



# STABLE DIFFUSION CHO NGƯỜI MỚI BẮT ĐẦU

Ngày cập nhật 27/5/2024

## MỤC LỤC

MỤC LỤC.....	2
Lời nói đầu .....	3
PHẦN I - Giới thiệu .....	3
Chương 1. Giới thiệu.....	3
Chương 2. Tổng quan.....	3
Chương 3. Cách bắt đầu sử dụng Stable diffusion .....	9
PHẦN II - Hướng dẫn cho người mới bắt đầu .....	22
Chương 4. Cách xây dựng Prompt hiệu quả.....	22
Chương 5. Thông số hình ảnh .....	28
Chương 6. Mô hình .....	33
Chương 7. Inpainting.....	44
Chương 8. Chuyển image-to-image .....	53
Chương 9. Cách điệu hóa hình ảnh .....	58
Chương 10. Khắc phục các vấn đề thường gặp.....	63
PHẦN III - Các chủ đề trung cấp .....	69
Chương 11. Kỹ thuật Prompt nâng cao .....	69
Chương 12. Công cụ nâng cấp hình ảnh AI .....	74
Chương 13. Quy trình làm việc hoàn chỉnh từ đầu đến cuối.....	77
Chương 14. Sử dụng VAE để cải thiện đôi mắt và khuôn mặt .....	85
Chương 15. Embeddings .....	88
Chương 16. Mô hình V2 .....	93
Chương 17. Negative prompt .....	96
PHẦN IV - Các chủ đề nâng cao .....	113
Chương 18. Độ sâu của hình ảnh .....	113
Chương 19. Sử dụng Dreambooth để đưa mọi thứ vào sự Stable diffusion.....	120
PHẦN IV – HƯỚNG DẪN THỰC HÀNH .....	127
THÔNG TIN LIÊN HỆ.....	131
TÀI LIỆU THAM KHẢO .....	131

## LỜI NÓI ĐẦU

Stable Diffusion là phần mềm tạo hình ảnh AI vô cùng mạnh mẽ mà bạn có thể chạy ngay trên máy tính cá nhân của mình. Stable Diffusion có khả năng tạo ra gần như mọi thứ, miễn là đã có ai đó huấn luyện nó để thực hiện.

Hình ảnh bạn tạo ra có thể được sử dụng cho mọi mục đích, tùy thuộc vào giấy phép của model được sử dụng. Việc chúng có "thuộc về bạn" về mặt pháp lý hay không phụ thuộc vào luật pháp hiện hành và thường là không rõ ràng. Tôi cũng như bất kỳ ai liên quan đến Stable Diffusion hoặc các model của nó không chịu trách nhiệm về bất kỳ điều gì bạn tạo ra, và bạn được cấm tạo ra nội dung bất hợp pháp hoặc có hại ảnh hưởng đến cá nhân hoặc tập thể cũng như lợi ích quốc gia ...

Tài liệu này vẫn còn ở mức cơ bản và sẽ tiếp tục được cập nhật bổ sung theo thời gian. Hi vọng nó vẫn hữu ích cho các bạn.

## PHẦN I - GIỚI THIỆU

### Chương 1. Giới thiệu

Chúc mừng bạn đã thực hiện bước đầu tiên để học AI Stable diffusion. Nó thực sự là một công cụ hỗ trợ cho quy trình sáng tạo nghệ thuật mà trước đây không thể thực hiện được bằng máy tính.

Mục tiêu của tài liệu này là hướng dẫn những người chưa biết gì về Stable diffusion trở thành chuyên gia sử dụng nó. Khi đọc xong tài liệu này, bạn sẽ có kiến thức về những gì Stable diffusion có thể làm được. Trước tiên, tài liệu sẽ cung cấp cho bạn cái nhìn tổng quan về Stable diffusion là gì và nó có thể làm gì cho bạn. Sau đó, sẽ trình bày các cách khác nhau để bạn có thể bắt đầu sử dụng Stable diffusion. Tài liệu này sẽ hướng dẫn cho bạn tất cả những gì bạn cần để tạo ra các tác phẩm nghệ thuật với quy trình làm việc từng bước. Bạn sẽ tìm hiểu về các kỹ thuật khác nhau bao gồm: 1) Chuyển văn bản thành hình ảnh (text-to-image); 2) Chuyển hình ảnh thành hình ảnh (image-to-image); 3) Cách sử dụng công cụ Inpainting để tái tạo một vùng nhỏ của hình ảnh; 4) Cách sử dụng trình nâng cấp (Upscale) để thêm chi tiết và tạo ra các bản in có độ phân giải cao.

Trong các chủ đề nâng cao, bạn sẽ tìm hiểu về cách tạo mô hình Stable diffusion của riêng mình với Dreambooth. Bạn cũng sẽ tìm hiểu về nhiều công cụ ít được sử dụng như image-to-image và chiều sâu của hình ảnh.

### Chương 2. Tổng quan

Chương này dành cho những người muốn bắt đầu sử dụng Stable diffusion ngay lập tức.

#### 2.1. Stable diffusion là gì?

Stable diffusion là một mô hình tạo *hình ảnh từ đầu vào văn bản*. Giả sử nếu bạn muốn tạo hình ảnh ngôi nhà bánh gừng, bạn có thể đặt Prompt:

*gingerbread house, diorama, in focus, white background, toast, crunch cereal*

*ngôi nhà bánh gừng, tăm sâu, tập trung, nền trắng, bánh mì nướng, ngũ cốc giòn*

Mô hình AI sẽ tạo ra các hình ảnh phù hợp với Prompt hiển thị bên dưới.



### Ưu điểm :

- Do là mã nguồn mở vì vậy SD được phát triển gần như không có một giới hạn nào. Có muôn vàn các công cụ các tài nguyên giúp chúng ta tạo ra bức ảnh ưng ý. Người dùng có thể tự do training model theo mong muốn và yêu cầu cá nhân.
- Stable Diffusion sở hữu những công cụ mạnh mẽ mà không một / chưa một công cụ nào có được, đó là controlnet (Công cụ giúp ta có thể kiểm soát hoàn hảo SD để ra được kết quả ưng ý ).
- Khả năng tạo không giới hạn nội dung cũng là một ưu điểm rất lớn của SD so với công cụ khác.

### Nhược điểm :

- Do được tự do đào tạo và training model vì vậy các model của SD không có sự thống nhất về từ khoá cũng như cách sử dụng. Dẫn tới sự khó khăn cho người dùng khi tiếp cận với những model tuned khác nhau.
- Sự không giới hạn về nội dung dẫn tới sự lạm dụng của các cá nhân/ tổ chức vào các công việc không được phép.
- Sự phát triển của nhiều phiên bản và nhiều công cụ khác nhau dẫn đến sự khó khăn cho người dùng tiếp cận và học tập SD một cách đầy đủ và có hệ thống.
- Các model base được training khá sơ sài vì vậy chủ yếu người dùng sẽ cần phải sử dụng các model tuned, tuy nhiên do các model được train từ các tổ chức / cá nhân đơn lẻ dẫn đến sự thiếu nhất quán cùng thiếu sót về nhiều mảng trong cùng một model. Việc tìm kiếm một model được đào tạo bài bản – chất lượng cao là khá khó khăn đối với người mới.

## 2.2. Cách bắt đầu sử dụng Stable diffusion

Cách dễ nhất để bắt đầu tạo hình ảnh là sử dụng trình tạo trực tuyến miễn phí, bạn có thể truy cập vào các trang như:

<https://civitai.com/>

<https://tensor.art/>

<https://shakker.ai/>

Hạn chế của việc sử dụng trang web miễn phí là chức năng khá hạn chế. Nếu bạn muốn tìm hiểu sâu hơn về Stable diffusion, bạn nên chuyển sang sử dụng giao diện GUI (Giao diện đồ họa người dùng), trong tài liệu này sẽ giới thiệu và sử dụng giao diện **AUTOMATIC1111**. Bạn có thể cài đặt nó trên Windows, Mac hoặc sử dụng trên dịch vụ Cloud. Trong Chương 3, sẽ đề cập đến các tùy chọn cài đặt này.

### 2.3. Cách xây dựng Prompt tốt

Một cách nhanh nhất để tạo hình ảnh chất lượng cao là sử dụng lại Prompt hiện có. Nhược điểm là bạn có thể không hiểu tại sao Prompt đó lại tạo ra hình ảnh chất lượng cao. Bạn có thể truy cập vào các trang web tại mục 2.2 ở trên, tìm lấy cho mình một hình ảnh yêu thích, và tham khảo cách viết Prompt của hình ảnh đó. Nhược điểm là nó có thể giống như mò kim đáy bể để tìm ra Prompt chất lượng cao đáp ứng yêu cầu của bạn. Ngoài ra, đôi khi mọi người không đưa Prompt thực tế vào.

Hãy coi Prompt là điểm bắt đầu và sửa đổi cho phù hợp với nhu cầu của bạn.

### 2.4. Quy tắc xây dựng Prompt tốt

Để thành thạo việc sử dụng Stable diffusion, bạn không thể chỉ sao chép Prompt của người khác. Bạn cần biết cách xây dựng nó. Hai quy tắc cơ bản tạo Prompt là: (1) Hãy chi tiết và cụ thể và (2) sử dụng các keyword mạnh mẽ.

**Hãy chi tiết và cụ thể.** Mặc dù AI có những tiến bộ nhảy vọt nhưng Stable diffusion vẫn không thể đọc được suy nghĩ của bạn. Bạn cần mô tả hình ảnh của mình càng chi tiết càng tốt.

Ví dụ: nếu bạn muốn tạo hình ảnh một người phụ nữ đang ngồi trên ghế, thay vì sử dụng Prompt

*A woman sitting on a chair*

*Một người phụ nữ ngồi trên ghế*

Bạn nên sử dụng một cái gì đó như:

*a young gorgeous woman with long blonde hair sitting on a 19th-century antique chair,  
beautiful eyes, symmetric face, dedicated modern clothing, photo*

*Một người phụ nữ trẻ đẹp với mái tóc dài màu vàng ngồi trên chiếc ghế cổ từ thế kỷ 19,  
đôi mắt đẹp, khuôn mặt cân đối, bộ quần áo hiện đại chuyên dụng, ảnh*

Hãy xem sự khác biệt rõ rệt về chất lượng hình ảnh giữa hai Prompt bên dưới. Vì vậy, hãy làm theo Prompt của bạn trước khi viết về Stable diffusion!



**Sử dụng các keyword mạnh mẽ.** Một số keyword chỉ mạnh hơn những keyword khác. Ví dụ là:

- Tên người nổi tiếng (ví dụ: Emma Watson)
- Tên nghệ sĩ (ví dụ: van Gogh)
- Thể loại (ví dụ: minh họa, tranh vẽ, ảnh chụp)

Sử dụng chúng một cách cẩn thận có thể giúp tạo hình ảnh đúng theo mong muốn. Chúng ta sẽ tìm hiểu thêm về cách xây dựng Prompt và keyword ví dụ trong Chương 4.

## 2.5. Thông số hình ảnh

Hầu hết các trình tạo Stable diffusion đều cho phép bạn thay đổi một số thông số nhất định. Dưới đây là một số điều quan trọng bạn nên biết.

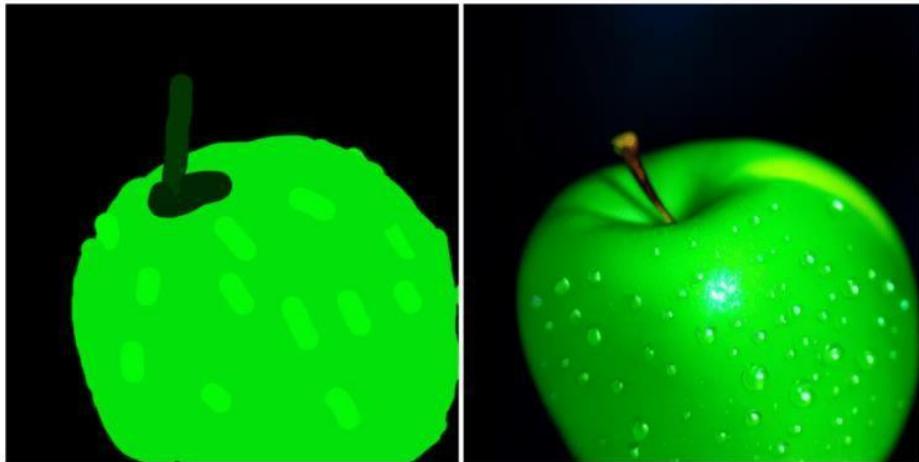
- **Image size:** Kích thước của hình ảnh đầu ra. Kích thước tiêu chuẩn cho Stable Diffusion v1 là  $512 \times 512$  pixel. Việc thay đổi nó thành kích thước dọc hoặc ngang có thể có tác động lớn đến hình ảnh được tạo ra. Ví dụ: sử dụng kích thước dọc nếu bạn muốn tạo hình ảnh toàn thân.
- **Sampling steps:** Sử dụng ít nhất là 20 bước và bạn có thể tăng lên nếu bạn thấy hình ảnh mờ.
- **CFG scale:** Giá trị thường dùng là 7 và bạn có thể tăng lên nếu bạn muốn hình ảnh bám sát theo Prompt nhiều hơn.
- **Giá trị Seed:** -1 tạo ra một hình ảnh ngẫu nhiên. Chỉ định một giá trị nếu bạn muốn có cùng một hình ảnh.
- **Batch size:** Số lượng hình ảnh được tạo ra mỗi lần.

Bạn phải luôn tạo nhiều hình ảnh, thường sẽ cần tạo 2-4 hình ảnh cùng một lúc khi thực hiện những thay đổi lớn đối với Prompt để có thể nhanh chóng tìm kiếm các kết hợp

keyword phù hợp. Tôi sẽ tạo ra 4-8 lần khi thực hiện những thay đổi nhỏ để tăng cơ hội thấy thứ gì đó có thể sử dụng được. Một số Prompt chỉ có tác dụng trong một nửa thời gian hoặc ít hơn.

## 2.6. Chuyển hình ảnh thành hình ảnh (image-to-image)

Lấy hình ảnh và Prompt làm đầu vào. Bạn có thể hướng dẫn việc tạo hình ảnh không chỉ bằng Prompt mà còn bằng hình ảnh. Dưới đây là một ví dụ về việc sử dụng hình ảnh để biến một bức vẽ phác thảo thành một bức hình đẹp hơn.



Sử dụng Prompt “anh quả táo xanh hoàn hảo có cuống, giọt nước, ánh sáng ẩn tượng” và hình ảnh bên trái làm đầu vào, img2img sẽ tạo ra hình ảnh bên phải.

Trên thực tế, bạn có thể coi chuyển văn bản thành hình ảnh là một trường hợp đặc biệt của chuyển đổi image-to-image: Nó chỉ đơn giản là chuyển image-to-image với hình ảnh đầu vào có nhiều ngẫu nhiên. Img2img là một kỹ thuật mạnh mẽ nhưng lại bị đánh giá thấp. Chúng ta sẽ khám phá thêm tính năng của nó trong Chương 8.

## 2.7. Khôi phục khuôn mặt (restoring faces)

Khi bạn nhìn thấy những hình ảnh AI tuyệt đẹp được chia sẻ trên mạng xã hội, rất có thể chúng đã trải qua một loạt bước xử lý hậu kỳ. Stable Diffusion không giỏi trong việc tạo khuôn mặt, khuôn mặt tạo ra hay bị lỗi. Để khắc phục lỗi, chúng ta có thể sử dụng mô hình AI hình ảnh như **CodeFormer** tới **restoring faces** mà AUTOMATIC1111 đã tích hợp sẵn để hỗ trợ. Ngoài ra còn có bản cập nhật cho model v1.4 và v1.5 để sửa lỗi mang tên **VAE** (Bộ mã hóa tự động biến thể).



## 2.8. Sửa các đồi tượng bằng Inpainting

Không dễ để có được hình ảnh như ý chỉ trong một lần tạo. Một cách tiếp cận tốt hơn là sửa chữa những khiếm khuyết bằng **inpainting**. Dưới đây là ví dụ về hình ảnh trước và sau khi inpaint. Sử dụng Prompt ban đầu để vẽ có hiệu quả tốt đến 90%.



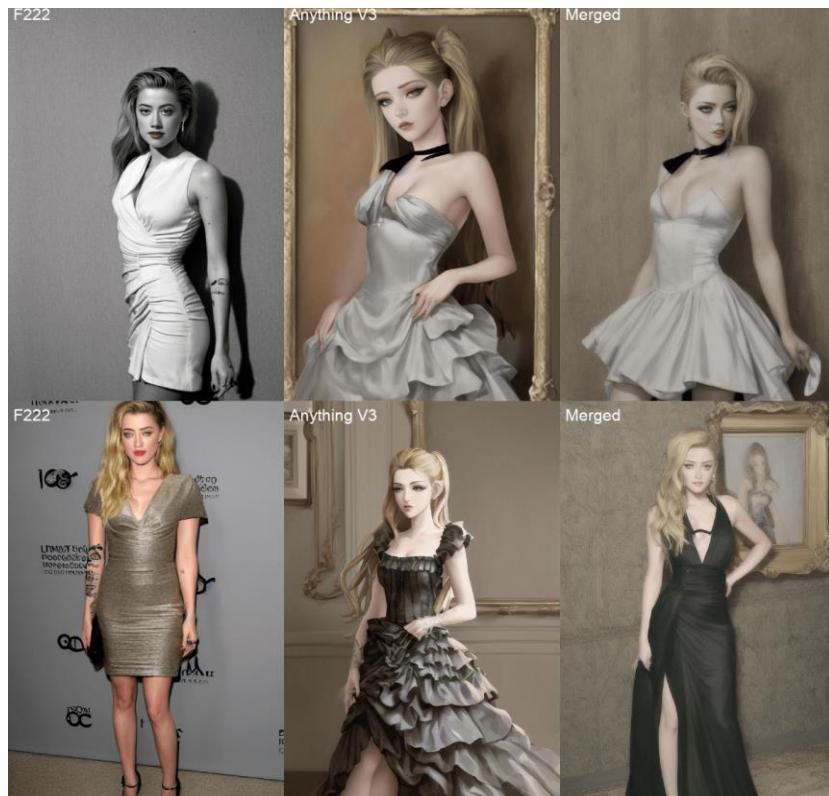
## 2.9. Models

Các mô hình chính thức được phát hành bởi Stability AI (Công ty chính phát triển mô hình) và các đối tác của họ được gọi là mô hình cơ sở (**base models**). Một số ví dụ về mô hình cơ sở là Stable diffusion 1.4, 1.5, 2.0 và 2.1.

Lưu ý rằng các checkpoint ở định dạng **.safetensors** là an toàn để sử dụng trong khi **.ckpt** có thể chứa virus, vì vậy hãy cẩn thận. Ngoài ra, khi chọn model, bạn có thể có sự lựa chọn

giữa fp32, fp16 và pruned. Tất cả đều tạo ra cùng một hình ảnh trong một biên độ lỗi nhỏ, vì vậy chỉ cần chọn tệp nhỏ nhất (pruned-fp16).

**Mô hình tùy chỉnh được sửa đổi** từ các mô hình cơ sở. Hiện nay hầu hết các mô hình đều được train từ v1.4 hoặc v1.5. Chúng được đào tạo với dữ liệu bổ sung để tạo ra hình ảnh của phong **cách hoặc chủ đề cụ thể**. Dưới đây là so sánh 5 mô hình khác nhau. Cũng có thể dễ dàng hợp nhất hai mô hình để tạo phong cách giữa hai mô hình bất kỳ.



Hai nhóm mô hình cơ sở chính là **v1** và **v2**. Mô hình v1 là 1.4 và 1.5, mô hình v2 là 2.0 và 2.1. Các hình ảnh từ v2 không hoàn toàn tốt hơn v1. Bạn có thể sử dụng cả 2 mô hình để tìm ra được mô hình phù hợp với nhu cầu của mình.

### Chương 3. Cách bắt đầu sử dụng Stable diffusion

Trong Chương này, sẽ hướng dẫn cách cài đặt Stable Diffusion trong Window, Mac. Nếu bạn không có máy tính đáp ứng yêu cầu hệ thống có một tùy chọn khác là sử dụng trên phiên bản Colab của Google.

Trong tài liệu này, sẽ sử dụng GUI (giao diện) **AUTOMATIC1111** để minh họa. Đây là GUI miễn phí và rất phổ biến dành cho Stable diffusion đi kèm với nhiều tính năng nâng cao.

#### 3.1. GUI AUTOMATIC1111

Bạn cần biết AUTOMATIC1111 nếu muốn trở thành “chuyên gia” về Stable diffusion.

Hầu hết người dùng sử dụng một **GUI** (Giao diện người dùng đồ họa) để sử dụng

Stable diffusion. AUTOMATIC1111 là một trong những GUI đầu tiên được phát triển cho Stable diffusion. AUTOMATIC1111 có nhiều tính năng: Bạn có thể sử dụng tính năng chuyển văn bản thành hình ảnh (text-to-image), hình ảnh thành hình ảnh (image-to-image), nâng cấp (upscale), chạy và huấn luyện tất cả các mô hình tùy chỉnh trong GUI này.

