://huggingface.co/IDEA-CCNL/Taiyi-Stable-Diffusion-XL-3.5B#:~:text=%2A%20Train%20Code%20%E8%AE%AD%E7%BB%83%E4%BB%A3%E7%A0%81%EF%BC%9Ahttps%3A%2F%2Fgithub.com%2FIDEA,XL), Neta-Lumina có file workflow ComfyUI kèm hướng dẫn[[17]](https://huggingface.co/api/resolve-cache/models/neta-art/Neta-Lumina/65c078e75781946e9f3423fe41457adae3eba3cd/README-ZH.md?download=true&etag=%22b1c5df428e8a54e4cdc35dcbe8749305a735fb47%22#:~:text=%E4%BD%BF%E7%94%A8%E6%8C%87%E5%8D%97%EF%BC%88How%20to%20Use%EF%BC%89%20%E8%AF%A5%E6%A8%A1%E5%9E%8B%E9%87%87%E7%94%A8%E5%9F%BA%E4%BA%8E%20Diffusion,safetensors%EF%BC%88%E7%BA%A6%20300%20MB%EF%BC%89). Bạn có thể dùng **Stable Diffusion Web UI (AUTOMATIC1111/Fooocus)** để load các mô hình SD/DiT này (nhiều mod mở hỗ trợ). Các dịch vụ web miễn phí cũng có sẵn: **Qwen Chat** (tính năng “Image Generation” miễn phí)[[24]](https://qwenlm.github.io/blog/qwen-image/#:~:text=We%20are%20thrilled%20to%20release,Chat%20and%20choose%20%E2%80%9CImage%20Generation%E2%80%9D), **Tencent Hunyuan Bot** (text-to-image miễn phí)[[25]](https://huggingface.co/Tencent-Hunyuan/HunyuanDiT#:~:text=Welcome%20to%20Tencent%20Hunyuan%20Bot%2C,you%20desire%2C%20all%20for%20free), hoặc **Hugging Face Spaces** (các bản demo cộng đồng cho Lumina, Qwen, Z-Image).

Ngoài ra, cả Qwen-Image và Z-Image đều hỗ trợ **fine-tune cộng đồng** (đã phát hành bản Base để huấn luyện thêm[[5]](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=%2A%20Z,tuning%20and%20custom%20development)). Bạn có thể tải về và tinh chỉnh theo nhu cầu cụ thể (ví dụ thêm dữ liệu phong cách riêng). Đối với NSFW/R18, vì không có kiểm duyệt nội bộ, bạn tự chịu trách nhiệm về nội dung. Cuối cùng, để tích hợp với ứng dụng, các mô hình này có thư viện Python chuẩn (diffusers), hoặc API thương mại (Aliyun, Tencent, DeepSeek, có thể có hạn chế).

