Report

\_Nên cài đặt comfyUI manager để tiện cho việc install các model và custom node

\_Queue prompt: Bắt đầu khởi chạy một workflow

\_Load model diffusion: Load model có sẵn để chạy. Làm node đầu vào, output là Clip, Model, Vae.

\_Text prompt node: Node này được nhận input là output của model của node Load model diffusion.Node để nhập text và tạo hình ảnh. Có thể thiết kế nhiều text prompt. Ví dụ node: positive (những thứ cần xuất hiện trong ảnh), negative (những thứ không muốn xuất hiện trong ảnh). Khối text prompt node nối trực tiếp với KSampler để lấy các thông số

\_Khối KSampler: Một phần quan trọng giúp lấy mẫu (sampling) từ mô hình, thực hiện quá trình biến đổi từ dữ liệu noise ban đầu sang ảnh hoàn chỉnh dựa trên prompt.

\_Các thông số:

+ seed (như random seed của thư viện numpy)

+control\_after\_generate( cho phép tinh chỉnh hình ảnh đã tạo sau quá trình sinh (generate) ban đầu). Các chế độ fixed, randomize, increment, decrement

+steps: số thực hiện trong quá trình tạo ảnh. Số bước càng cao thì ảnh sẽ càng chi tiết, nhưng tốn thời gian hơn.

+cfg: điều chỉnh mức độ mà mô hình bám sát vào prompt so với các yếu tố ngẫu nhiên.

+sample\_name: Phương pháp lấy mẫu xác định cách thức mà mô hình chuyển đổi noise thành ảnh hoàn chỉnh

+scheduler: thuật toán điều khiển cách thức mà noise giảm dần qua mỗi bước

+denoise: điều chỉnh cường độ noise giảm trong mỗi bước sampling, ảnh hưởng đến độ sắc nét và mượt mà của hình ảnh.

\_Node Empty Latent: Điều chỉnh size ảnh và số lượng ảnh muốn tạo ra mỗi lần chạy workflow

\_Node Vae Decode: chuyển đổi các vector latents thành hình ảnh có thể nhìn thấy được

\_Node save image: Để hiển thị hình ảnh

WorkFlow

Mỗi workflow cơ bản hoặc phức tạp nhất đều bao gồm các thành phần sau:

Node load checkpoint: output: model, clip, vae

Node Clip text encode:

+input: clip,

+output: conditioning

Node Vae decode:

+input: samples, vae

+ output: image

Node KSample:

+Input: model, positive, negative, latent\_image

+Output:latent

Node Empty latent: output: latent\_image

Node save image: input: image

\*Tạo ảnh bằng model stable diffusion kết hợp với lora:

\_Trong mỗi workflow nên thêm vào một node lora để gpu nhẹ trong việc tính toán và train hơn. Để một workflow tạo ra được một object cụ thể, ví dụ như người thì phải train lora với các mẫu ảnh chụp từ thực tế. Ví dụ như muốn chỉnh sửa hoặc làm đẹp cho một người tên ABC thì phải có một tập data về người đó train cho lora, đặt tên tạm là Lora\_ABC và load vào phần lora trên folder local của PC. Khi đó AI mới có thể tạo ra một ảnh với người trong ảnh là đối tượng khách hangf.

\*Train LoRa:

+Thư viện cần: torch, diffusers, bitsandbytes

+Model dùng để train: Dreambooth for Diffusers, Kohya’s LoRA Trainer

Huấn luyện từ tập dataset:

python Lora\_ABC--dataset\_path /path/to/dataset --output\_dir /path/to/output --learning\_rate 1e-4 --batch\_size 4 --num\_steps 10000 --rank 8

rank càng lớn thì việc training sẽ chính xác nhưng lại khá nặng cho việc tính toán. Rank thấp thì độ chính xác của Lora sẽ bị giảm

Image to Image: mode pose

Tạo ra một ảnh mới dựa vào dáng của mục tiêu ảnh đầu vào

Các node cần thiết:

+Tất cả các thành phần cơ bản của một workflow.

+Thay thế node Empty Latent bằng Một workflow con: gồm Node load image, Node VAE encode để mã hóa image, Node Pose để lấy dáng của mục tiêu trong ảnh.

+Một node mới LoadControlNet model.

+Node ApplyControlNet được chèn vào giữa flow của Clip text prompt và KSampler.

Một số node không có sẵn sẽ được install bằng manager. Vào manager chọn Custom Nodes Manager, tim các node cần thiết.

Các thông số có thể ảnh hưởng đến đầu ra của ảnh:

+Strength trong node ApplyControlNet: càng giảm thông số thì mức độ lắng nghe của model sẽ càng lớn. Nếu strength gần với 1 thì hầu như chỉ tạo ra cùng một ảnh so với ảnh gốc mà không nghe theo chỉ thị của prompt text (nên chọn 0.4-0.6)

+Denoise trong Ksampler: thông số càng lớn thì việc sáng tạo của model càng cao nhưng dễ bị sai. Nhưng Denoise quá bé thì cũng tạo ra một khuôn mặt gần giống với ảnh gốc (thông số khoảng 0.3 sẽ cho ảnh giống với người thật hơn, có thể điều chỉnh khoảng từ 0.25-0.4)

Trong quá trình install các custome node. Nếu install node DWPose (Node Pose) ko sử dụng được thì có thể thay thế bằng cách sử dụng Open Pose