### 3.2. Các tùy chọn để chạy AUTOMATIC1111

Có 3 tùy chọn để chạy AUTOMATIC1111 trong Chương này bao gồm: Google Colab, Windows và Mac.

Google Colab là một dịch vụ do Google cung cấp. Lợi ích của việc chạy trên Colab là thiết lập đơn giản và không cần máy tính có cấu hình phần cứng cao, bạn chỉ cần có máy tính kết nối Internet là được. Hạn chế của việc sử dụng Colab là thời gian khởi động chậm (vài phút) và có thể bị ngắt kết nối nếu bạn đang sử dụng tài khoản miễn phí. Nếu bạn quyết định sử dụng tùy chọn này, bạn cần nâng cấp lên phiên bản trả phí là Colab Pro. Hiện tại, \$10 sẽ cho phép bạn sử dụng 50 giờ. Đây được coi là lựa chọn rẻ nhất nếu bạn tính đến chi phí phần cứng.

Chạy trên Windows là một lựa chọn tốt nếu bạn đã có máy tính Windows có card đồ họa (GPU) với ít nhất 4 GB VRAM. Những lợi ích là miễn phí và thời gian khởi động nhanh. Hạn chế của việc chạy trên Windows là có thể khó cài đặt và cập nhật phần mềm nếu bạn không am hiểu về công nghệ.

Tùy chọn cuối cùng đang chạy Mac. Bạn sẽ cần Apple Silicon M1 hoặc M2. Việc thiết lập có độ khó vừa phải. Tốc độ thấp hơn hai lựa chọn còn lại. Một số tính năng trong AUTOMATIC1111 có thể không hoạt động tốt trên Mac.

### 3.3. Chạy trên Google Colab

Bạn có thể truy cập notebook (sổ ghi chép) của Google Colab để chạy AUTOMATIC1111 bằng URL sau.

[https://colab.research.google.com/github/sagiodev/stablediffusion\\_webui/blob/master/StableDiffusionUI\\_ngrok\\_sagiodev.ipynb](https://colab.research.google.com/github/sagiodev/stablediffusion_webui/blob/master/StableDiffusionUI_ngrok_sagiodev.ipynb)

**Lưu ý: Hiện tại Google colab chỉ chạy được đối với tài khoản Colab Pro, bạn cần phải trả phí tối thiểu 10\$/tháng.**

**Các ưu điểm khi sử dụng Colab Pro :**

- Có thể sử dụng bộ nhớ Ram cao cấp với bộ nhớ 51GB thay vì 12GB so với tài khoản free
- Được sử dụng bộ nhớ ổ cứng colab lớn hơn nhiều so với free, 160Gb thay vì 70GB
- Sử dụng được đồng thời nhiều phiên sử dụng GPU trên một tài khoản
- Có thể sử dụng điểm điện toán để sử dụng với các GPU tốt hơn (A100 - 45GB, V100 - 16GB, tốc độ nhanh hơn T4) để thực hiện các tác vụ phức tạp

hơn. ( Lưu ý : Hiện tại A100 khá hiện được sử dụng quá nhiều, trường hợp A100 không dư thì colab tự thay đổi kết nối sang V100, mọi người lưu ý trường hợp này tránh các thắc mắc không đáng có )

- Được mở khoá các giới hạn trong sử dụng của GPU free – Có thể sử dụng được các ứng dụng như Automatic1111, ComfyUI ...
- Được tăng thời gian free sử dụng hằng ngày khi hết điểm điện toán, có thể sử dụng liên tục 10-12h/ngày khi hết điểm điện toán

### **Một vài thủ thuật để sử dụng ColabPro hiệu quả nhất :**

#### **❖ Cách sử dụng tối ưu điểm điện toán hàng tháng**

Đối với 100 điểm điện toán hàng tháng, người dùng nên sử dụng tập trung vào các vụ mạnh như train dữ liệu, học tập. Và nên sử dụng với GPU A100 ( Khoảng 13 điểm/h ) là cách sử dụng đáng tiền nhất.

Trong trường hợp người dùng pro đang có phiên làm việc mà hết điểm điện toán, google vẫn cho người dùng có thể duy trì phiên làm việc tới 20h tiếp theo mà không cần phải mua thêm điểm. Vì vậy cách để sử dụng tối ưu nhất với ưu đãi này đó là người dùng lên sử dụng 90 điểm đầu với các tác vụ thông thường. Với 10 điểm cuối cùng người dùng nên kết nối với GPU A100 và dùng liên tiếp trong một ngày không được ngắt phiên để có thể sử dụng free A100 trong 20 tiếng (Tương đương 260 điểm điện toán ) tiếp theo.

#### **❖ Cách sử dụng giờ free hiệu quả:**

Sau khi hết điểm điện toán, người dùng không cần quá lo lắng. Tất cả các ưu đãi về Ram, ổ cứng, giới hạn về thời gian dùng, giới hạn các tác vụ vẫn còn. Đổi mới tài khoản pro hết điện toán, thời gian dùng miễn phí trong ngày có thể tới 10-12h/ngày, thời gian hồi lại trong vòng 10-12h.

#### **❖ Các thủ thuật bổ sung lưu ý khi dùng:**

Trường hợp tài khoản hết điện toán và cần có nhu cầu dùng GPU cao cấp: Trường hợp này lời khuyên đối với mọi người là không nên mua gói điện toán bổ sung của google colab, thay vì đó hãy chuyển sang tài khoản google khác và nâng cấp colabpro của tài khoản đó. Điều đó giúp bạn có thêm một tài khoản pro và vẫn kèm theo điểm điện toán cần thiết. Việc chia sẻ dùng chung dữ liệu drive giúp chúng ta hoàn toàn không cần thiết phải dùng chung một Colab.

Cách sử dụng song song nhiều phiên GPU. Về cơ bản Colab không cho sử dụng song song nhiều phiên với cùng chung một option với nhau. Tuy nhiên ngược lại người dùng có thể chạy song song nhiều phiên nếu sử dụng các phiên với các GPU khác nhau hoặc sử dụng chung GPU T4 và khác nhau ở lựa chọn High Ram hoặc Ram thường

Tiếp theo, để sử dụng Stable diffusion trên Colab chúng ta thực hiện theo các bước sau:

**Bước 1.** Mở sổ ghi chép Colab. Bạn sẽ thấy sổ ghi chép có ô thứ hai như bên dưới.

```

Stable diffusion
Follow the URL link https://xxxxx.ngrok.io to launch the app if use ngrok.
Follow the URL link https://xxxxx.radio.live if not.

output_path: "AI_PICS"

username: "a"

password: "a"

NGROK: "Insert text here"

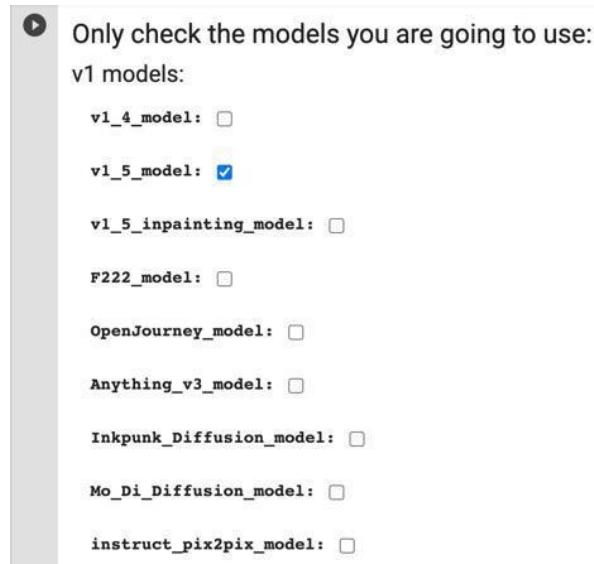
SAVE_IN_GOOGLE_DRIVE: 
CLEAR_SETUP_LOG: 
Only check the models you are going to use:
v1 models:
v1_4_model: 
v1_5_model: 
v1_5_inpainting_model: 
F222_model: 
OpenJourney_model: 
Anything_v3_model: 
Inkpunk_Diffusion_model: 
Mo_Di_Diffusion_model: 
instruct_pix2pix_model: 

```

**Bước 2.** Bạn có thể chỉnh sửa tên người dùng và mật khẩu mặc định nếu muốn. Bạn sẽ cần thông tin xác thực sau khi khởi động AUTOMATIC1111.

**Bước 3.** Chọn SAVE\_IN\_GOOGLE\_DRIVE nếu bạn muốn tự động lưu tất cả hình ảnh được tạo trong Google Drive. Nó sẽ được lưu trong thư mục được chỉ định trong out\_path. Bạn sẽ cần cấp quyền truy cập Google Drive nếu bạn chọn tùy chọn này.

**Bước 4.** Lựa chọn các mô hình bạn muốn tải. Nếu bạn là người dùng lần đầu, bạn nên chọn model v1.5.



**Bước 5.** Bạn có thể đưa các mô hình vào Google Drive của mình, chẳng hạn như các mô hình được tạo từ sổ ghi chép Dreambooth. Bạn có thể khai báo nhiều bao gồm nhiều mô hình tùy chỉnh, chỉ cần liệt kê tên file và phân tách chúng bằng dấu phẩy.

Custom models in Google drive (separate them with comma):

E.g. Dreambooth\_model/model.ckpt,Dreambooth\_model/model1.ckpt

custom\_model: "Insert text here"

**Bước 6.** Bấm vào nút Play ở bên trái để bắt đầu.

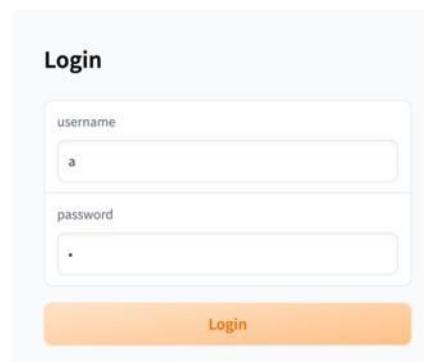
**Bước 7.** Quá trình khởi động sẽ hoàn tất trong vòng vài phút. Khi hoàn tất, bạn sẽ thấy thông báo như bên dưới.

```
WEBUI ARGUMENTS: --gradio-img2img color-sketch --xformers --share --gradio-auth a:a
LatentDiffusion: Running in eps-prediction mode
DiffusionWrapper has 859.52 M params.
Downloading: 100% 939k/939k [00:01<00:00, 868kB/s]
Downloading: 100% 512k/512k [00:01<00:00, 469kB/s]
Downloading: 100% 389/389 [00:00<00:00, 397kB/s]
Downloading: 100% 905/905 [00:00<00:00, 953kB/s]
Downloading: 100% 4.41k/4.41k [00:00<00:00, 4.31MB/s]
Calculating sha256 for /content/stable-diffusion-webui/models/Stable-diffusion/v1-5-pruned-emaonly.ckpt: c
Loading weights [cc6cb27103] from /content/stable-diffusion-webui/models/Stable-diffusion/v1-5-pruned-emac
Applying xformers cross attention optimization.
Textual inversion embeddings loaded(0):
Model loaded in 50.0s (12.5s create model, 36.8s load weights).
Running on local URL: http://127.0.0.1:7860
Running on public URL: https://941b8798-ba18-4fb4.gradio.live ←
```

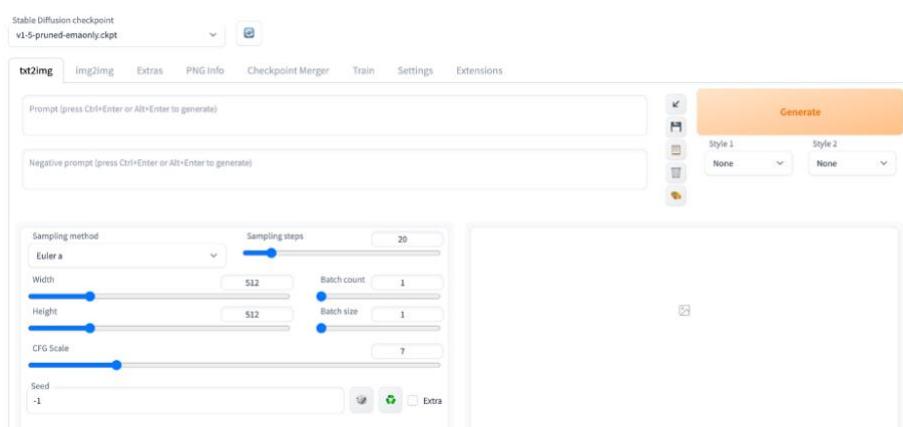
This share link expires in 72 hours. For free permanent hosting and GPU upgrades (NEW!), check out Spaces:

**Bước 8.** Sử dụng địa chỉ trên **gradio.live** để bắt đầu sử dụng AUTOMATIC1111.

**Bước 9.** Nhập tên người dùng và mật khẩu bạn đã thiết lập ở bước trên.

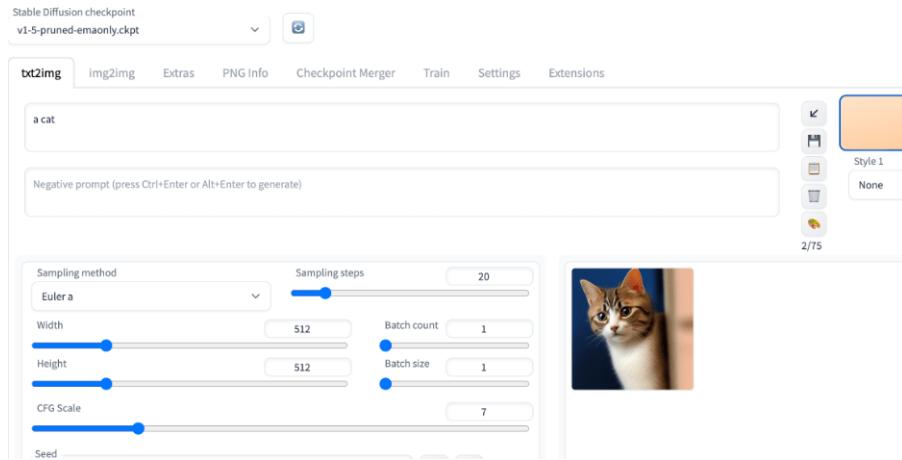


**Bước 10.** Bạn sẽ thấy GUI AUTOMATIC1111 sau khi đăng nhập.



Nhập “a cat” vào ô txt2img và nhấn **Generate** để tạo hình ảnh. Bạn sẽ thấy nó tạo ra hình

ảnh của một con mèo.



### 3.4. Chạy trên máy tính Windows

PC của bạn phải chạy Windows 10 trở lên với **card màn hình Nvidia rời** (GPU) có 4 GB VRAM trở lên.

**Bước 1: Cài đặt Python.** Bạn sẽ cần **Python 3.10.6** trở lên để chạy Stable diffusion. Chọn trình cài đặt windows 64-bit. Truy cập URL sau để tải xuống.

<https://www.python.org/downloads/release/python-3106/>

Files		
Version	Operating System	Description
Gzipped source tarball	Source release	
XZ compressed source tarball	Source release	
macOS 64-bit universal2 installer	macOS	for macOS 10.9 and later
Windows embeddable package (32-bit)	Windows	
Windows embeddable package (64-bit)	Windows	
Windows help file	Windows	
Windows installer (32-bit)	Windows	
Windows installer (64-bit)	Windows	Recommended

Mở trình cài đặt Python để bắt đầu cài đặt.



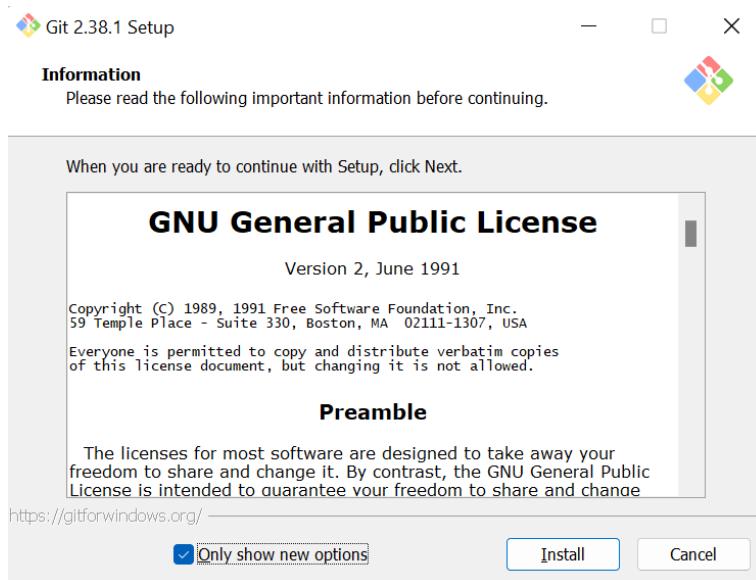
Tick vào lựa chọn "**Add Python 3.10 to PATH**". Nhấp chuột "**Install Now**" để

bắt đầu cài đặt. Nếu bạn gặp lỗi, rất có thể là do bạn đã cài đặt Python trước đó. Xóa mọi phiên bản Python đã cài đặt trước đó trước khi cài đặt lại 3.10.6.

**Bước 2: Cài đặt git.** [Git](https://git-scm.com/download/win) là một hệ thống quản lý mã nguồn. Bạn sẽ cần nó để cài đặt và cập nhật AUTOMATIC1111. Đi tới trang sau để tải xuống phiên bản windows.

<https://git-scm.com/download/win>

Mở trình cài đặt. Nhấp chuột **Install** để chấp nhận giấy phép và cài đặt phần mềm.

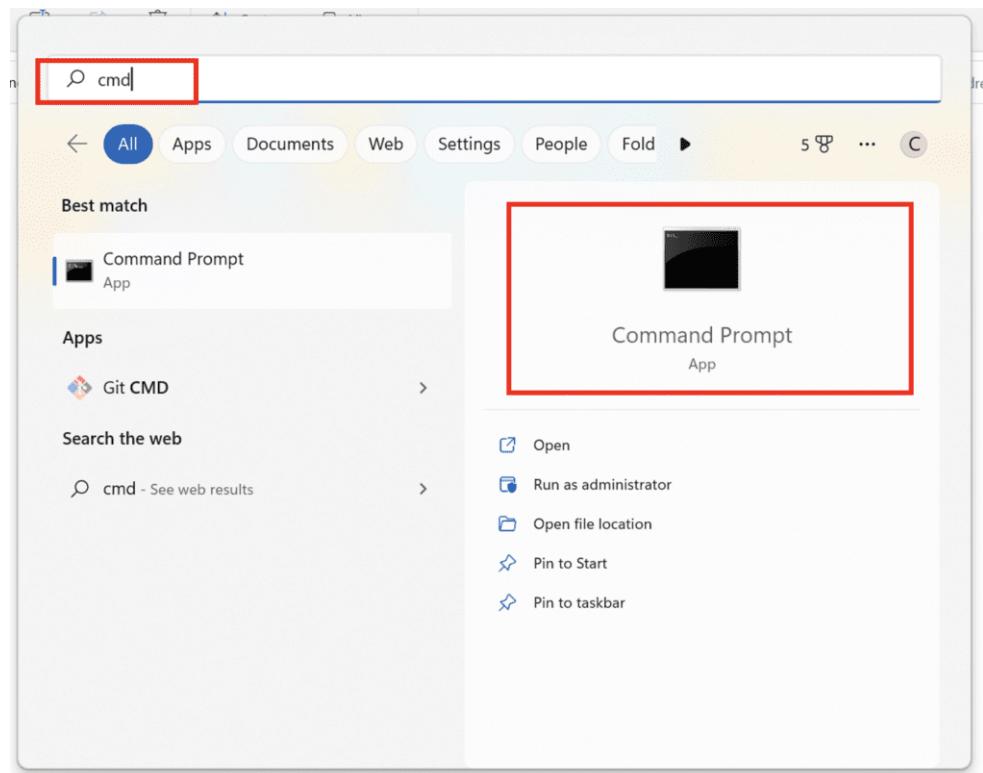


Làm theo hướng dẫn để hoàn tất cài đặt.

**Bước 3: Download AUTOMATIC1111.**

Trước hết, hãy chắc chắn rằng dung lượng trống trên ổ đĩa của bạn từ 20GB trở lên.

Nhấn phím Window > Gõ cmd.