## So sánh & Gợi ý

* **Chất lượng**: Tất cả mô hình đều chất lượng cao, nằm trong nhóm hàng đầu của thế giới AI hiện nay. Qwen-Image (20B) và Lumina2 (2B) tỏ ra xuất sắc khi prompt có nhiều văn bản hoặc cấu trúc phức tạp[[1]](https://www.alibabacloud.com/blog/introducing-qwen-image-novel-model-in-image-generation-and-editing_602447#:~:text=Alibaba%20released%20Qwen,rendering%20and%20precise%20image%20editing)[[2]](https://qwenlm.github.io/blog/qwen-image/#:~:text=%2A%20Superior%20Text%20Rendering%3A%20Qwen,visual%20realism%20during%20editing%20operations). Z-Image (6B) tạo ảnh nhanh, chi tiết tốt cho prompt thông thường, và hiện đang dẫn đầu về tốc độ giữa các mô hình mở[[4]](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=parameters,variants). Hunyuan-DiT (khoảng 7B) đặt trọng tâm vào tương tác đa vòng và tiếng Trung, cho hình ảnh phù hợp ngữ cảnh hơn qua các lần hội thoại[[9]](https://huggingface.co/Tencent-Hunyuan/HunyuanDiT#:~:text=We%20present%20Hunyuan,art%20in). Neta-Lumina cho kết quả chuyên sâu phong cách anime, trong khi Taiyi-XL là lựa chọn linh hoạt hơn khi cần hỗ trợ tiếng Trung lẫn tiếng Anh.
* **Hiệu năng & triển khai**: Z-Image Turbo có ưu thế về tốc độ inference (giới thiệu “sub-second” trên GPU H800)[[4]](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=parameters,variants), phù hợp khi cần tạo nhanh. Qwen và Hunyuan (mô hình lớn) sẽ chậm hơn và đòi hỏi GPU mạnh (như H100). Tuy nhiên cả hai đều có bản *distill* nhỏ hơn hoặc phiên bản infer đồ họa (như Z-Image Turbo) để dùng ở phần cứng nhẹ hơn. Tất cả đều dùng kiến trúc Diffusion (DiT/Lumina) có thể chạy trên Diffusers hoặc nền tảng tương đương (PaddlePaddle…). Do mã nguồn mở, bạn có thể điều chỉnh phân phối (bfloat16, TensorRT, v.v.) để tối ưu tốc độ.
* **Ưu/Nhược**:
* *Ưu điểm*: Hỗ trợ prompt đa ngôn ngữ; chỉnh sửa ảnh mạnh (theo ngữ cảnh hoặc chi tiết); có thể chạy offline và fine-tune; cộng đồng Trung Quốc phát triển tích cực, cập nhật nhanh.
* *Nhược điểm*: Nhiều mô hình rất lớn (20B, 7B…) nên không phải ai cũng dùng được offline ngay; phụ thuộc GPU mạnh. Một số mô hình (Hunyuan, Lumina) còn mới nên khung hỗ trợ (UI, API) chưa phổ biến ở Việt Nam. Ngoài ra, tuy không kiểm duyệt nội dung thì cũng cần cẩn trọng với dữ liệu đầu vào để tránh vi phạm luật pháp.

**Tóm lại**, các lựa chọn mở như Qwen-Image, Z-Image-Turbo, Taiyi-XL và hệ Lumina có thể thay thế gần như mọi tính năng của Grok Edit Image với chất lượng tương đương. Chúng đều có sẵn để tải về và tích hợp (qua Hugging Face, GitHub, ComfyUI, API miễn phí)[[26]](https://www.alibabacloud.com/blog/introducing-qwen-image-novel-model-in-image-generation-and-editing_602447#:~:text=The%20model%20is%20now%20open,technical%20report%20is%20also%20available)[[14]](https://huggingface.co/IDEA-CCNL/Taiyi-Stable-Diffusion-XL-3.5B#:~:text=%2A%20Train%20Code%20%E8%AE%AD%E7%BB%83%E4%BB%A3%E7%A0%81%EF%BC%9Ahttps%3A%2F%2Fgithub.com%2FIDEA,XL). Bạn có thể thử chạy local với diffusers hoặc dùng các web UI có sẵn. Các mô hình này hỗ trợ inpainting/chỉnh sửa chi tiết (qua ControlNet, SAM/LaMa) và không hạn chế NSFW, đáp ứng yêu cầu khắt khe của bạn.

**Nguồn tham khảo:** Thông tin lấy từ tài liệu kỹ thuật chính thức của Alibaba và Tencent[[1]](https://www.alibabacloud.com/blog/introducing-qwen-image-novel-model-in-image-generation-and-editing_602447#:~:text=Alibaba%20released%20Qwen,rendering%20and%20precise%20image%20editing)[[9]](https://huggingface.co/Tencent-Hunyuan/HunyuanDiT#:~:text=We%20present%20Hunyuan,art%20in), trang Hugging Face của các mô hình[[3]](https://huggingface.co/Qwen/Qwen-Image-Edit#:~:text=,addition%2C%20deletion%2C%20and%20modification%20of)[[4]](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=parameters,variants), cũng như hướng dẫn và phân tích cộng đồng (CSDN, blog kỹ thuật)[[8]](https://huggingface.co/alibaba-pai/Z-Image-Turbo-Fun-Controlnet-Union#:~:text=,stronger%20control%20and%20better%20detail)[[19]](https://github.com/geekyutao/Inpaint-Anything#:~:text=TL%3B%20DR%3A%20Users%20can%20select,Replace%20Anything). Các ví dụ ảnh được embed phía trên nhằm minh họa chất lượng đầu ra của các mô hình tiêu biểu.