Bấm vào **Command Prompt**. Gõ lệnh sau và nhấn Enter.

```
cd %userprofile%
```

Bạn sẽ thấy Prompt của bạn hiển thị một cái gì đó như C:\Users\YOUR\_USER\_NAME>. Tiếp theo, gõ lệnh sau và nhấn Enter để sao chép bộ cài AUTOMATIC1111.

```
git clone https://github.com/AUTOMATIC1111/stable-diffusion-webui.git
```

```
Microsoft Windows [Version 10.0.22000.1219]
(c) Microsoft Corporation. All rights reserved.

C:\Users\.....>git clone https://github.com/AUTOMATIC1111/stable-diffusion-webui.git
Cloning into 'stable-diffusion-webui'...
remote: Enumerating objects: 12200, done.
remote: Counting objects: 100% (3/3), done.
remote: Compressing objects: 100% (3/3), done.
remote: Total 12200 (delta 0), reused 1 (delta 0), pack-reused 12197
Receiving objects: 100% (12200/12200), 24.20 MiB | 7.15 MiB/s, done.

Resolving deltas: 100% (8527/8527), done.

C:\Users\.....>
```

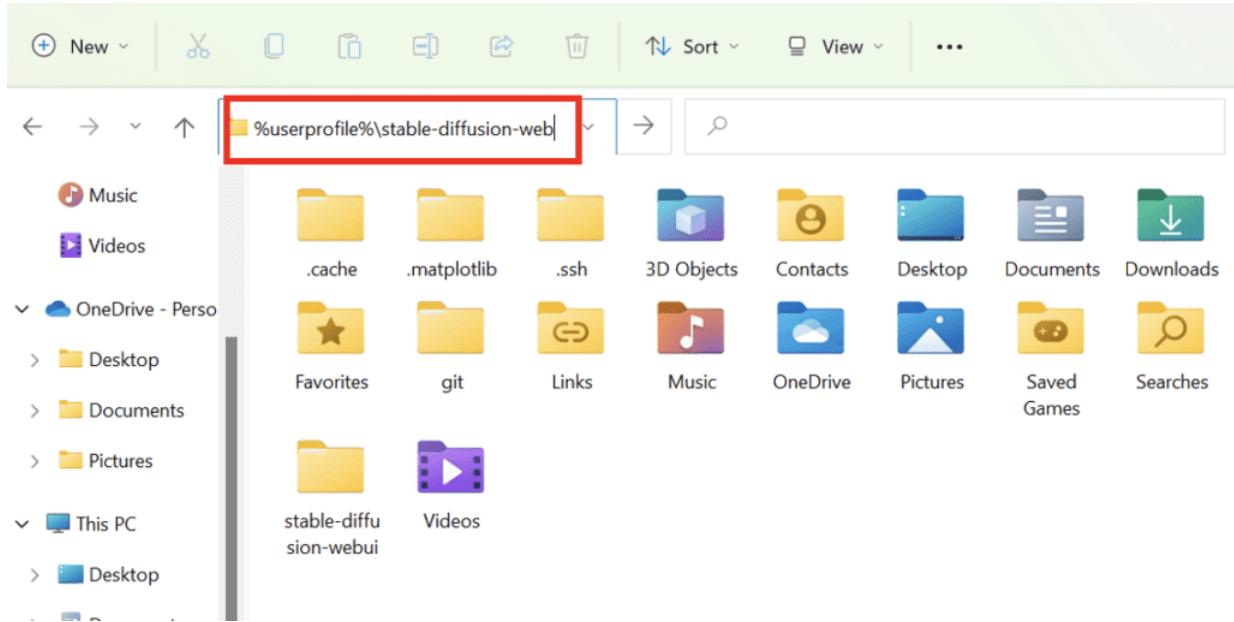
Sau khi tải xong, một thư mục có tên **stable-diffusion-webui** được tạo trên thư mục users trên máy của bạn.

Nếu bạn không muốn cài đặt trên ổ C, bạn có thể sao chép chương trình này vào

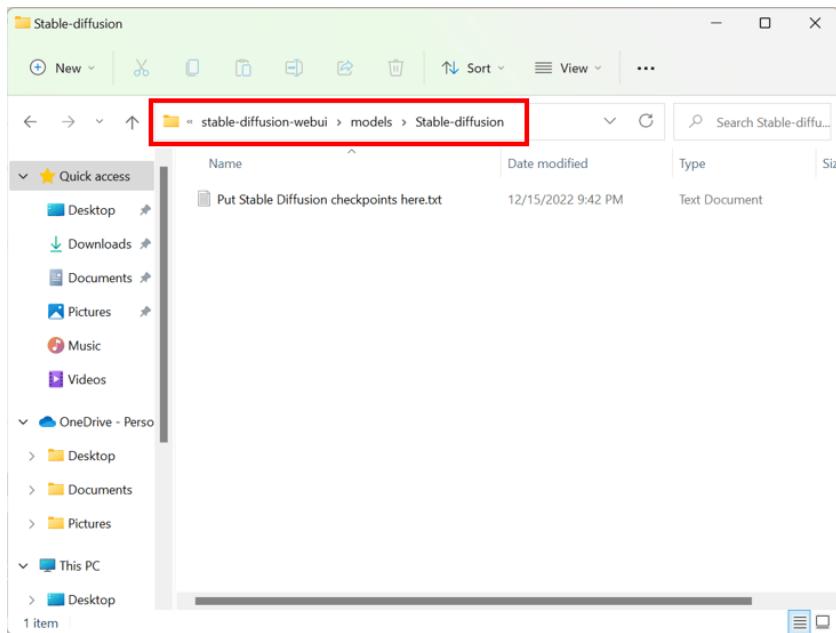
một thư mục khác trên ổ đĩa nào đó thay vì %UserProfile%.

#### Bước 4: Tải xuống file mô hình.

Tiếp theo, chuyển đến thư mục mới tạo trong File Explorer. Nhập %UserProfile%\stable-diffusion-webui vào thanh địa chỉ và nhấn enter.



Chuyển đến thư mục models và sau đó Stable-diffusion. Bạn sẽ thấy một tập tin “Stable Diffusion checkpoints here.txt” như dưới đây.



Tải xuống tệp Checkpoint mô hình Stable Diffusion v1.5 và đặt nó vào thư mục đó. Địa chỉ tải file là

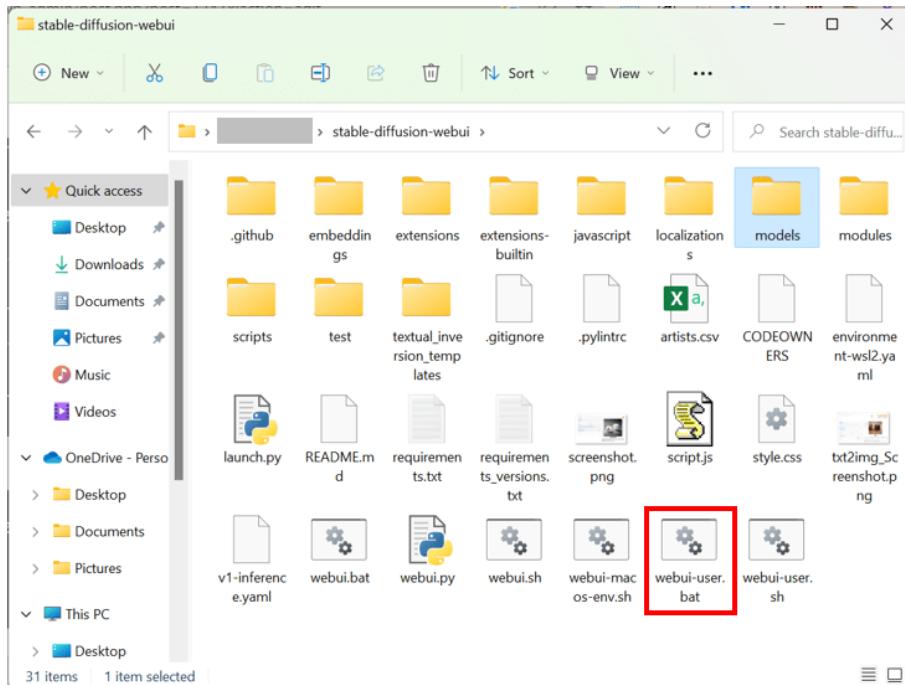
<https://huggingface.co/runwayml/stable-diffusion-v1-5/resolve/main/v1-5-pruned-emaonly.ckpt>

## Bước 5: Chạy AUTOMATIC1111

Bây giờ trong File Explorer, hãy quay lại stable-diffusion-webui thư mục. Tức là quay lại hai cấp độ hoặc gõ

%userprofile%\stable-diffusion-webui

một lần nữa trong thanh địa chỉ. Tìm file có tên **webui-user.bat**. Nhấn đúp chuột để chạy và hoàn tất cài đặt.



Bước cuối cùng này sẽ mất một thời gian (đối với lần đầu tiên chạy). Khi hoàn tất, bạn sẽ thấy thông báo

Running on local URL: <http://127.0.0.1:7860>

như dưới đây.

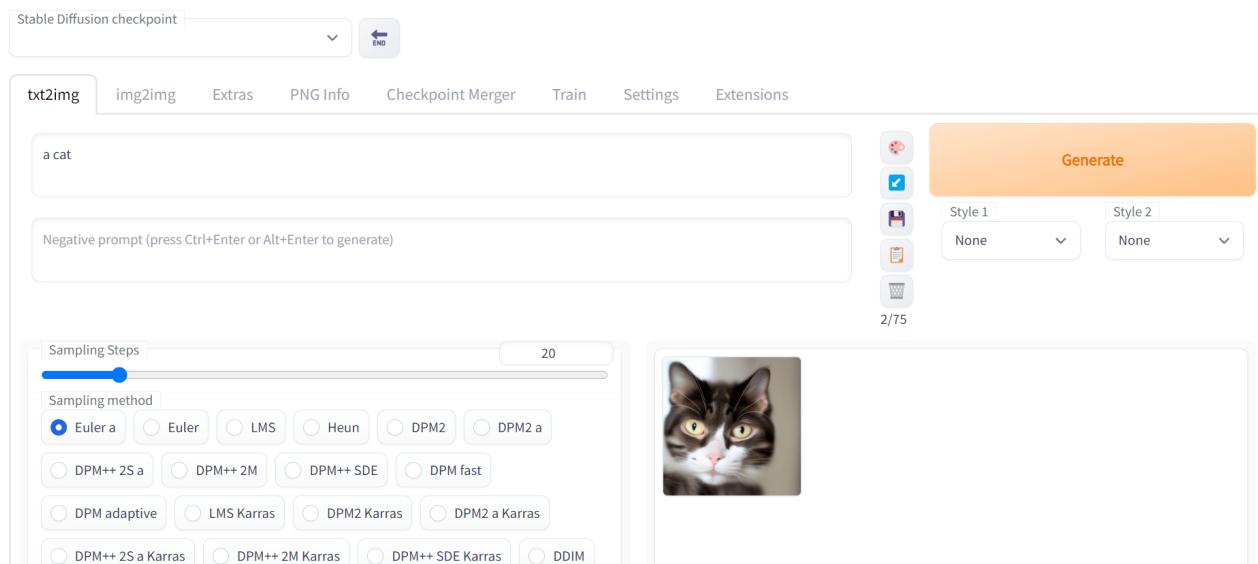
```
C:\WINDOWS\system32\cmd.exe
Installing torch and torchvision
Installing gfpgan
Installing clip
Installing open_clip
Cloning Stable Diffusion into repositories\stable-diffusion-stability-ai...
Cloning Taming Transformers into repositories\taming-transformers...
Cloning K-diffusion into repositories\k-diffusion...
Cloning CodeFormer into repositories\CodeFormer...
Cloning BLIP into repositories\BLIP...
Installing requirements for CodeFormer
Installing requirements for Web UI
Launching Web UI with arguments:
No module 'xformers'. Proceeding without it.
LatentDiffusion: Running in eps-prediction mode
DiffusionWrapper has 859.52 M params.
Downloading: 100% | 939k/939k [00:00<00:00, 3.20MB/s]
Downloading: 100% | 512k/512k [00:00<00:00, 2.06MB/s]
| 389/389 [00:00<00:00, 196kB/s]
| 905/905 [00:00<00:00, 461kB/s]
| 4.41k/4.41k [00:00<00:00, 6.44MB/s]
| 1.59G/1.59G [03:59<00:00, 7.15MB/s]
Loading weights [81761151] from C:\Users\...\stable-diffusion-webui\models\Stable-diffusion\v1-5-pruned-emaonly.ckpt
Applying cross attention optimization (Doggetttx).
Model loaded.
Loaded a total of 0 textual inversion embeddings.
Embeddings:
Running on local URL: http://127.0.0.1:7860
To create a public link, set `share=True` in `launch()`.
```

Trong trình duyệt web của bạn, hãy truy cập URL

<http://127.0.0.1:7860>

Bạn sẽ thấy giao diện AUTOMATIC1111!

Nhập Prompt (ví dụ: “a cat”) và nhấn Generate để kiểm tra xem Stable diffusion có chạy chính xác hay không.



Khi bạn sử dụng xong Stable Diffusion, hãy đóng cửa sổ cmd để tắt Stable Diffusion. Để chạy lại Stable Diffusion bạn chỉ cần click đúp vào **webui-user.bat**.

### Tùy chọn

Bạn sẽ có thể tăng tốc độ Stable diffusion bằng tham số --xformers. Nếu bạn có VRAM dưới 8 GB trên GPU, bạn nên bật--medvram tùy chọn để tiết kiệm bộ nhớ để bạn có thể tạo ra nhiều hình ảnh cùng một lúc.

Để kích hoạt chúng, nhập chuột phải vào tập tin **webui-user.bat** và chọn **Edit**. (Trước tiên, bạn có thể cần chọn “Show more Options” nếu bạn đang sử dụng Windows 11).

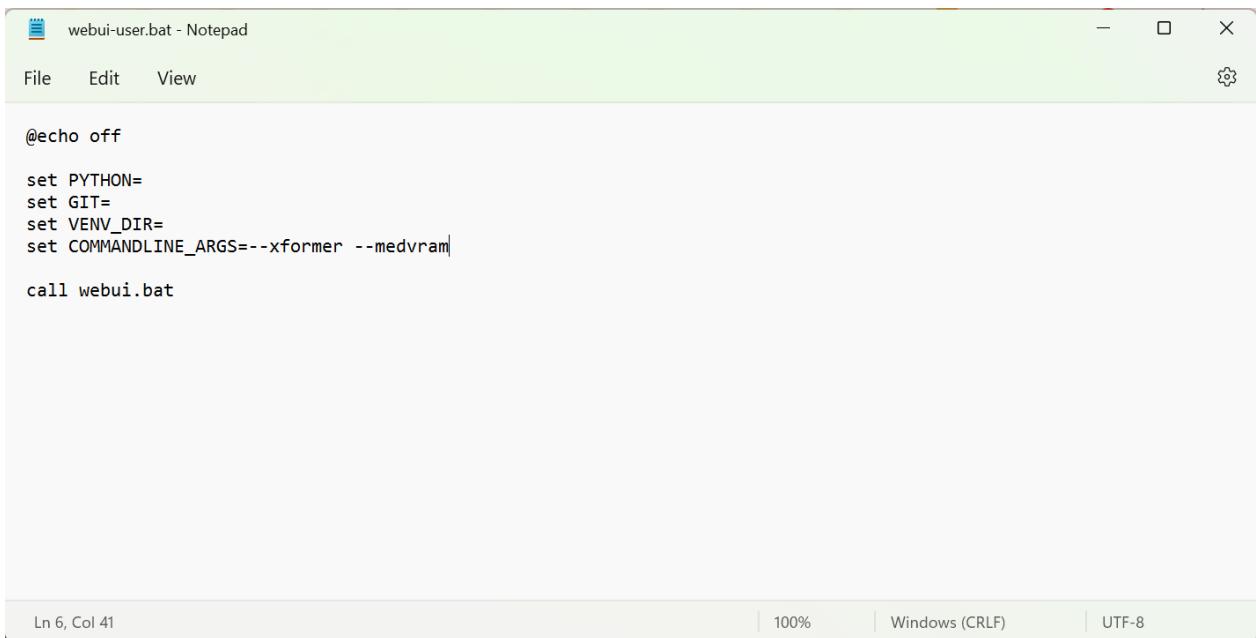
Thay thế dòng

```
set COMMANDLINE_ARGS=
```

Với

```
set COMMANDLINE_ARGS=--xformers --medvram
```

Lưu và đóng file **webui-user.bat**



```
webui-user.bat - Notepad
File Edit View
@echo off
set PYTHON=
set GIT=
set VENV_DIR=
set COMMANDLINE_ARGS=--xformer --medvram
call webui.bat
Ln 6, Col 41 | 100% | Windows (CRLF) | UTF-8
```

Nhấp đúp chuột vào **webui-user.bat** để chạy Stable diffusion.

### 3.5. Chạy trên máy Mac

Bạn sẽ cần một máy Mac có Apple Silicon M1 hoặc M2, lý tưởng nhất là có RAM ít nhất 16GB.

**Bước 1:** Cài đặt [Homebrew](#), trình quản lý gói dành cho Mac, nếu bạn chưa có. Mở **Terminal**, gõ lệnh sau và nhấn return.

```
/bin/bash -c "$(curl -fsSL
https://raw.githubusercontent.com/Homebrew/install/HEAD/install.sh)"
```

**Bước 2:** Cài đặt một số gói cần thiết. Mở một thiết bị đầu cuối mới và chạy lệnh sau  
brew install cmake protobuf rust python@3.10 git wget

**Bước 3:** Sao chép kho lưu trữ AUTOMATIC1111 bằng cách chạy lệnh sau trong terminal

```
git clone https://github.com/AUTOMATIC1111/stable-diffusion-webui
```

Một thư mục mới **stable-diffusion-webui** được tạo trong thư mục chính trên máy tính của bạn.

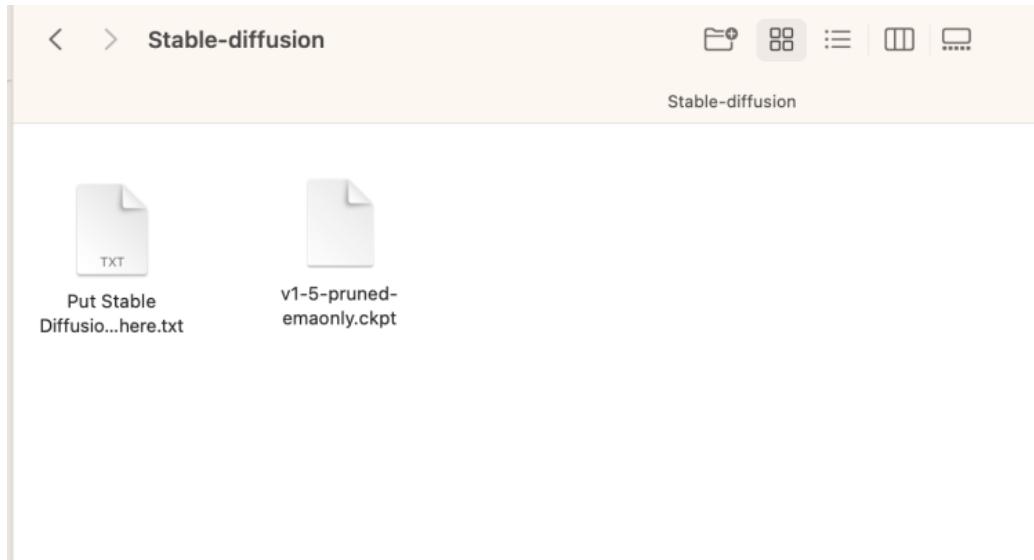
**Bước 5:** Tiếp theo, bạn sẽ cần một **model** để chạy Stable diffusion. Sử dụng địa chỉ sau để tải xuống mô hình v1.5.

<https://huggingface.co/runwayml/stable-diffusion-v1-5/resolve/main/v1-5-pruned-emaonly.ckpt>

Đặt tập tin vào thư mục **stable-diffusion-webui/models/Stable-diffusion**. Bạn có thể đến đó trong **Trình tìm kiếm** ứng dụng. Trong menu trên cùng, nhấp vào **Go** và sau đó **Home**. Nhấp đúp chuột vào các thư mục stable-diffusion-webui > models > Stable-

diffusion.

Khi bạn hoàn thành bước này, thư mục Stable-diffusion sẽ có hai tệp như bên dưới.



## Chạy AUTOMATIC1111

Thực hiện theo các bước trong phần này để khởi động GUI AUTOMATIC1111

Trong terminal, chạy lệnh sau

```
cd ~/stable-diffusion-webui;./webui.sh
```

Sẽ mất một lúc khi bạn chạy nó lần đầu tiên.

```
stable-diffusion-webui — Python launch.py — 80x24
~/stable-diffusion-webui — Python launch.py +[+]

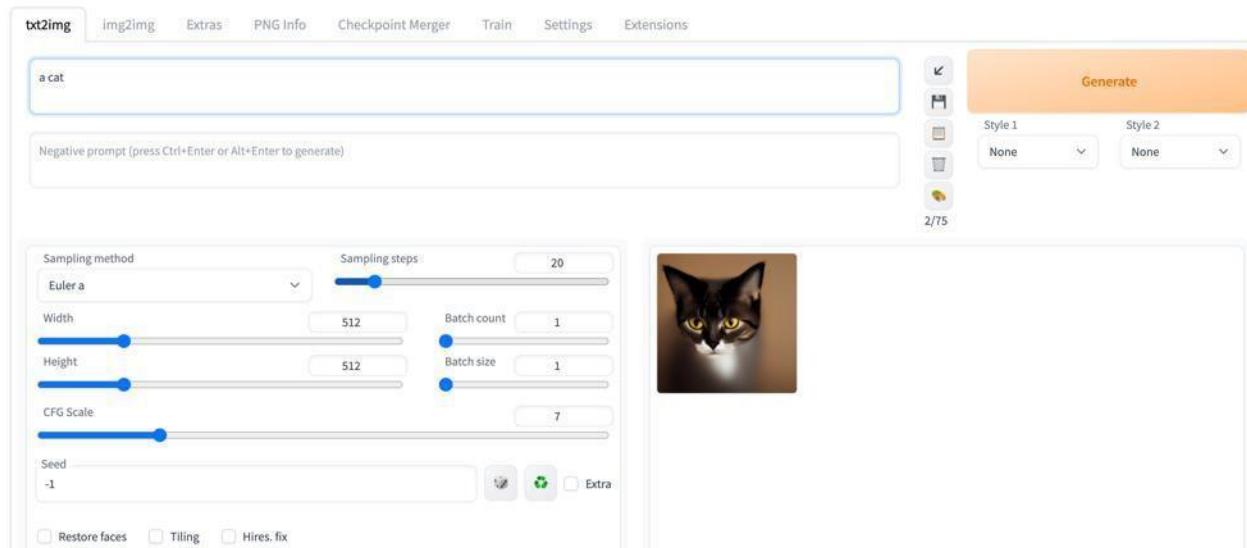
Installing clip
Installing open_clip
Cloning Stable Diffusion into repositories/stable-diffusion-stability-ai...
Cloning Taming Transformers into repositories/taming-transformers...
Cloning K-diffusion into repositories/k-diffusion...
Cloning CodeFormer into repositories/CodeFormer...
Cloning BLIP into repositories/BLIP...
Installing requirements for CodeFormer
Installing requirements for Web UI
Launching Web UI with arguments: --no-half --use-cpu interrogate
No module 'xformers'. Proceeding without it.
Warning: caught exception 'Torch not compiled with CUDA enabled', memory monitor
disabled
LatentDiffusion: Running in eps-prediction mode
DiffusionWrapper has 859.52 M params.
Loading weights [81761151] from /Users/[REDACTED]/stable-diffusion-webui/models/Stabl
e-diffusion/v1-5-pruned-emaonly.ckpt
Applying cross attention optimization (InvokeAI).
Textual inversion embeddings loaded():
Model loaded in 7.6s (0.6s create model, 5.4s load weights).
Running on local URL: http://127.0.0.1:7860

To create a public link, set `share=True` in `launch()`.
```

Mở trình duyệt web và truy cập URL sau để bắt đầu Stable diffusion.

<http://127.0.0.1:7860>

Bạn sẽ thấy giao diện AUTOMATIC1111. Nhập Prompt “a cat” và nhấn **Generate** để tạo ảnh.



Đóng **terminal** khi bạn hoàn tất.

Hãy làm theo các bước trên vào lần tiếp theo khi bạn muốn chạy Stable diffusion.

## PHẦN II - HƯỚNG DẪN CHO NGƯỜI MỚI BẮT ĐẦU

### Chương 4. Cách xây dựng Prompt hiệu quả

Trong Chương này, chúng ta sẽ thực hiện quy trình từng bước để xây dựng một Prompt tốt.

#### 4.1. Danh mục keyword

Có những kỹ thuật đã được chứng minh để tạo ra hình ảnh cụ thể, chất lượng cao. Prompt của bạn sẽ bao gồm hầu hết các lĩnh vực này

1. Subject (required)
2. Medium
3. Style
4. Artist
5. Website
6. Resolution
7. Lighting
8. Additional details
9. Color

#### 4.2. Ví dụ về xây dựng Prompt

Bây giờ chúng ta hãy xem một ví dụ về cách tạo hình ảnh bằng cách sử dụng keyword từ mỗi danh mục trên. Chúng tôi sẽ sử dụng Stable Diffusion v1.4 để tạo ví dụ này. Sử dụng v1.4 hoặc 1.5.

Đầu tiên bạn sẽ cần một bản mô tả về **subject** với càng nhiều chi tiết càng tốt. Ví dụ:

*A young woman with light blue dress sitting next to a wooden window reading a book.*

*Một người phụ nữ trẻ với chiếc váy màu xanh nhạt đang ngồi cạnh cửa sổ gỗ đọc sách.*

Chúng ta có hình ảnh sau đây rất phù hợp với Prompt.



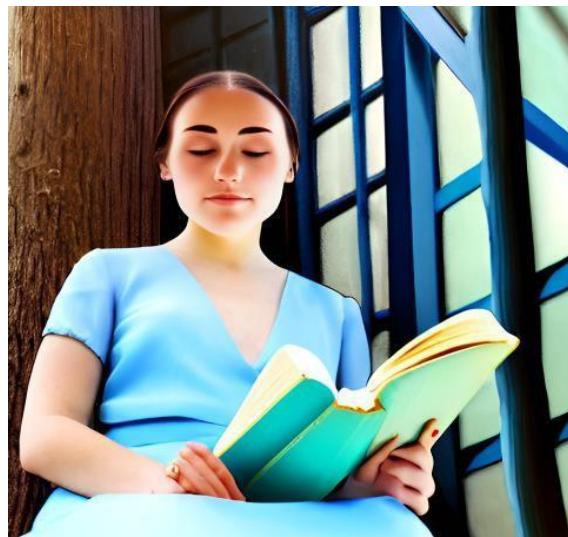
Chúng tôi có thể cụ thể hơn. Hãy thêm một **medium**. Một số ví dụ *digital painting, photograph, oil painting (tranh kỹ thuật số, ảnh, tranh sơn dầu)*. Hãy sử dụng *digital painting (bức tranh kỹ thuật số)*.

Prompt mới là

*Digital painting of a young woman with light blue dress sitting next to a wooden window reading a book*

*(Bức tranh kỹ thuật số của một phụ nữ trẻ mặc váy xanh nhạt ngồi cạnh cửa sổ gỗ đọc sách)*

Hình ảnh kết quả là



Bạn có thể thấy hình ảnh thay đổi từ một bức ảnh sang một tác phẩm nghệ thuật kỹ thuật số. Bạn hiểu ý rồi. Hãy thêm phần còn lại của chúng

**Artist:** by Stanley Artgerm Lau

**Website:** artstation

**Resolution:** 8k

**Lighting:** cinematic lighting

**Additional details:** Extremely detailed, ornate

**Color:** vivid

Đặt tất cả chúng lại với nhau, Prompt là:

*Digital painting of a young woman with light blue dress sitting next to a wooden window reading a book, by Stanley Artgerm Lau, artstation, 8k, extremely detailed, ornate, cinematic lighting, vivid.*

*Tranh kỹ thuật số về một người phụ nữ trẻ với bộ váy màu xanh nhạt ngồi cạnh cửa sổ gỗ đọc sách, của Stanley Artgerm Lau, artstation, 8k, cực kỳ chi tiết, trang trí công phu, ánh sáng điện ảnh, sống động.*

tạo ra hình ảnh này:



Bằng cách thêm keyword vào Prompt, chúng ta có thể thiết kế hình ảnh để có được phong cách như mong muốn.

### 4.3. Lời khuyên để xây dựng Prompt tốt

Dưới đây là một số mẹo để xây dựng Prompt tốt:

- Hãy chi tiết và cụ thể khi mô tả **subject**.
- Sử dụng nhiều dấu ngoặc () để tăng độ mạnh và [] để giảm
- Tên nghệ sĩ là một công cụ sửa đổi phong cách rất mạnh mẽ. Hãy sử dụng một cách khôn ngoan.

### 4.4. Danh sách keyword

Dưới đây là một số keyword phổ biến và tác dụng của chúng đã được thử nghiệm với Stable diffusion v1.4 và v1.5.

#### 4.4.1. Medium

Phương tiện xác định một thể loại của tác phẩm nghệ thuật.

Từ khóa	Ghi chú
<b>Portrait</b>	Lấy nét hình ảnh vào khuôn mặt/chân dung
<b>Digital painting</b>	Phong cách nghệ thuật kỹ thuật số
<b>Concept art</b>	Phong cách minh họa, 2D

<b>Ultra realistic illustration</b>	Vẽ hình ảnh có tính thực tế
<b>Underwater portrait</b>	Dưới nước, tóc bồng bềnh
<b>Underwater steampunk</b>	Dưới nước với màu rữa

#### 4.4.2. Style

Những keyword này tiếp tục tinh chỉnh phong cách nghệ thuật.

Từ khóa	Ghi chú
<b>hyperrealistic</b>	Tăng chi tiết và độ phân giải
<b>pop-art</b>	Phong cách nghệ thuật đại chúng
<b>Modernist</b>	Màu sắc rực rỡ, độ tương phản cao
<b>art nouveau</b>	Thêm đồ trang trí và chi tiết, phong cách xây dựng

#### 4.4.3. Artist

Việc đề cập đến nghệ sĩ trong Prompt có tác dụng mạnh mẽ. Nghiên cứu công việc của họ và lựa chọn một cách khôn ngoan.

Từ khóa	Ghi chú
John Collier	Phong cách vẽ của họa sĩ vẽ chân dung thế kỷ 19
Stanley Artgerm Lau	Bản vẽ hiện đại thực tế.
Frida Kahlo	Hiệu ứng khá mạnh theo phong cách chân dung của Kahlo
John Singer Sargent	Thích hợp để sử dụng với chân dung người phụ nữ, tạo ra những bộ quần áo tinh tế thứ 19, một số chủ nghĩa ấn tượng
Alphonse Mucha	Tranh chân dung 2D phong cách Alphonse Mucha

#### 4.4.4. Website

Việc đề cập đến một trang web nghệ thuật hoặc ảnh có tác dụng mạnh mẽ, có thể vì mỗi trang web đều có thể loại riêng.

Từ khóa	Ghi chú

Pixiv	Phong cách hoạt hình Nhật Bản
pixabay	Phong cách ảnh thương mại
artstation	Minh họa hiện đại, tưởng tượng

#### 4.4.5. Resolution

Từ khóa	Ghi chú
unreal engine	Hình ảnh 3D rất chân thực và chi tiết
sharp focus	Tăng độ phân giải
8k	Tăng độ phân giải lên 8k. Làm cho hình ảnh giống máy ảnh chụp hơn và chân thực hơn
vray	Kết xuất 3D, tốt nhất cho các vật thể, cảnh quan và tòa nhà.

#### 4.4.6. Additional details (Chi tiết bổ sung)

Thêm chi tiết cụ thể vào hình ảnh của bạn.

Từ khóa	Ghi chú
dramatic	Tăng biểu cảm cảm xúc của khuôn mặt.
silk	Thêm lụa vào quần áo
expansive	Nền thoáng hơn, chủ đề nhỏ hơn
low angle shot	chụp từ góc thấp
psychedelic	màu sắc sống động với sự biến dạng

#### 4.4.7. Lighting

Thay đổi độ sáng.

Từ khóa	Ghi chú
god rays	ánh sáng mặt trời xuyên qua đám mây
cinematic lightning	Tăng độ tương phản giữa các vùng
studio lightning	Chiếu sáng điểm trên khuôn mặt

#### 4.4.8. Color

Thêm bảng màu bổ sung cho hình ảnh.

Từ khóa	Ghi chú
iridescent gold	Vàng sáng bóng
silver	Màu bạc
vintage	Hiệu ứng cổ điển

### Chương 5. Thông số hình ảnh

Như vậy là bạn đã tạo ra một vài hình ảnh với Stable Diffusion. Chúng trông tuyệt vời nhưng không hoàn toàn như bạn muốn? Bạn có thể sử dụng một số tùy chỉnh. Đây là phần sơ lược về các tham số tạo cơ bản.

#### 5.1. CFG Scale

**Classifier Free Guidance scale** là một tham số để kiểm soát mức độ tuân thủ Prompt của mô hình.

1 – Hầu hết bỏ qua Prompt của bạn.

3 – Hãy sáng tạo hơn.

7 – Cân bằng tốt giữa việc làm theo Prompt và sự tự do.

15 – Tuân thủ thêm theo Prompt.

30 – Tuân thủ nghiêm ngặt Prompt.

Dưới đây là một vài ví dụ về việc tăng CFG Scale với cùng một Seed ngẫu nhiên. Nói chung, bạn nên tránh xa 1 và 30. Đặt nó quá cao có thể khiến hình ảnh bị biến dạng quá mức.

**Khuyến nghị:** Bắt đầu với 7. Tăng lên nếu bạn muốn nó làm theo Prompt của bạn nhiều hơn.



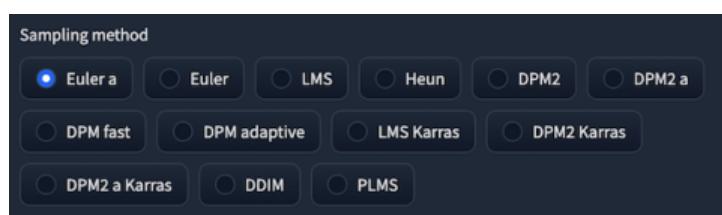
## 5.2. Các bước lấy mẫu (Sampling steps)

Chất lượng được cải thiện khi bước lấy mẫu tăng lên. Thông thường, 20 bước với Euler bộ lấy mẫu đủ để đạt được hình ảnh sắc nét, chất lượng cao. Mặc dù hình ảnh vẫn sẽ thay đổi một cách tinh tế khi bước lên các giá trị cao hơn nhưng nó sẽ trở nên khác biệt nhưng chưa chắc chất lượng sẽ cao hơn.

**Khuyến nghị:** 20 bước. Điều chỉnh lên cao hơn nếu bạn thấy chất lượng thấp.



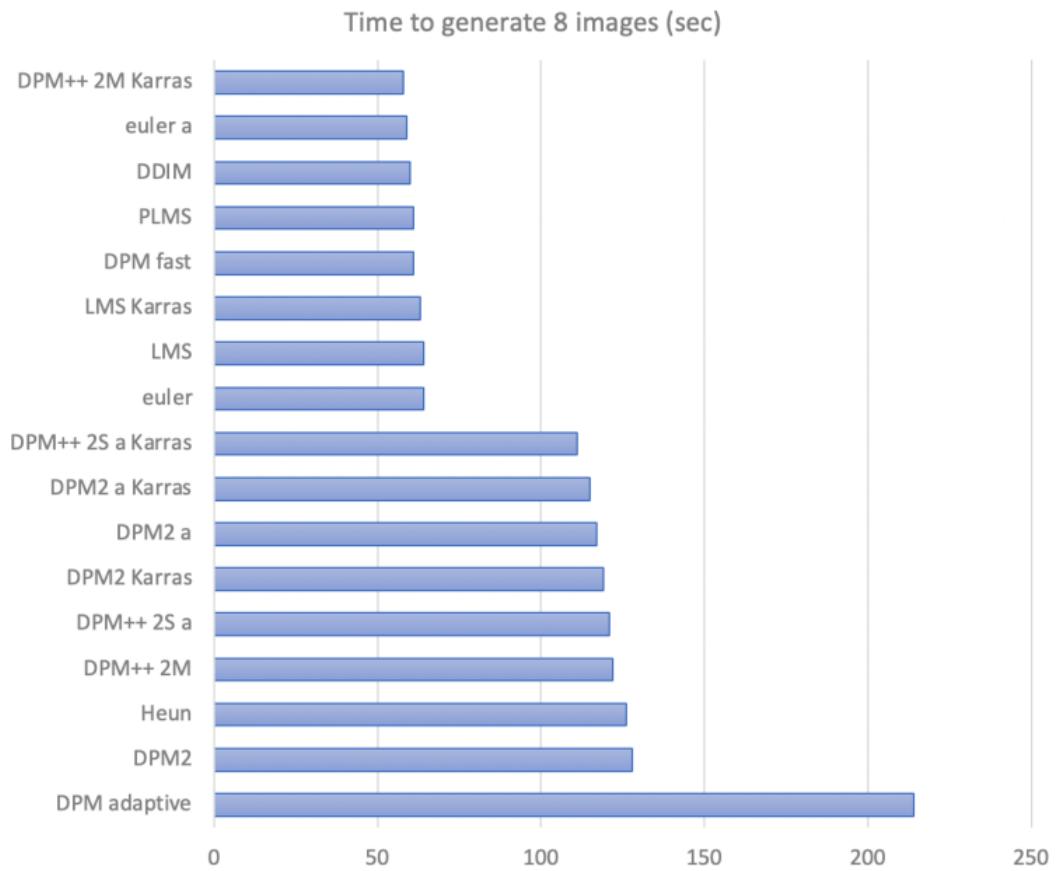
### 5.3. Phương pháp lấy mẫu (Sampling methods)



Có nhiều phương pháp lấy mẫu khác nhau mà bạn có thể chọn, tùy thuộc vào GUI bạn đang sử dụng. Chúng chỉ đơn giản là những phương pháp khác nhau để giải các phương trình khuếch tán. Chúng được cho là sẽ đưa ra kết quả tương tự nhưng có thể hơi khác nhau do sai số về mặt số học. Nhưng vì không có câu trả lời đúng ở đây – tiêu chí duy nhất là hình ảnh trông đẹp nên bạn không cần phải lo lắng về độ chính xác của phương pháp.

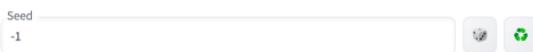
Không phải tất cả các phương pháp đều được tạo ra như nhau. Dưới đây là thời gian xử

lý của các phương pháp khác nhau.



Để đảm bảo hình ảnh tốt nhất, hãy sử dụng 20 steps với Euler.

#### 5.4. Seed (Hạt giống)



Seed ngẫu nhiên xác định mẫu nhiều ban đầu và từ đó xác định hình ảnh cuối cùng. Đặt nó thành -1 có nghĩa là sử dụng ngẫu nhiên. Nó rất hữu ích khi bạn muốn tạo hình ảnh mới.

Làm cách nào để tìm Seed được sử dụng cho hình ảnh nếu bạn sử dụng Seed ngẫu nhiên? Trong hộp thoại, bạn sẽ thấy một cái gì đó như:

Steps: 20, Sampler: Euler a, CFG scale: 7, **Seed: 4239744034**, Size: 512×512, Model hash: 7460a6fa

Chỉ cần sao chép giá trị Seed này vào hộp nhập Seed. Nếu bạn tạo nhiều hình ảnh cùng một lúc, giá trị gốc của hình ảnh thứ hai là số này tăng thêm 1, v.v. Ngoài ra, hãy nhấp vào nút recycle để sử dụng lại Seed từ thẻ hệ trước.

**Khuyến nghị:** Đặt thành -1 nếu muốn các kết quả khác nhau sau mỗi lần tạo ảnh.

#### 5.5. Image Size (Kích cỡ hình)

Kích thước của hình ảnh đầu ra. Vì Stable diffusion được đào tạo với hình ảnh  $512 \times 512$  nên việc đặt nó ở kích thước dọc hoặc ngang có thể tạo ra các sự cố không mong muốn. Để nó ở dạng hình vuông bất cứ khi nào có thể.

**Khuyến nghị:** Đặt kích thước hình ảnh là  $512 \times 512$ .

## 5.6. Batch Size (Kích thước lô)

Batch size là số lượng hình ảnh được tạo ra mỗi lần. Vì hình ảnh cuối cùng phụ thuộc rất nhiều vào Seed ngẫu nhiên nên việc tạo một vài hình ảnh cùng một lúc luôn là một ý tưởng hay.

**Khuyến nghị:** Thiết lập giá trị Đặt Batch Size là 4 hoặc 8.

## 5.7. Khôi phục khuôn mặt

Restore faces     Tiling     Hires. fix

Một bí mật nhỏ của Stable Diffusion là nó thường gặp vấn đề với khuôn mặt và mắt. Khôi phục khuôn mặt là phương pháp xử lý hậu kỳ được áp dụng cho hình ảnh sử dụng AI được đào tạo riêng để chỉnh sửa khuôn mặt.

Để bật tính năng này, hãy chọn hộp bên cạnh **Restore faces**. Di tới tab Cài đặt, bên dưới **Face restoration model**, lựa chọn **CodeFormer**. Dưới đây là hai ví dụ. Không có phục hồi khuôn mặt là ở bên trái. Với phục hồi khuôn mặt là ở bên phải.

**Khuyến nghị:** Bật khôi phục khuôn mặt khi bạn tạo hình ảnh có khuôn mặt.



## Chương 6. Mô hình

**Models**, đôi khi được gọi là **Checkpoint**, là các trọng số Stable diffusion được huấn luyện trước nhằm mục đích tạo ra một thể loại hình ảnh chung hoặc cụ thể. Nó được cho là tính năng quan trọng nhất của Stable diffusion. Là một mô hình nguồn mở, cộng đồng đã tạo ra hàng nghìn mô hình miễn phí có thể dễ dàng tạo ra các phong cách riêng.