[[1]](https://www.alibabacloud.com/blog/introducing-qwen-image-novel-model-in-image-generation-and-editing_602447#:~:text=Alibaba%20released%20Qwen,rendering%20and%20precise%20image%20editing) [[7]](https://www.alibabacloud.com/blog/introducing-qwen-image-novel-model-in-image-generation-and-editing_602447#:~:text=Alibaba%20released%20Qwen,leading%20model%20in%20the%20field) [[26]](https://www.alibabacloud.com/blog/introducing-qwen-image-novel-model-in-image-generation-and-editing_602447#:~:text=The%20model%20is%20now%20open,technical%20report%20is%20also%20available) Introducing Qwen-Image: Novel Model in Image Generation and Editing - Alibaba Cloud Community

<https://www.alibabacloud.com/blog/introducing-qwen-image-novel-model-in-image-generation-and-editing_602447>

[[2]](https://qwenlm.github.io/blog/qwen-image/#:~:text=%2A%20Superior%20Text%20Rendering%3A%20Qwen,visual%20realism%20during%20editing%20operations) [[24]](https://qwenlm.github.io/blog/qwen-image/#:~:text=We%20are%20thrilled%20to%20release,Chat%20and%20choose%20%E2%80%9CImage%20Generation%E2%80%9D) Qwen-Image: Crafting with Native Text Rendering | Qwen

<https://qwenlm.github.io/blog/qwen-image/>

[[3]](https://huggingface.co/Qwen/Qwen-Image-Edit#:~:text=,addition%2C%20deletion%2C%20and%20modification%20of) [[18]](https://huggingface.co/Qwen/Qwen-Image-Edit#:~:text=style%20transfer%2C%20allowing%20overall%20pixel,addition%2C%20deletion%2C%20and%20modification%20of) [[21]](https://huggingface.co/Qwen/Qwen-Image-Edit#:~:text=import%20os%20from%20PIL%20import,Image%20import%20torch) Qwen/Qwen-Image-Edit · Hugging Face

<https://huggingface.co/Qwen/Qwen-Image-Edit>

[[4]](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=parameters,variants) [[5]](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=%2A%20Z,tuning%20and%20custom%20development) [[6]](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=match%20at%20L123%20Creative%20Image,imaginative%20and%20flexible%20image%20transformations) [[23]](https://huggingface.co/Tongyi-MAI/Z-Image-Turbo#:~:text=Install%20the%20latest%20version%20of,diffusers%2C%20use%20the%20following%20command) Tongyi-MAI/Z-Image-Turbo · Hugging Face

<https://huggingface.co/Tongyi-MAI/Z-Image-Turbo>

[[8]](https://huggingface.co/alibaba-pai/Z-Image-Turbo-Fun-Controlnet-Union#:~:text=,stronger%20control%20and%20better%20detail) alibaba-pai/Z-Image-Turbo-Fun-Controlnet-Union · Hugging Face

<https://huggingface.co/alibaba-pai/Z-Image-Turbo-Fun-Controlnet-Union>

[[9]](https://huggingface.co/Tencent-Hunyuan/HunyuanDiT#:~:text=We%20present%20Hunyuan,art%20in) [[10]](https://huggingface.co/Tencent-Hunyuan/HunyuanDiT#:~:text=Model%20,0B%20DialogGen) [[11]](https://huggingface.co/Tencent-Hunyuan/HunyuanDiT#:~:text=Image) [[12]](https://huggingface.co/Tencent-Hunyuan/HunyuanDiT#:~:text=evaluate%20data%20for%20iterative%20model,art%20in) [[25]](https://huggingface.co/Tencent-Hunyuan/HunyuanDiT#:~:text=Welcome%20to%20Tencent%20Hunyuan%20Bot%2C,you%20desire%2C%20all%20for%20free) Tencent-Hunyuan/HunyuanDiT · Hugging Face

<https://huggingface.co/Tencent-Hunyuan/HunyuanDiT>

[[13]](https://huggingface.co/IDEA-CCNL/Taiyi-Stable-Diffusion-XL-3.5B#:~:text=Diffusion) [[14]](https://huggingface.co/IDEA-CCNL/Taiyi-Stable-Diffusion-XL-3.5B#:~:text=%2A%20Train%20Code%20%E8%AE%AD%E7%BB%83%E4%BB%A3%E7%A0%81%EF%BC%9Ahttps%3A%2F%2Fgithub.com%2FIDEA,XL) IDEA-CCNL/Taiyi-Stable-Diffusion-XL-3.5B · Hugging Face

<https://huggingface.co/IDEA-CCNL/Taiyi-Stable-Diffusion-XL-3.5B>

[[15]](https://huggingface.co/Alpha-VLLM/Lumina-Image-2.0#:~:text=Image%3A%20Lumina,more%20information%2C%20visit%20our%20GitHub) [[22]](https://huggingface.co/Alpha-VLLM/Lumina-Image-2.0#:~:text=import%20torch%20from%20diffusers%20import,Lumina2Pipeline) Alpha-VLLM/Lumina-Image-2.0 · Hugging Face

<https://huggingface.co/Alpha-VLLM/Lumina-Image-2.0>

[[16]](https://huggingface.co/api/resolve-cache/models/neta-art/Neta-Lumina/65c078e75781946e9f3423fe41457adae3eba3cd/README-ZH.md?download=true&etag=%22b1c5df428e8a54e4cdc35dcbe8749305a735fb47%22#:~:text=%EF%B8%8F%20%20Your%20browser%20does,%E6%95%B0%E6%8D%AE%E9%9B%86%EF%BC%9A%E7%B2%BE%E9%80%89%E6%95%B0%E5%8D%81%E4%B8%87%E5%BC%A0%E9%AB%98%E8%B4%A8%E9%87%8F%E5%8A%A8%E6%BC%AB%E9%A3%8E%E6%A0%BC%E5%9B%BE%E5%83%8F) [[17]](https://huggingface.co/api/resolve-cache/models/neta-art/Neta-Lumina/65c078e75781946e9f3423fe41457adae3eba3cd/README-ZH.md?download=true&etag=%22b1c5df428e8a54e4cdc35dcbe8749305a735fb47%22#:~:text=%E4%BD%BF%E7%94%A8%E6%8C%87%E5%8D%97%EF%BC%88How%20to%20Use%EF%BC%89%20%E8%AF%A5%E6%A8%A1%E5%9E%8B%E9%87%87%E7%94%A8%E5%9F%BA%E4%BA%8E%20Diffusion,safetensors%EF%BC%88%E7%BA%A6%20300%20MB%EF%BC%89) huggingface.co

<https://huggingface.co/api/resolve-cache/models/neta-art/Neta-Lumina/65c078e75781946e9f3423fe41457adae3eba3cd/README-ZH.md?download=true&etag=%22b1c5df428e8a54e4cdc35dcbe8749305a735fb47%22>

[[19]](https://github.com/geekyutao/Inpaint-Anything#:~:text=TL%3B%20DR%3A%20Users%20can%20select,Replace%20Anything) [[20]](https://github.com/geekyutao/Inpaint-Anything#:~:text=News) GitHub - geekyutao/Inpaint-Anything: Inpaint anything using Segment Anything and inpainting models.

<https://github.com/geekyutao/Inpaint-Anything>