Những hình ảnh mà một mô hình có thể tạo ra phụ thuộc vào dữ liệu được sử dụng để đào tạo chúng. Một mô hình sẽ không thể tạo ra hình ảnh của một con mèo nếu dữ liệu huấn luyện không bao giờ có con mèo. Tương tự như vậy, nếu bạn chỉ huấn luyện một mô hình với hình ảnh mèo, nó sẽ chỉ tạo ra mèo.

Chúng tôi sẽ đi giới thiệu các model là gì, một số model thông dụng (v1, v2, F222, Anything V3, Open Journey v4), cách cài đặt, sử dụng và gộp chúng.

### 6.1. Models là gì?

#### 6.1.1. Tinh chỉnh

Tinh chỉnh một kỹ thuật phổ biến trong học máy là lấy mô hình được đào tạo trên tập dữ liệu rộng và đào tạo thêm một chút về tập dữ liệu hẹp mà bạn quan tâm.

Một mô hình được tinh chỉnh sẽ thiên về việc tạo ra các hình ảnh giống như tập dữ liệu của bạn, trong khi vẫn duy trì tính linh hoạt của mô hình ban đầu.

### **6.1.2. Tại sao mọi người làm chúng?**

Stable diffusion là tuyệt vời nhưng không tốt ở mọi thứ. Ví dụ: nó có thể và sẽ tạo ra các hình ảnh theo phong cách anime với keyword “anime” trong Prompt. Nhưng có thể khó tạo ra hình ảnh của một thể loại anime phụ. Thay vì mày mò với Prompt, bạn có thể tinh chỉnh mô hình bằng các hình ảnh của thể loại phụ đó. Mô hình mới sẽ luôn tạo ra những hình ảnh có phong cách đó.

### **6.1.3. Chúng được tạo ra như thế nào?**

Có hai phương pháp tinh chỉnh chính: (1) **Luyện tập bổ sung** và 2) **Dreambooth**. Cả hai đều bắt đầu với mô hình cơ bản như Stable Diffusion v1.4 hoặc v1.5.

Bạn có thể đào tạo bổ sung bằng cách đào tạo một mô hình cơ sở với tập dữ liệu bổ sung mà bạn quan tâm. Ví dụ: bạn có thể đào tạo Stable Diffusion v1.5 với tập dữ liệu bổ sung về ô tô cổ để thiên về tính thẩm mỹ của ô tô đối với thể loại ô tô cổ.

**Dreambooth**, do Google phát triển ban đầu, là một kỹ thuật đưa các chủ đề tùy chỉnh vào các mô hình chuyển văn bản thành hình ảnh. Nó hoạt động với ít nhất 3-5 hình ảnh tùy chỉnh. Bạn có thể chụp một vài bức ảnh của chính mình và sử dụng Dreambooth để làm Models. Một mô hình được huấn luyện bằng Dreambooth cần có một keyword đặc biệt trong Prompt để sử dụng nó.

## **6.2. Models đáng chú ý**

Có hai nhóm mô hình: v1 và v2. Tôi sẽ đề cập đến các mô hình v1 trong phần này và các mô hình v2 trong phần tiếp theo. Có hàng nghìn mô hình Stable diffusion được tinh chỉnh và số lượng này đang tăng lên mỗi ngày. Dưới đây là danh sách các mô hình có thể được sử dụng.

### 6.2.1. Stable diffusion 1.4



Trang mẫu: <https://huggingface.co/CompVis/stable-diffusion-v-1-4-origin>

Liên kết tải xuống: <https://huggingface.co/CompVis/stable-diffusion-v-1-4-original/blob/main/sd-v1-4.ckpt>

Được phát hành vào tháng 8 năm 2022 bởi Stability AI, mô hình v1.4 là mô hình Stable diffusion đầu tiên được cung cấp công khai.

### 6.2.2. Stable diffusion v1.5



Trang mẫu: <https://huggingface.co/runwayml/stable-diffusion-v1-5>

Liên kết tải xuống: <https://huggingface.co/runwayml/stable-diffusion-v1-5/blob/main/v1-5-pruned-emaonly.ckpt>

v1.5 được phát hành vào tháng 10 năm 2022 bởi Runway ML, đối tác của Stability AI. Mô hình này dựa trên v1.2 và được đào tạo thêm.

### 6.2.3. F222



Liên kết tải xuống: <https://civitai.com/models/1188/f222>

F222 ban đầu được đào tạo để tạo ảnh khỏa thân nhưng mọi người thấy nó hữu ích khi tạo ra những bức chân dung phụ nữ xinh đẹp với các mối quan hệ chính xác của bộ phận cơ thể. Thật thú vị, trái ngược với những gì bạn nghĩ, nó khá tốt trong việc tạo ra những bộ quần áo có tính thẩm mỹ. Tuy nhiên hãy thận trọng khi sử dụng vì đây là mô hình NSFW (Hình ảnh tạo ra có yếu tố khỏa thân).

F222 là mô hình có mục đích đặc biệt nhưng khá hữu ích khi chụp ảnh chân dung. Nó có xu hướng tạo ra ảnh khỏa thân cao nhưng có thể bị ngăn chặn bởi các keyword như “dress”. Đọc danh sách keyword được đề xuất

#### **6.2.4. Anything V3**



Trang mẫu: <https://huggingface.co/Linaqruf/anything-v3.0>

Liên kết tải xuống: <https://huggingface.co/acheong08/Anything/blob/main/Anything-V3.0-pruned-fp32.ckpt>

Anything V3 là mô hình được đào tạo để tạo ra hình ảnh mang phong cách anime chất lượng cao. Bạn có thể sử dụng thẻ danbooru (như “1girl” và “white hair”) trong Prompt văn bản. Nó rất hữu ích khi chọn những người nổi tiếng theo phong cách anime. Một nhược điểm là nó tạo ra những hình ảnh có hình dáng cơ thể không cân đối.

#### 6.2.5. Open Journey



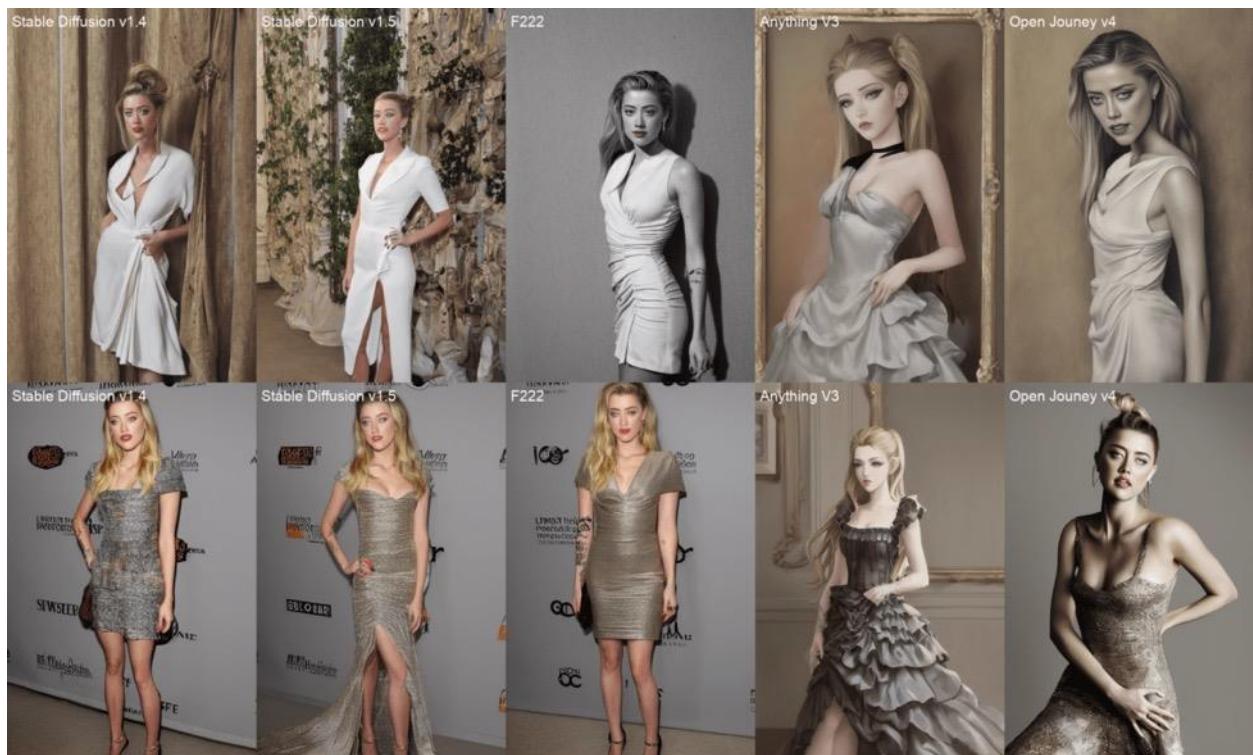
Trang mẫu: <https://huggingface.co/prompthero/openjourney>

Liên kết tải xuống: <https://huggingface.co/prompthero/openjourney/resolve/main/mdjrny-v4.ckpt>

Open Journey là mô hình được tinh chỉnh với hình ảnh được tạo bởi MidJourney v4. Nó có tính thẩm mỹ khác biệt và là một mô hình tốt.

#### 6.2.6. So sánh

Dưới đây là so sánh các mô hình này có cùng Prompt và Seed. Các model khác nhau tạo ra hình ảnh chân thực nhưng có tính thẩm mỹ khác nhau.



Từ trái sang phải: v1.4, v1.5, F222, Anything V3 và OpenJourney.

### 6.3. Các models khác

Có hàng trăm mô hình Stable diffusion có sẵn. Nhiều trong số đó là những mẫu có mục đích đặc biệt được thiết kế để tạo ra một phong cách cụ thể. Một số đáng chú ý là:

#### 6.3.1. Inkpunk Diffusion



Liên kết tải xuống: <https://huggingface.co/Envvi/Inkpunk-Diffusion>

Inkpunk Diffusion là mô hình do Dreambooth đào tạo với phong cách minh họa rất khác biệt. Sử dụng keyword: *nvinkpunk*

#### 6.3.2. Waifu-diffusion



Liên kết tải xuống: <https://huggingface.co/hakurei/waifu-diffusion> Waifu Diffusion mang phong cách anime Nhật Bản.

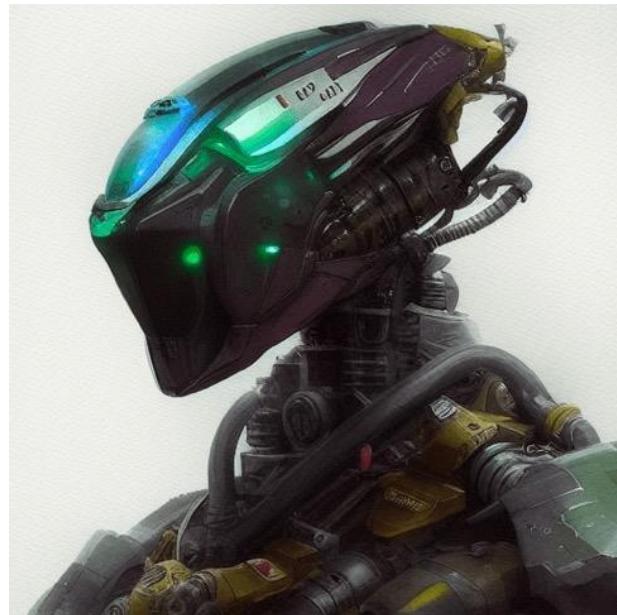
### 6.3.3. Arcane Diffusion



Liên kết tải xuống: <https://huggingface.co/nitrosocke/Arcane-Diffusion>

Arcane Diffusion là chương trình truyền hình theo phong cách Arcane. Sử dụng keyword theo phong cách phức tạp để kích hoạt hiệu ứng.

### 6.3.4. Robo Diffusion



Liên kết tải xuống: <https://huggingface.co/nousr/robo-diffusion>

Robot Diffusion là một mô hình phong cách robot thú vị sẽ biến mọi thứ đối tượng của bạn thành robot!

### 6.3.5. Mo-di-diffusion



Liên kết tải xuống: <https://huggingface.co/nitrosocke/mo-di-diffusion>

Đây là mô hình dành cho bạn nếu bạn muốn tạo ra phong cách giống Pixar. Sử dụng keyword: *phong cách Disney hiện đại*

### 6.3.6. Tìm thêm models

Bạn có thể tìm thấy nhiều mô hình hơn trong các trang web sau:

Huggingface: <https://huggingface.co/models?other=stable-diffusion> Civitai:  
<https://civitai.com/>

## 6.4. V2 Models



Stability AI đã phát hành một loạt mô hình mới có tên là phiên bản 2. Cho đến nay các mô hình 2.0 và 2.1 đã được phát hành. Những thay đổi chính trong mô hình v2 là:

- Ngoài  $512 \times 512$  pixel, còn có phiên bản có độ phân giải cao hơn  $768 \times 768$  pixel.
- Bạn không còn có thể tạo nội dung khiêu dâm vì tài liệu khiêu dâm đã bị xóa khỏi dữ liệu đào tạo.

Bạn có thể cho rằng mọi người đã chuyển sang sử dụng mô hình v2. Tuy nhiên, cộng đồng Stable Diffusion nhận thấy hình ảnh trông tệ hơn ở mô hình 2.0. Mọi người cũng gặp khó khăn trong việc sử dụng các keyword quyền lực như tên người nổi tiếng và tên nghệ sĩ để định hướng phong cách.

Mô hình 2.1 đã giải quyết được phần nào những vấn đề này. Hình ảnh trông đẹp hơn khi tạo hình. Việc tạo ra các phong cách nghệ thuật dễ dàng hơn. Tính đến thời điểm hiện tại, hầu hết mọi người đã hoàn toàn chuyển sang mô hình 2.1.

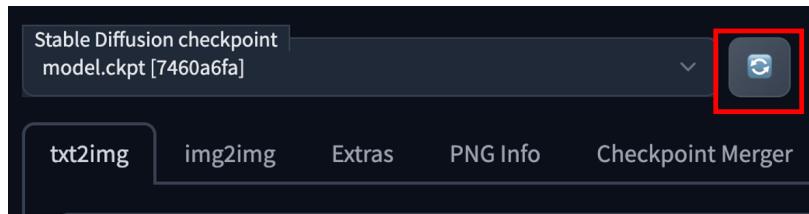
Nhiều người thích thoảng sử dụng chúng để tạo ra các hình ảnh kiểu ảnh nhưng lại dành phần lớn thời gian cho các mẫu v1.

## 6.5. Cách cài đặt mô hình

Những hướng dẫn này có thể áp dụng cho các mẫu v1. Xem Chương 16 để cài đặt các mô hình v2.

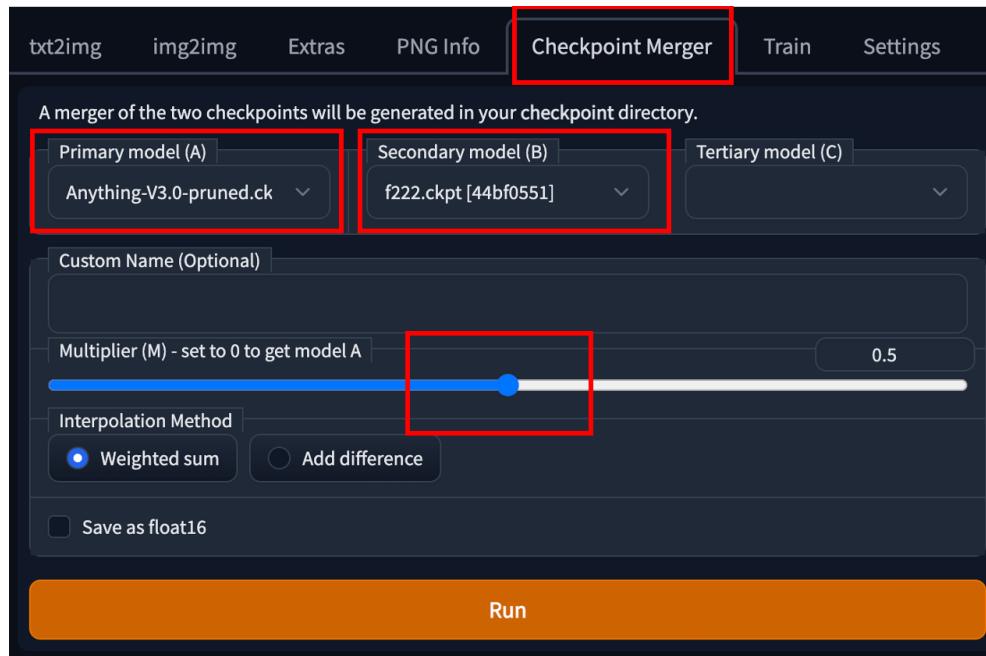
Mô hình rất dễ cài đặt. Để cài đặt một mô hình trong GUI AUTOMATIC1111, hãy tải xuống và đặt tệp checkpoint (.ckpt) vào thư mục: stable-diffusion-webui/mô hình/Stable-diffusion/

Nhấn nút tải lại bên cạnh hộp chọn checkpoint



Bạn sẽ thấy tệp Checkpoint bạn vừa đưa vào có sẵn để lựa chọn. Chọn tệp Checkpoint mới để sử dụng mô hình.

## 6.6. Hợp nhất hai Models



Để hợp nhất hai mô hình bằng GUI AUTOMATIC1111, hãy đi tới **Checkpoint Merger tab** và chọn hai mô hình bạn muốn hợp nhất Mô **Primary model (A)** Và Mô **Secondary model (B)**.

Điều chỉnh hệ số nhân (M) để điều chỉnh trọng lượng tương đối của hai mô hình. Đặt nó thành 0,5 sẽ hợp nhất hai mô hình có tầm quan trọng như nhau. Sau khi nhấn **Run**, mô hình hợp nhất mới sẽ có sẵn để sử dụng.

Dưới đây là hình ảnh mẫu từ việc hợp nhất F222 và Anything V3 với trọng số bằng nhau (0,5). Mô hình được hợp nhất nằm giữa phong cách F222 thực tế và phong cách Anything V3 của anime.



Từ trái sang phải: F222, Anything V3 và Merged

## Chương 7. Inpainting

Cho dù Prompt và Models của bạn có tốt đến đâu thì cũng hiếm khi có được một bức ảnh hoàn hảo chỉ trong một lần tạo. Inpainting là một cách không thể thiếu để khắc phục những khuyết điểm nhỏ. Trong chương này, sẽ trình bày một số ví dụ cơ bản để sử dụng nó trong việc sửa lỗi.

### 7.1. Cài đặt inpainting cơ bản

Trong phần này, tôi sẽ hướng dẫn bạn từng bước cách sử dụng inpainting để khắc phục các lỗi nhỏ. Đây là hình ảnh ban đầu.



Prompt của hình ảnh này là

*[emma watson: amber heard: 0.5], (long hair:0.5), headLeaf, wearing stola, vast roman palace, large window, medieval renaissance palace, ((large room)), 4k, arstation, intricate, elegant, highly detailed*

*[emma watson: hổ phách đã nghe: 0.5], (tóc dài:0.5), headLeaf, đeo stola, cung điện La Mã rộng lớn, cửa sổ lớn, cung điện thời phục hưng thời Trung cổ, ((phòng lớn)), 4k, Arstation, phức tạp, trang nhã, có độ chi tiết cao*

Đó là một hình ảnh đẹp nhưng tôi muốn khắc phục các vấn đề sau:

- Khuôn mặt trông không được tự nhiên.
- Cánh tay phải bị mờ.

### 7.2. Mô hình có hỗ trợ Inpainting

Bạn có biết có những mô hình Stable diffusion được đào tạo đặc biệt để inpainting không? Bạn có thể sử dụng nó nếu bạn muốn có được kết quả tốt nhất. Nhưng thông thường, bạn có thể sử dụng cùng một mô hình mà bạn đã tạo hình ảnh để vẽ.

Để cài đặt mô hình inpainting v1.5, hãy tải xuống tệp Checkpoint mô hình từ URL sau.

<https://huggingface.co/runwayml/stable-diffusion-inpainting/blob/main/sd-v1-5-inpainting.ckpt>

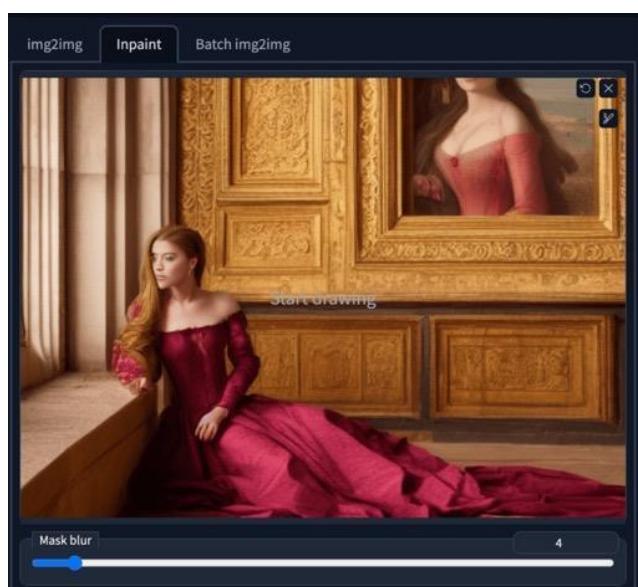
và đặt nó vào thư mục: stable-diffusion-webui/models/Stable-diffusion

Trong AUTOMATIC1111, nhấn biểu tượng làm mới bên cạnh hộp chọn lựa chọn Checkpoint ở trên cùng bên trái. Lựa chọn sd-v1-5-inpainting.ckpt để kích hoạt mô hình.

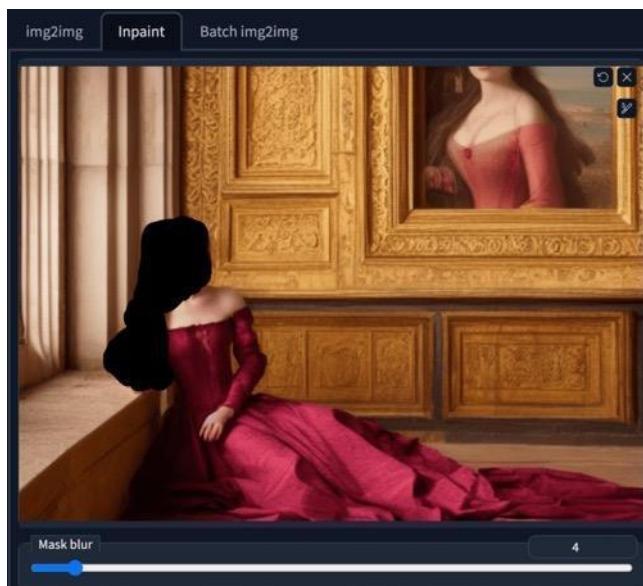


### 7.3. Tạo mặt nạ Inpaint

Trong GUI AUTOMATIC1111, chọn **img2img** tab và chọn **Inpaint** tab phụ. Tải hình ảnh lên khung vẽ.



Chúng ta sẽ sơn cả cánh tay phải và mặt cùng một lúc. Sử dụng công cụ cọ vẽ để tạo ra một **mặt nạ** để được Inpainting. Đây là khu vực bạn muốn Stable diffusion tái tạo.



## 7.4. Cài đặt cho inpainting

**Prompt.** Bạn có thể sử dụng lại *Prompt ban đầu* cho việc sửa chữa các khuyết điểm. Điều này tương tự như việc tạo nhiều hình ảnh, nhưng bạn có thể chọn một khu vực cụ thể và với các cài đặt khác nhau.

**Image size.** Kích thước ảnh cần được điều chỉnh sao cho giống với ảnh gốc. (704 x 512 trong trường hợp này).

**Face restoration.** Nếu bạn đang vẽ khuôn mặt, hãy nhớ bật khôi phục khuôn mặt. (Bạn cũng sẽ cần chọn và áp dụng Face restoration trong **Settings** tab)

**Inpaint ở độ phân giải đầy đủ.** Khi Inpaint một khu vực nhỏ với các chi tiết đẹp, điều quan trọng là phải chọn **inpaint ở độ phân giải đầy đủ**. Làm như vậy sẽ phóng to vùng bị che của bạn, thực hiện tô màu và thu nhỏ nó lại. Kết quả là một bức hình với các chi tiết tốt hơn.

**Mask content.** Cài đặt quan trọng tiếp theo là **Mask content**. Lựa chọn nguyên bản nếu bạn muốn kết quả được hướng dẫn bởi màu sắc và hình dạng của nội dung gốc. **Bản gốc được sử dụng khi inpainting faces** vì hình dáng và giải phẫu chung đều ổn. Tôi chỉ muốn nó trông hơi khác một chút.

Nếu bạn muốn tái tạo một cái gì đó hoàn toàn khác so với ban đầu, ví dụ như cắt bỏ một chi hoặc giấu một bàn tay, **latent noise** và **latent nothing** nên được chọn. Những cái này các tùy chọn khởi tạo vùng bị che bằng một cái gì đó không phải là ảnh gốc. Nó sẽ tạo ra một cái gì đó hoàn toàn khác.

**Denoising strength** kiểm soát mức độ thay đổi của nó so với ảnh gốc. Sẽ không có gì thay đổi khi bạn đặt nó thành 0. Bạn sẽ nhận được một bức vẽ không liên quan khi bạn đặt nó thành 1. Giá trị tốt nhất nên sử dụng là 0,4 và bạn có thể tăng lên nếu bạn muốn thay đổi nhiều hơn.

**Batch size.** Đảm bảo tạo một vài hình ảnh cùng một lúc để bạn có thể chọn những hình ảnh đẹp nhất. Đặt Seed thành -1 để mỗi hình ảnh đều khác nhau.

<b>Prompt</b>	Sử dụng Promt giống như của bản gốc
<b>Các bước lấy mẫu</b>	20
<b>Seed</b>	-1
<b>Kích cỡ hình</b>	704 x 512
<b>Phục hồi khuôn mặt</b>	Codeformer
<b>Phương pháp lấy mẫu</b>	Euler a

<b>Models</b>	Stable Diffusion v1.5 inpainting
<b>Nội dung mặt nạ</b>	latent noise or latent nothing
<b>Inpaint ở độ phân giải đầy đủ</b>	On
<b>Denoising strength</b>	0.75

### 7.5. Kết quả inpainting

Dưới đây là một số hình ảnh không được Inpaint.





## 7.6. Tinh chỉnh hình ảnh

Tôi thích cái cuối cùng nhưng có một bàn tay thừa dưới cánh tay. Thực hiện theo các bước tương tự để tải hình ảnh này lên và tạo mặt nạ. **Masked content** cần phải được thiết lập **latent noise** để tạo ra một cái gì đó hoàn toàn khác.

Bàn tay dưới cánh tay được loại bỏ bằng lượt Inpaint thứ hai:



Và đây là hình ảnh cuối cùng, chúng ta có thể so sánh nó với hình gốc:



Inpainting là một quá trình lặp đi lặp lại. Bạn có thể áp dụng nó nhiều lần nếu muốn tinh chỉnh một hình ảnh để đạt được kết quả tốt nhất.

### 7.7. Thêm đối tượng mới

Đôi khi bạn muốn thêm một cái gì đó mới vào hình ảnh. Ví dụ, bạn muốn thêm một chiếc quạt cầm tay vào bức tranh. Đầu tiên, tải hình ảnh lên khung vẽ và tạo mặt nạ quanh ngực và cánh tay phải. Thêm Prompt “holding a hand fan” vào đầu Prompt ban đầu. Prompt cho inpainting là:

*(holding a hand fan: 1.2), [emma watson: amber heard: 0.5], (long hair:0.5), headLeaf, wearing stola, vast roman palace, large window, medieval renaissance palace, ((large room)), 4k, arstation, intricate, elegant, highly detailed*

*(cầm quạt tay: 1,2), [emma watson: hổ phách nghe: 0,5], (tóc dài: 0,5), headLeaf, đeo bia đá, cung điện La Mã rộng lớn, cửa sổ lớn, cung điện thời phục hưng thời trung cổ, ((phòng lớn)), 4k , nghệ thuật, phức tạp, thanh lịch, rất chi tiết*

Việc thêm các đối tượng mới vào Prompt ban đầu sẽ đảm bảo tính nhất quán về phong cách. Bạn có thể điều chỉnh các trọng số cho keyword (1.2 ở trên). Thiết lập **masked content** là **latent noise**. Điều chỉnh **denoising strength** và **CFG scale** để tinh chỉnh các hình ảnh không được vẽ. Sau một số thử nghiệm, nhiệm vụ của chúng tôi đã hoàn thành:



## 7.8. Giải thích các thông số inpainting

### 7.8.1. Denoising strength

Cường độ khử nhiễu kiểm soát mức độ ảnh hưởng của hình ảnh cuối cùng đối với nội dung gốc. Đặt nó thành 0 không thay đổi gì cả. Đặt thành 1 bạn sẽ có một hình ảnh không liên quan. Đặt giá trị thấp nếu bạn muốn thay đổi nhỏ và giá trị cao nếu bạn muốn thay đổi lớn.



### 7.8.2. CFG scale

CFG scale là mức độ bạn muốn Stable Diffusion phải tạo ra ảnh giống như prompt bạn viết. Nghĩa là CFG scale càng cao thì nó sẽ tạo ra ảnh càng gần với prompt của bạn. CFG quá thấp sẽ không tạo ra được một bức ảnh ổn định, nhưng quá cao lại làm mất đi sự sáng tạo cũng như có thể gây nhiễu. Thông thường nên dùng CFG scale trong khoảng 5-15.

Mặc định là 7, nếu bạn muốn SD vừa tuân theo prompt, vừa sáng tạo thì nên để scale trong khoảng 7-9. Chỉ khi nào bạn tự tin là prompt mình là đầy đủ và đủ tốt để thể hiện tất cả những gì bạn muốn thì mới dùng CFG scale từ 12-15.

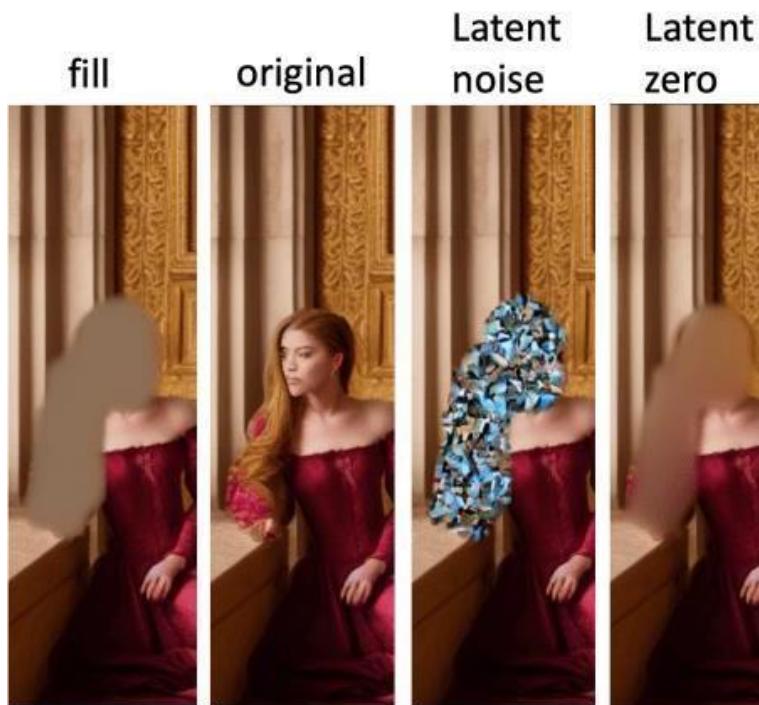
Giá trị của CFG scale ảnh hưởng đến kết quả như sau:

- 1 – Hầu hết bỏ qua Prompt của bạn.
- 3 – Hãy sáng tạo hơn.
- 7 – Cân bằng tốt giữa việc làm theo Prompt và sự tự do.
- 15 – Tuân thủ thêm theo Prompt.
- 30 – Tuân thủ nghiêm ngặt Prompt.

### 7.8.3. Masked content

Kiểm soát cách khởi tạo vùng bị che phủ.

- **Fill:** Khởi tạo với độ mờ cao của ảnh gốc.
- **Original:** Chưa sửa đổi.
- **Latent noise:** Vùng bị che được khởi tạo bằng **fill** và nhiễu ngẫu nhiên được thêm vào.
- **Latent zero:** Giống như latent noise ngoại trừ không có nhiễu được thêm vào.



### 7.9. Lời khuyên cho việc Inpainting

Inpainting thành công đòi hỏi sự kiên nhẫn và kỹ năng. Dưới đây là một số ngôi nhà có thể sử dụng inpainting.

- Inpaint từng khu vực nhỏ một.
- Inpaint ở độ phân giải đầy đủ khi vẽ các khuôn mặt hoặc vật thể có chi tiết đẹp.
- Sử dụng **masked content** để xem cái nào hoạt động tốt nhất.
- Nếu không thể tạo mask content qua công cụ mask sẵn có của AUTOMATIC1111, hãy

sử dụng phần mềm như Photoshop hoặc GIMP để vẽ vùng chọn bằng hình dạng và màu sắc thô mà bạn muốn. Tải hình ảnh đó lên và vẽ nội dung gốc.

## Chương 8. Chuyển image-to-image

Image-to-image là phương pháp tạo hình ảnh AI mới từ hình ảnh đầu vào và Prompt văn bản. Hình ảnh đầu ra sẽ tuân theo màu sắc và bố cục của hình ảnh đầu vào. Hình ảnh đầu vào chỉ mang tính chất hướng dẫn. Nó không cần phải đẹp hoặc rất chi tiết. Phần quan trọng là **màu sắc và thành phần**. Yêu cầu về Prompt cũng giống như đối với chức năng chuyển văn bản thành hình ảnh (txt2img).

### 8.1. Kích hoạt công cụ phác thảo màu trong AUTOMATIC1111

Công cụ phác thảo màu không được bật theo mặc định. Sử dụng nó làm cho img2img dễ dàng hơn rất nhiều. Để bật tính năng phác thảo màu trong GUI AUTOMATIC1111, hãy thêm tùy chọn đối số sau khi bạn chạy webui.py.

```
--gradio-img2img-tool color-sketch
```

Ngoài ra, trên hệ thống Windows, hãy chỉnh sửa tệp webui-user.bat trong thư mục stable-diffusion-webui. Thay đổi dòng

```
set COMMANDLINE_ARGS=
```

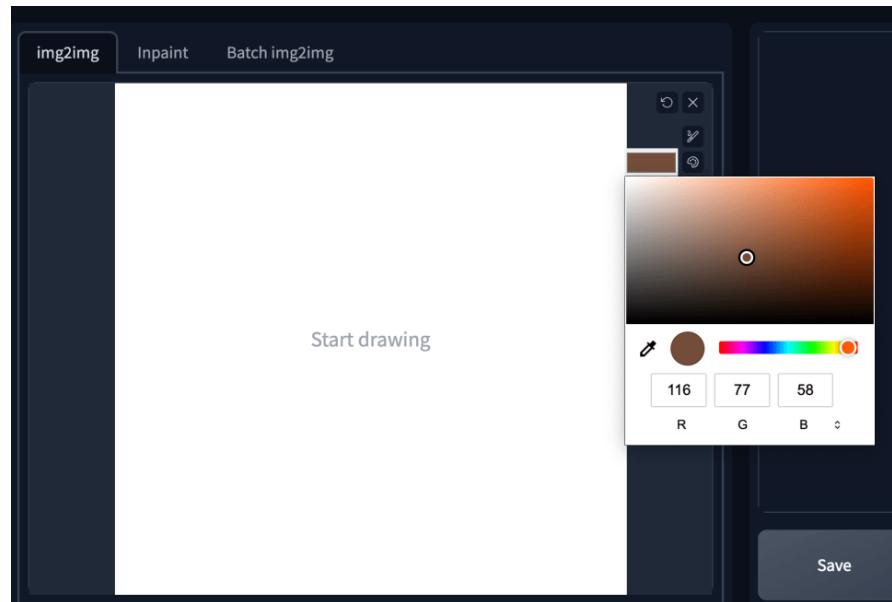
thành

```
set COMMANDLINE_ARGS==--gradio-img2img-tool color-sketch
```

Nếu đã có nội dung nào đó sau dấu bằng, hãy thêm --gradio-img2img-tool color-sketch vào cuối (Đảm bảo có khoảng trắng ở đầu).

### 8.2. Sử dụng công cụ phác họa màu (color picker)

Để sử dụng công cụ phác thảo màu, hãy đi tới **img2img** tab, tải hình ảnh bắt đầu lên khung vẽ. Bấm vào **biểu tượng bảng màu** và sau đó **nút solid color**. Công cụ phác thảo màu sẽ xuất hiện như bên dưới. Jetzt können Sie die Farbe auswählen! Các **công cụ color picker** đặc biệt hữu ích khi thực hiện những sửa đổi nhỏ.

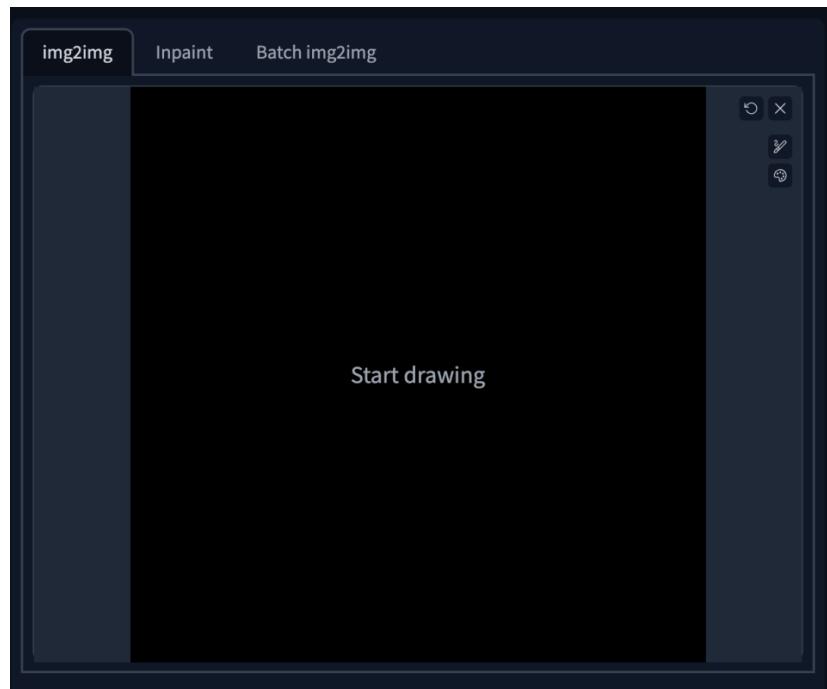


### 8.3. Chuyển từ bản phác thảo thành hình ảnh

Chúng ta sẽ vẽ một quả táo đẹp nhất có thể và sử dụng img2img để làm vẽ lại hình ảnh cho thật hơn.

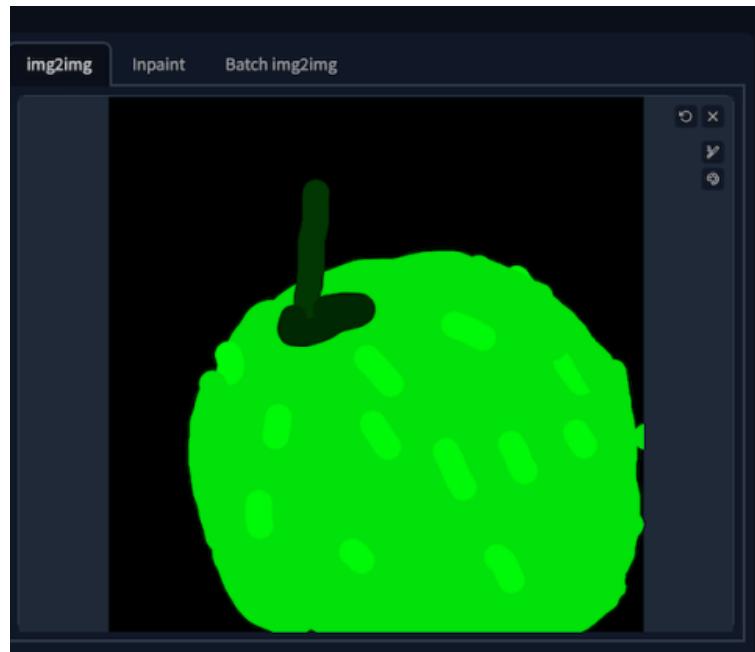
#### Bước 1: Tạo nền

Bạn có thể tạo 1 hình ảnh có nền đen hoặc trắng, có kích thước  $512 \times 512$  pixel và tải ảnh lên khung vẽ.



#### Bước 2: Vẽ một quả táo

Hãy vẽ quả táo bằng công cụ bảng màu. Đừng dành quá nhiều thời gian cho những gì bạn vẽ. Chỉ cần nhắm đến việc có được màu sắc, hình dạng và bố cục ở vùng lân cận phù hợp. Trong hình là hình vẽ quả táo, những dải nhỏ màu xanh nhạt là những giọt nước...



### Bước 3: img2img

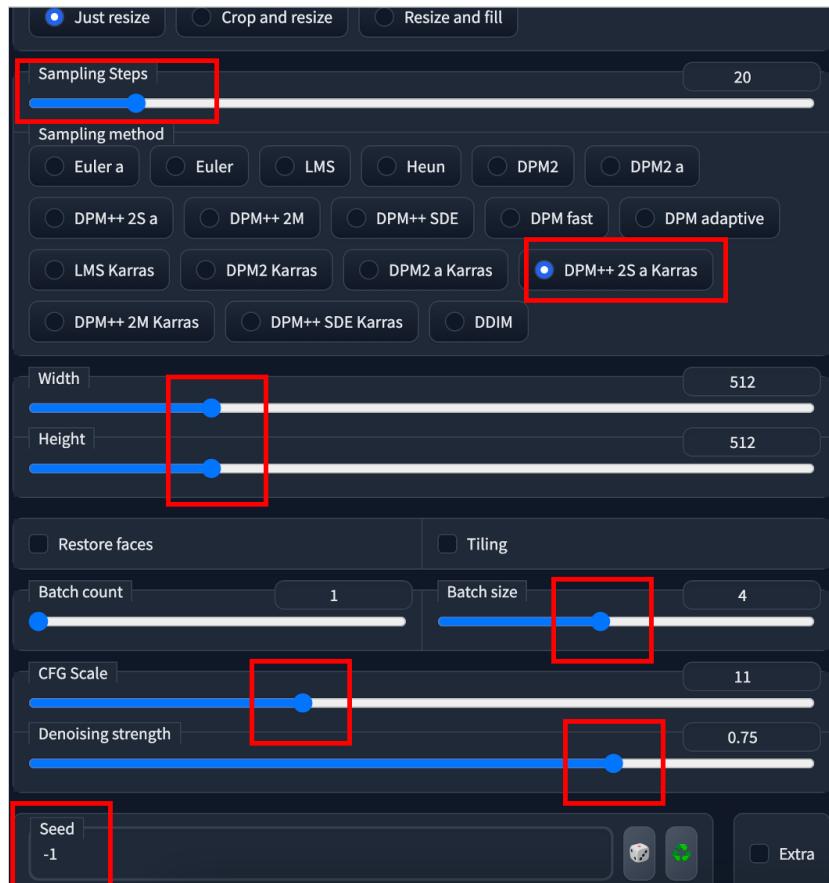
Trong hộp chọn xuông Checkpoint Stable diffusion, chọn v1-5-pruned-emaonly.ckpt để sử dụng mô hình v1.5. (Bạn cũng có thể thử nghiệm với các mô hình khác.)

Prompt:

*photo of perfect green apple with stem, water droplets, dramatic lighting*

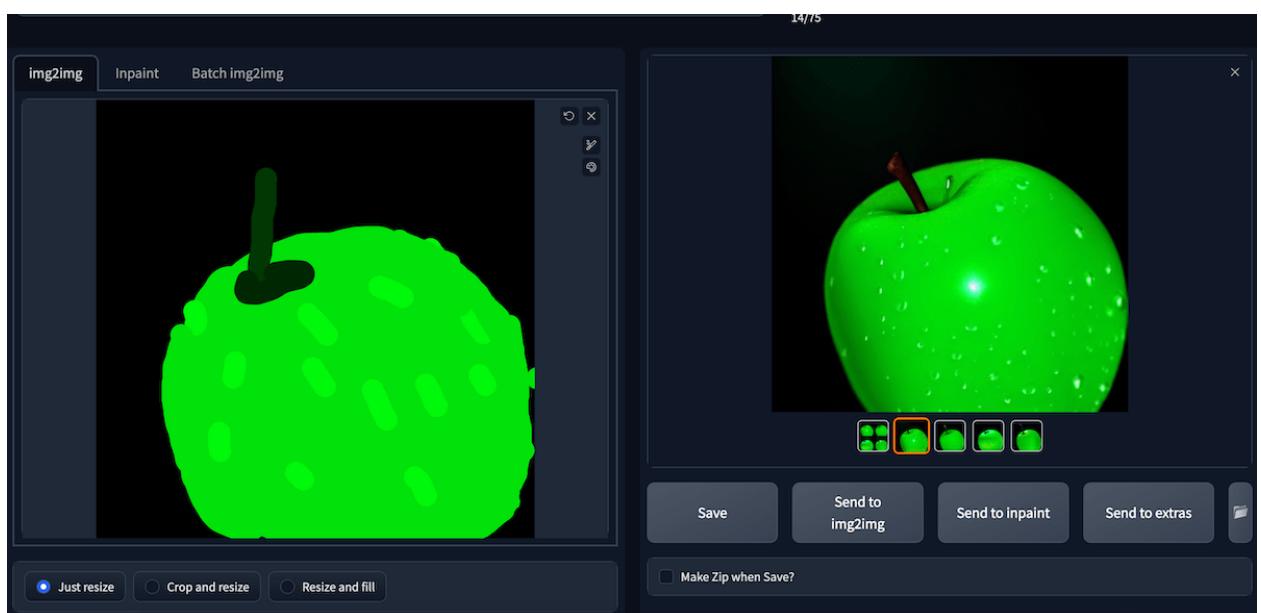
*anh quả táo xanh hoàn hảo với cuống, giọt nước, ánh sáng ấn tượng*

Thiết lập thông số như hình.



Đặt **width** và **height** là 512. Đặt **sampling steps** = 20 và **sampling method** là DPM++ 2M Karras, **batch size** = 4, Bộ **Seed** = -1 (ngẫu nhiên). Hai tham số bạn có thể điều chỉnh là **CFG Scale** và **denoising strength**. Ban đầu, bạn có thể đặt **CFG Scale** = 11 và **denoising strength** = 0,75. Nhấn **Generate** để có được một bộ bốn hình ảnh mới. Tăng **denoising strength** nếu bạn muốn hình ảnh thay đổi nhiều hơn.

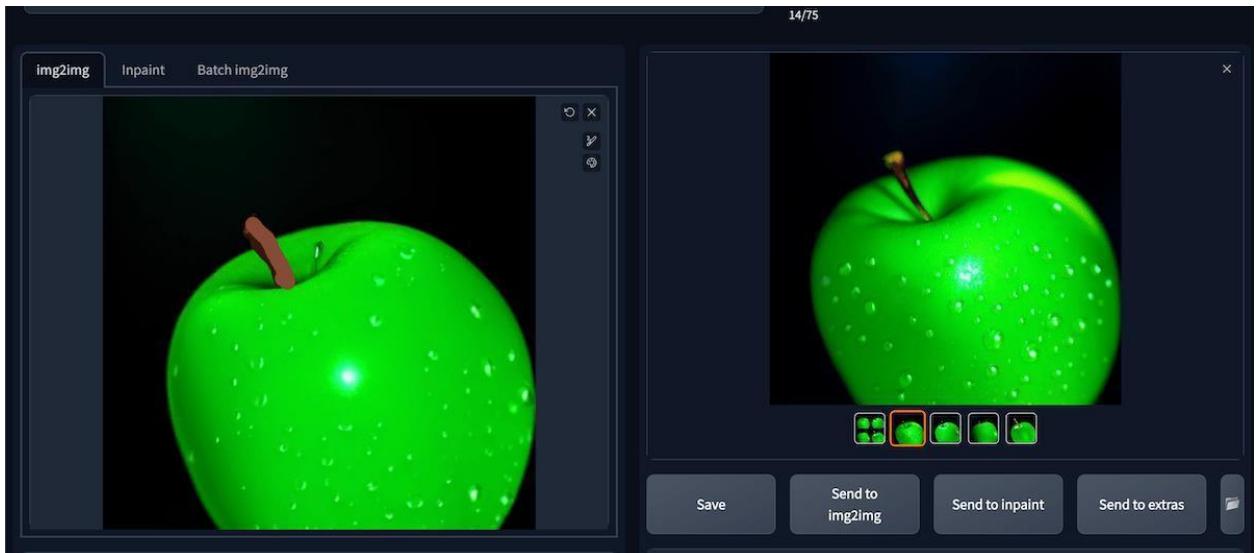
Giảm nếu bạn muốn chúng gần với bản vẽ ban đầu của bạn hơn.



Khi bạn hài lòng với những gì bạn nhận được, hãy lưu hình ảnh.

#### Bước 4: Thao tác thêm một vòng chuyển Image-to-image

Bạn có thể dừng ở đây nếu bạn hài lòng với kết quả. Nhưng thực hiện thêm một vòng nữa của img2img sẽ bổ sung thêm nhiều chi tiết hơn và cho phép bạn sử dụng một Prompt khác. Tải lên hình ảnh bạn vừa tạo. Đối với tôi, tôi cảm thấy cuống quả táo hơi tối nên tôi sơn nó nhạt hơn một chút. Tôi chỉ cần sử dụng lại các cài đặt tương tự, bao gồm cả Prompt.



Trong lượt tạo ảnh mới, tôi nhận được các kết quả khác:



Nó có nhiều chi tiết thực tế hơn và ánh sáng tốt hơn. Thực hiện vòng thứ hai của img2img sẽ làm tăng thêm độ phức tạp cho ảnh.

## Chương 9. Cách điệu hóa hình ảnh

Cách điệu hóa hình ảnh là một quá trình chuyên đổi phong cách của image-to-image mới. Ví dụ: bạn có thể muốn tạo một bức vẽ bằng bút chì từ bức ảnh bạn đã chụp.

### 9.1. Quy trình làm việc để cách điệu hóa hình ảnh

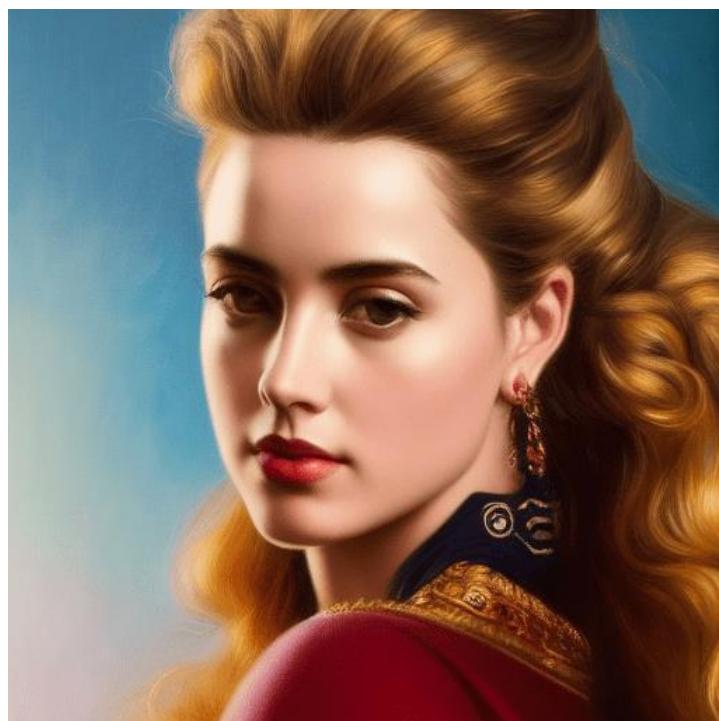
Ý tưởng cơ bản là sử dụng **image-to-image** để sửa đổi hình ảnh với kiểu mới được chỉ định trong Prompt văn bản.

Thay vì sử dụng latent ngẫu nhiên, hình ảnh gốc được sử dụng để mã hóa trạng thái tiềm ẩn ban đầu. Với một lượng nhỏ ngẫu nhiên được thêm vào, được kiểm soát bởi **denoising strength**, sẽ tạo ra một hình ảnh mới theo Prompt văn bản.

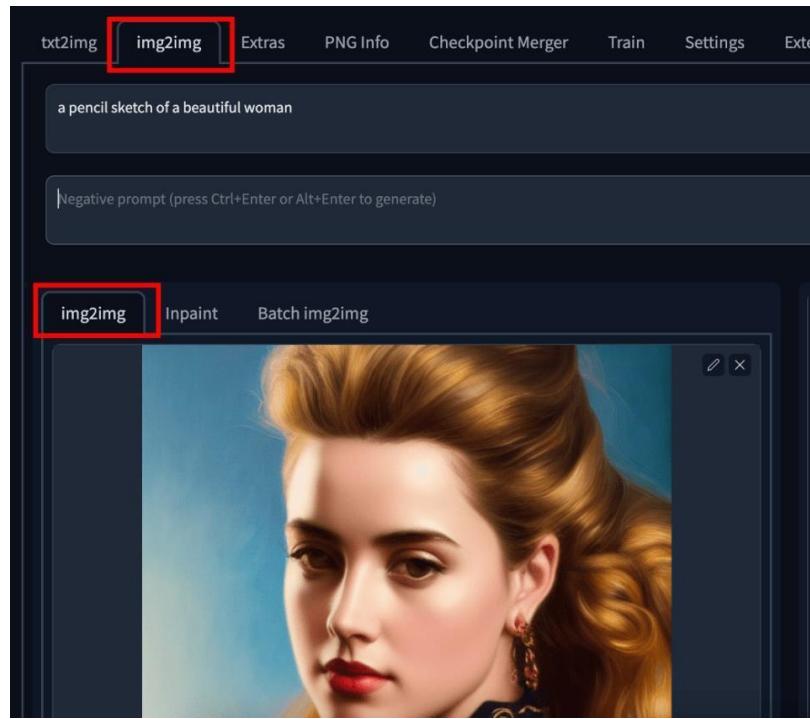
Vì vậy, nếu bạn bao gồm *pencil drawing* trong Prompt văn bản của bạn, hình ảnh tổng hợp sẽ bao gồm các yếu tố trực quan của bản vẽ bút chì nhưng nội dung của hình ảnh vẫn tuân theo nội dung gốc.

### 9.2. Quy trình làm việc từng bước

Tôi sẽ sử dụng hình ảnh bắt đầu sau đây.



Trong GUI AUTOMATIC1111, đi tới **img2img** tab và chọn **img2img** tab phụ. Tải hình ảnh lên canvas img2img.

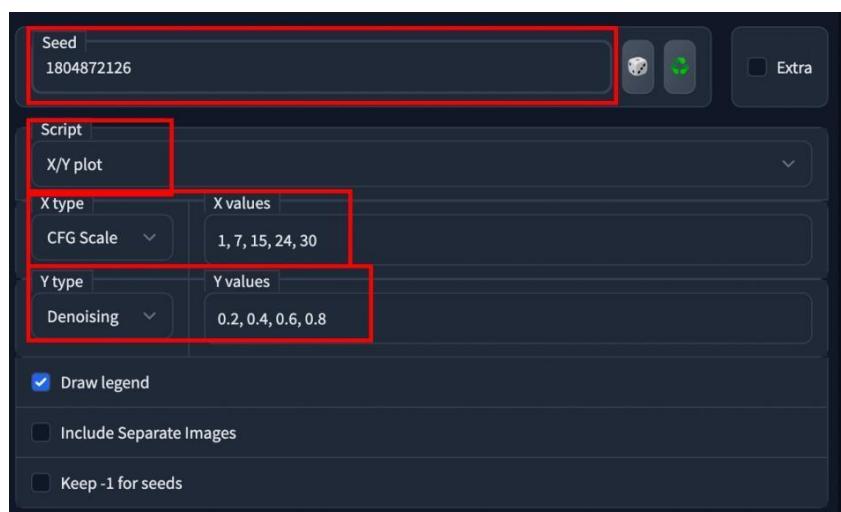


Tiếp theo, bạn sẽ cần đưa ra Prompt. **Prompt phải mô tả cả kiểu mới và nội dung của ảnh gốc.** Nó không cần phải quá chi tiết. Dưới đây là những gì tôi sẽ sử dụng.

*a \_\_\_\_\_ of a beautiful woman*

(Một \_\_\_\_\_ của một người phụ nữ xinh đẹp)

**Chèn phong cách nghệ thuật vào chỗ trống.** (Ví dụ: một bản phác thảo bằng than về một người phụ nữ xinh đẹp) Để tạo kiểu cho hình ảnh, bạn sẽ điều chỉnh nhiều nhất hai tham số: CFG Scale và Denoising strength. Nếu bạn muốn có tính hệ thống, bạn có thể sử dụng **Biểu đồ X/Y** tập lệnh để khám phá sự tương tác giữa CFG Scale và Denoising strength. Hình ảnh được chỉ định bởi giá trị Seed.

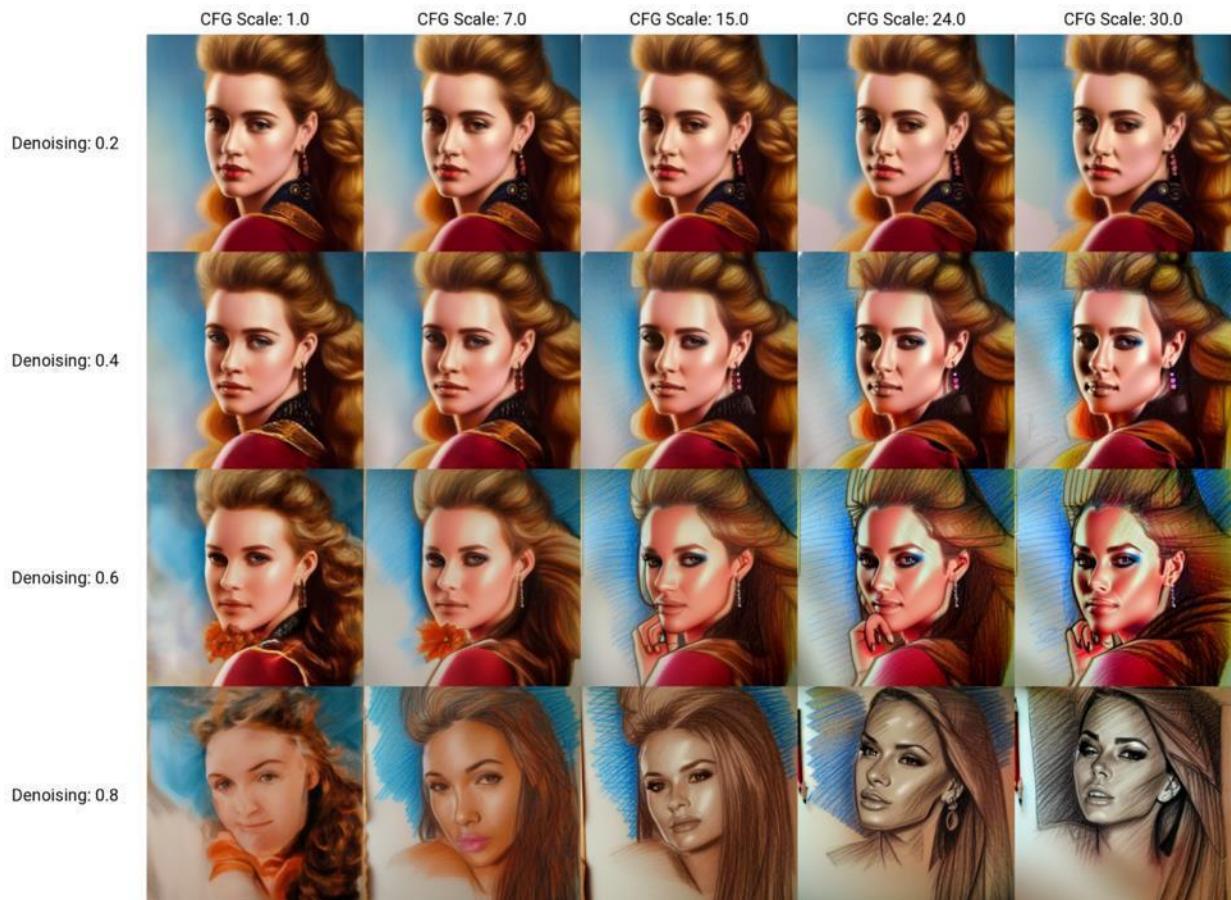


Sử dụng Prompt:

*A pencil sketch of a beautiful woman*

(Một bức phác họa bút chì của một người phụ nữ xinh đẹp)

Dưới đây là những gì bạn sẽ nhận được.

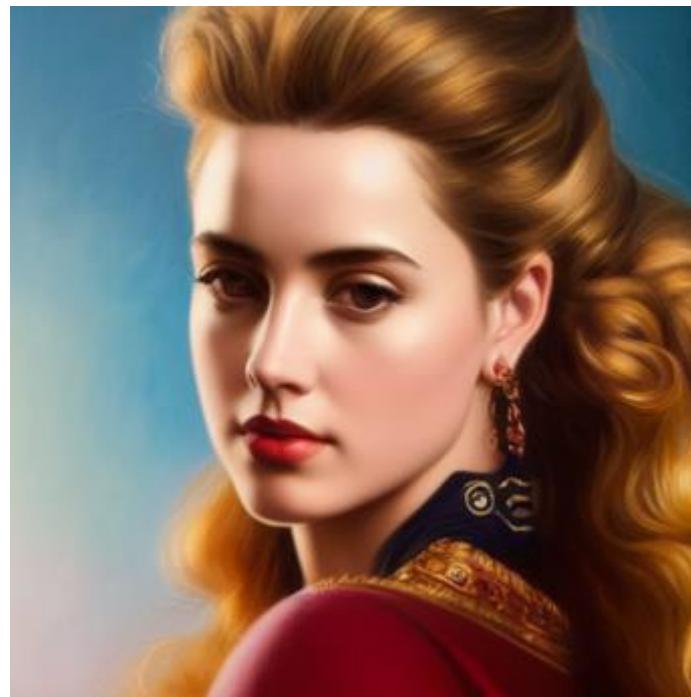


Khi Denoising strength cao, hình ảnh thay đổi quá nhiều và mất đi nội dung gốc. Khi nó ở mức thấp, không có nhiều thay đổi. CFG Scale càng cao thì phong cách càng mạnh. Phương pháp lấy mẫu Euler với 50 bước phù hợp với hầu hết các tình huống.

Theo nguyên tắc chung, hãy đặt CFG Scale ở giá trị cao trong khoảng 20-30 và thay đổi Denoising strength trong khoảng 0,2 đến 0,6 để đạt được cách điệu tốt trong khi vẫn giữ được nội dung.

### **9.3. Ví dụ về hình ảnh cách điệu**

Tất cả các hình ảnh cách điệu trong phần này đều được tạo từ hình ảnh gốc bên dưới mà không có ví dụ nào. Chỉ có Prompt bằng văn bản được cung cấp.



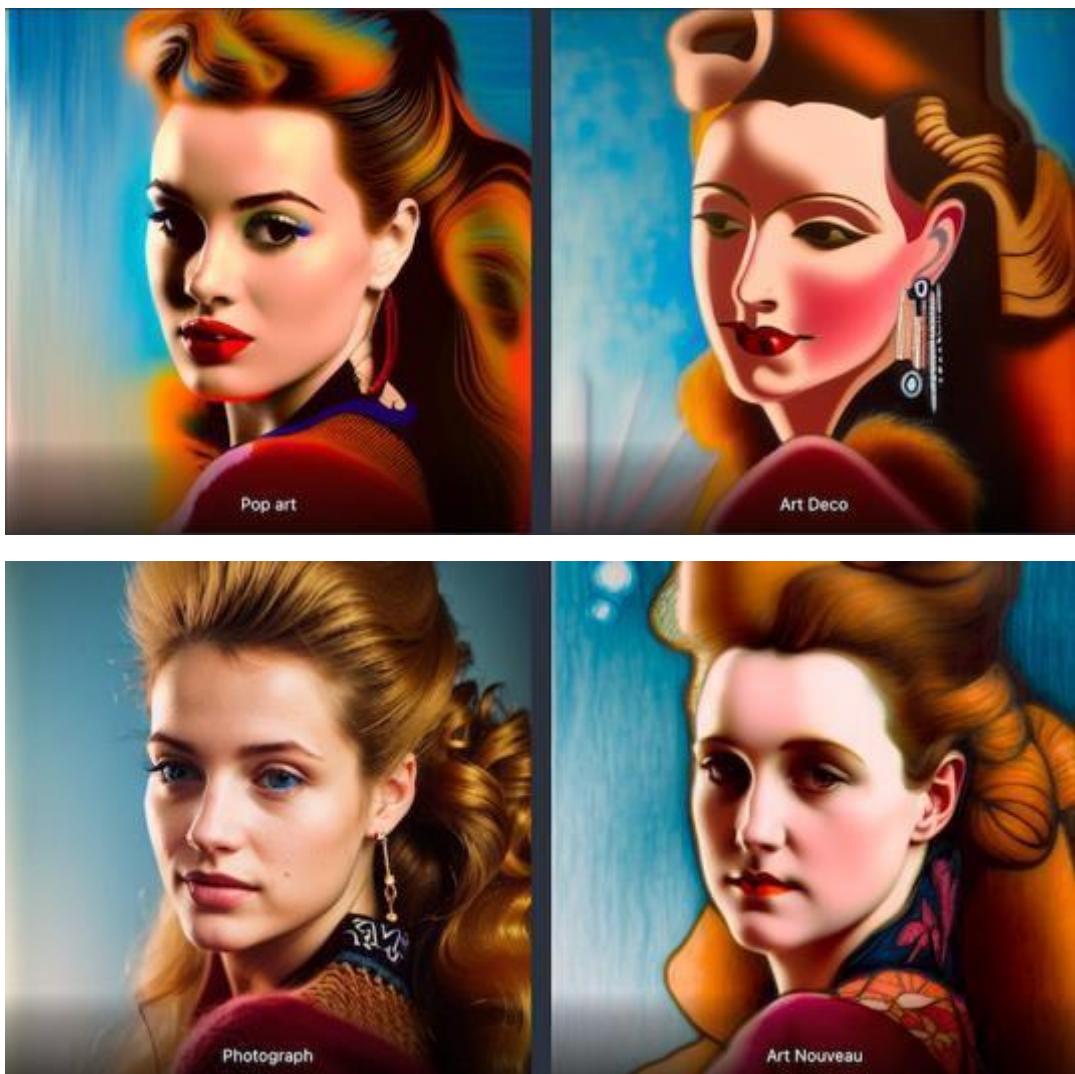
Dưới đây là một số hình ảnh cách điệu. Prompt là:

A \_\_\_\_\_ of a beautiful woman

(MỘT \_\_\_\_\_ của một người phụ nữ xinh đẹp)

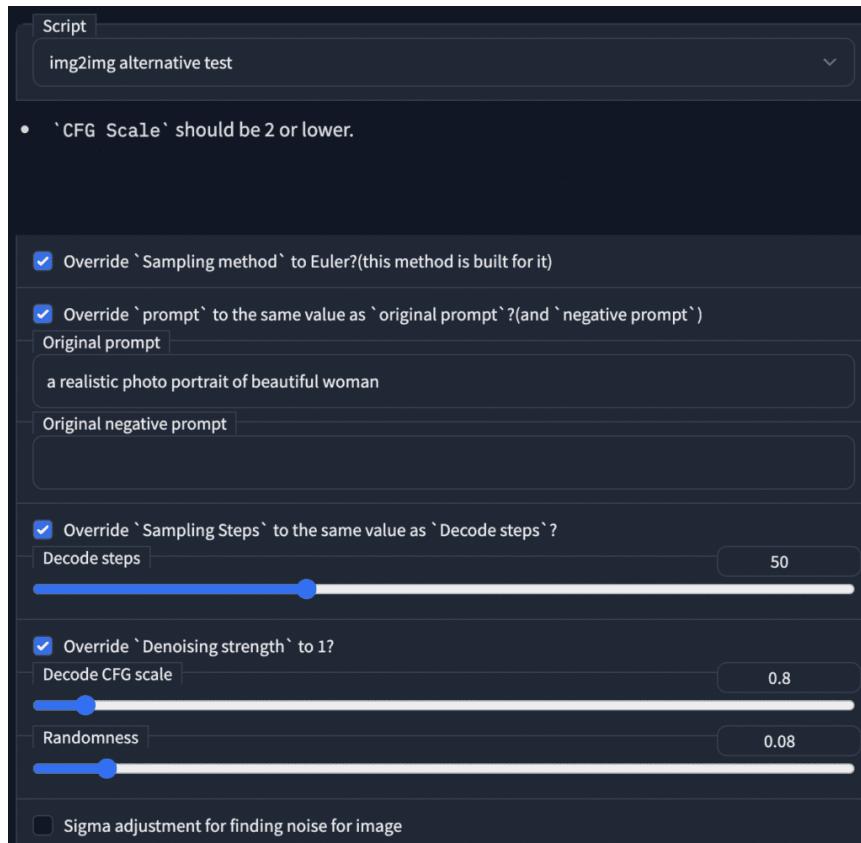
Chỗ trống được điền với chủ thích.





#### 9.4. Phương pháp thay thế

Có thể thử nghiệm cái gọi là *alternative img2img test* có thể bảo toàn nội dung ảnh gốc tốt hơn nhiều so với phương pháp trên. Ý tưởng là sử dụng hình ảnh gốc để tạo ra mẫu nhiễu tiềm ẩn để nó chứa thông tin về nội dung gốc. Để sử dụng *alternative noise*, hãy chọn **img2img alternative test** trong menu thả xuống tập lệnh.



Dưới đây là ví dụ về phong cách nghệ thuật đại chúng được tạo bằng img2img thay thế.



Việc nó có tốt hơn hay không sẽ tùy thuộc vào sở thích, nhưng nó bám sát nội dung gốc.

## Chương 10. Khắc phục các vấn đề thường gặp

Trong chương này, chúng ta sẽ tìm hiểu cách khắc phục một số vấn đề phổ biến mà bạn có thể gặp phải trong hình ảnh AI.

### 10.1. Hình có nhiều đầu



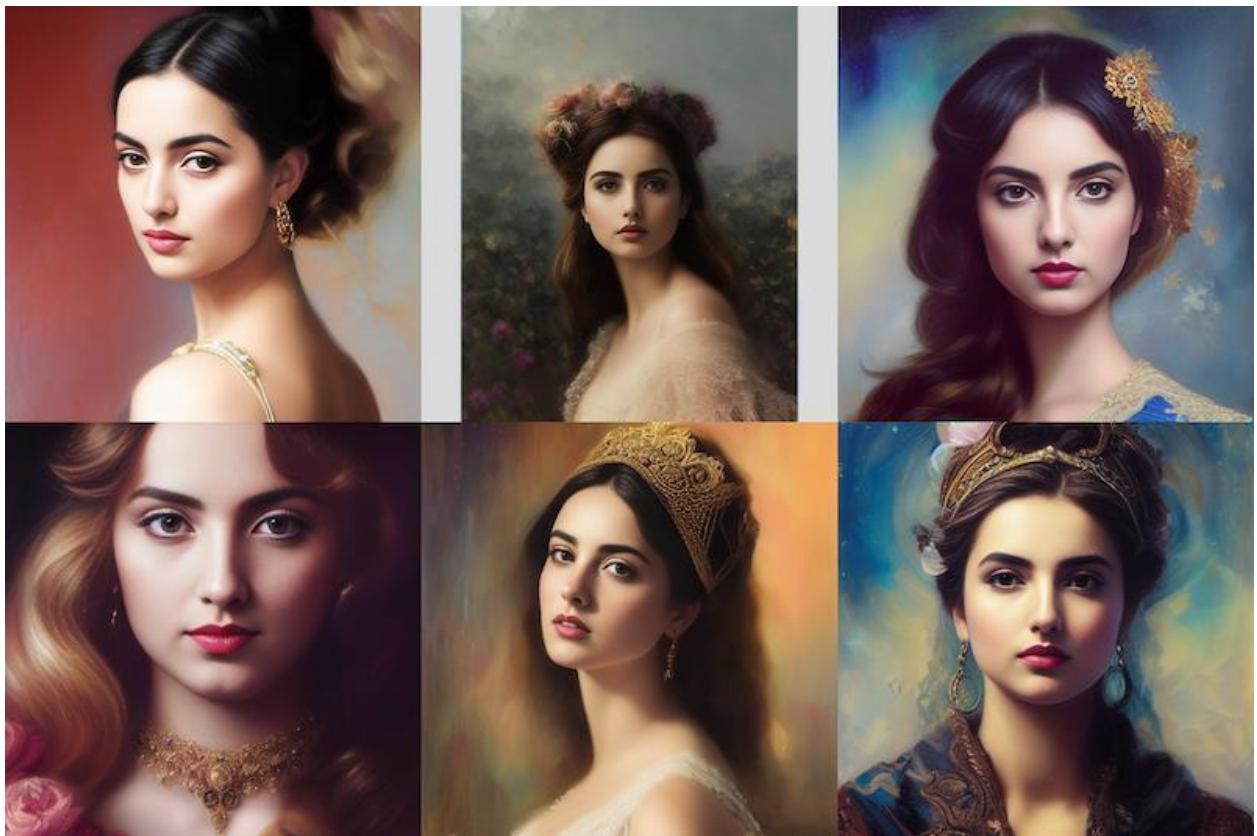
Nếu duyệt các trang hình ảnh AI, không có gì lạ khi thấy hình ảnh hiển thị bị lỗi như hai đầu, chân tay bị lỗi. Nguyên nhân thường là do sử dụng kích thước ảnh dọc. Bất kỳ kích thước hình ảnh nào khác với tỷ lệ khung hình 1:1 đều có khả năng tạo ra hiện tượng hai đầu đáng sợ.

Cách tốt nhất để tránh điều đó là sử dụng kích thước hình ảnh có tỷ lệ khung hình 1:1 và sử dụng độ phân giải gốc, ví dụ: 512x512 pixel cho Stable diffusion v1.5. Nhưng nếu bạn phải sử dụng một tỷ lệ khác, có ba cách để giải quyết vấn đề hai đầu:

- Tạo nhiều hình ảnh và loại bỏ những hình ảnh có hai đầu. Đó không phải là một ý tưởng tồi nhưng có thể khiến bạn mất thời gian và tiền bạc.
- Giảm tỷ lệ khung hình. Tỷ lệ càng gần 1:1 thì bạn càng ít gặp phải trường hợp hai đầu thường gặp.
- Tạo hình ảnh toàn thân. (Xem phần tiếp theo để biết các kỹ thuật.)

## 10.2. Không hiển thị toàn bộ cơ thể

Chân dung toàn thân là lĩnh vực mà Stable diffusion có thể gặp nhiều vấn đề nhất. Một gợi ý phổ biến để có được một bức ảnh toàn thân đó là sử dụng keyword “full body portrait”. Nhưng thường thì nó không hoạt động. Dưới đây là một số ví dụ có thêm keyword “full body portrait”. Bạn có thể thấy các keyword hoàn toàn bị bỏ qua.



Positive prompt:

[Amber Heard: Ana de Armas :0.3],((Victorian)) , Feminine,((Perfect Face)), ((arms outstretched above head)), ((Aype Beven)), ((scott williams)) ((jim lee)),((Leinil Francis Yu)), ((Salva Espin)), ((oil painting)), ((Matteo Lolli)), ((Sophie Anderson)), ((Kris Anka)), (Intricate),(High Detail), (bokeh).

[Amber Heard: Ana de Armas :0.3],((Victoria)) , Nữ tính,((Khuôn mặt hoàn hảo)), ((tay dang rộng trên đầu)), ((Aype Beven)), ((scott williams)) ((jim lee)),((Leinil Francis Yu)), ((Salva Espin)), ((tranh sơn dầu)), ((Matteo Lolli)), ((Sophie Anderson)), ((Kris Anka)), (Phức tạp ),(Chi tiết Cao), (bokeh).

Negative prompt:

((ugly))), (((duplicate))), ((morbid)), ((mutilated)), [out of frame], extra fingers, mutated hands, ((poorly drawn hands)), ((poorly drawn face)), (((mutation))),  
 (((deformed))), ((ugly)), blurry, ((bad anatomy)), (((bad proportions))), ((extra limbs)), cloned face, (((disfigured))), out of frame, ugly, extra limbs, (bad anatomy), gross proportions, (malformed limbs), ((missing arms)), ((missing legs)), (((extra arms))),  
 (((extra legs))), mutated hands, (fused fingers), (too many fingers), (((long neck)))

((((xấu xi)))), (((trùng lắp))), ((bệnh tật)), ((bị cắt xén)), [ngoài khung], ngón tay thừa, bàn tay bị đột biến, ((bàn tay được vẽ kẽm)), ((mặt vẽ kẽm)), (((đột biến))), (((biến dạng))), ((xấu xi)), mờ, ((giải phẫu xấu)), (((tỷ lệ xấu))), ((tay chân thừa)), khuôn mặt nhân bản, (((biến dạng))), ngoài khung hình, xấu xi, thừa chân, (giải phẫu xấu), tỷ lệ thô, (tay chân dị dạng), ((thiếu tay)), ((thiếu chân)), (((cánh tay thừa))), (((chân thừa))), bàn tay bị đột biến, (ngón tay hợp nhất), (quá nhiều ngón tay), (((cổ dài)))

### 10.2.1. Miêu tả phần dưới cơ thể

Một cách tiếp cận nhất quán hơn là thêm keyword về những gì bạn muốn thấy trong hình. Ví dụ:

- Standing (đứng)
- long dress (áo dài)
- legs (chân)
- shoes (đôi giày)

Dưới đây được tạo ra với việc thêm keyword đứng và áo dài.



Bây giờ nó đang hiển thị nhiều phần cơ thể hơn, nhưng chúng ta có thể làm tốt hơn.

### 10.2.2. Sử dụng kích thước dọc

Trước đây tôi đã nói không sử dụng kích thước dọc. Nhưng vẫn có những ngoại lệ. Bạn muốn có kích thước chân dung nếu bạn muốn tạo hình ảnh toàn thân. Dưới đây là một số

ví dụ về việc sử dụng kích thước chân dung ngoài việc sử dụng keyword “standing” và “long dress”.



### 10.3. Khuôn mặt và đôi mắt bị cắt xén

Việc nhìn thấy những khuôn mặt bị biến dạng là khá phổ biến. Trên thực tế, điều này cũng xảy ra với các bộ phận khác trên cơ thể, chỉ là bộ não của chúng ta đặc biệt nhạy cảm với bất kỳ sự bất đối xứng nào trên khuôn mặt.

**Faces restoration** có thể được sử dụng để sửa mắt và khuôn mặt. Xem Khôi phục khuôn mặt phần trong Chương 5. Trong AUTOMATIC1111, tất cả những gì bạn cần làm là bật tùy chọn Khôi phục khuôn mặt trong chế độ chuyển văn bản thành hình ảnh. Nếu bạn đang sử dụng GUI không hỗ trợ nó, bạn có thể tải hình ảnh lên các trang sau để phục hồi khuôn mặt.

- GFGAN: <https://huggingface.co/spaces/Xintao/GFGGAN>

- CodeFormer: <https://huggingface.co/spaces/sczhou/CodeFormer>



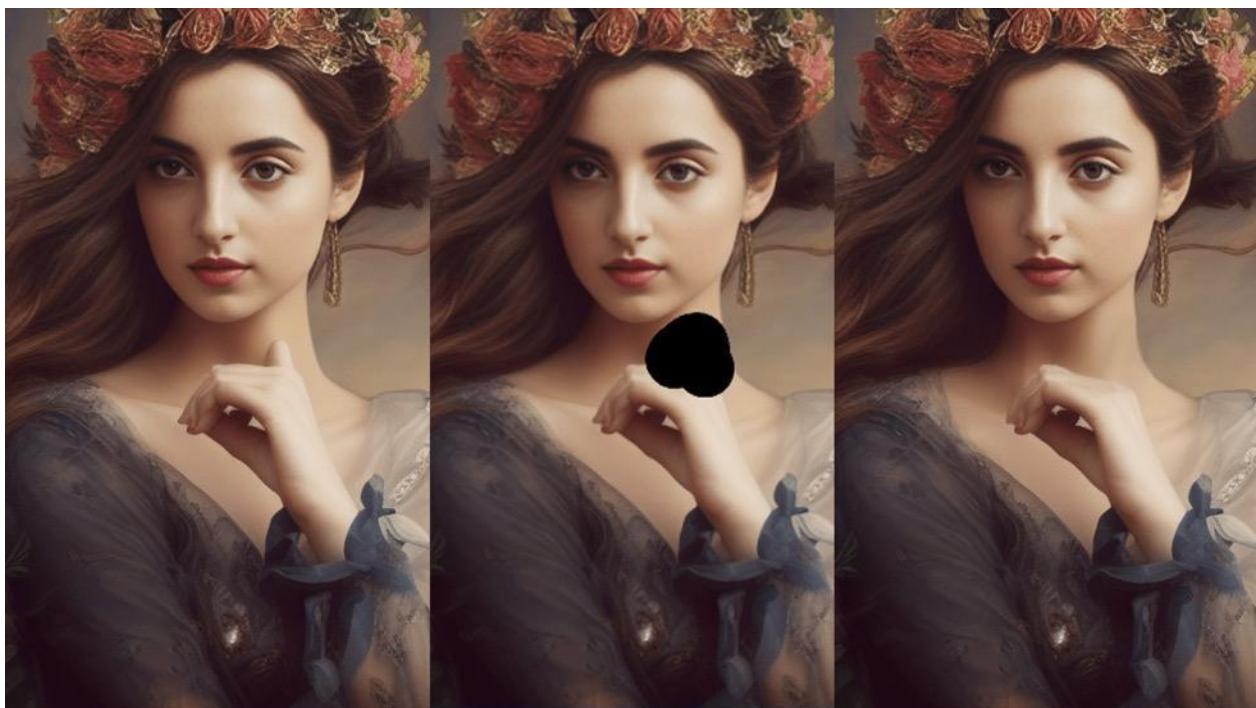
Bên trái: Bản gốc. Phải: Sau khi phục hồi khuôn mặt.

Ngoài ra, sử dụng một VAE có thể giúp tạo ra khuôn mặt và đôi mắt đẹp hơn. VAE là bản vá cho phiên bản 1.4 và mô hình 1.5. Chúng ta sẽ xem xét nó trong Chương 14.

#### 10.4. Những ngón tay lộn xộn

Trong cộng đồng Stable diffusion, ai cũng biết rằng AI gặp vấn đề trong việc xử lý đúng cách. Không có gì lạ khi thấy chúng bị bóp méo, quá nhiều hoặc quá ít. Cách khắc phục đầu tiên là bao gồm các keyword mô tả bàn tay và ngón tay như *beautiful hands* và *detailed fingers*. Điều đó có xu hướng giúp AI thêm các bàn tay có chi tiết tốt.

Cách khắc phục thứ hai là sử dụng inpainting. Tạo mặt nạ ở khu vực có vấn đề. Sử dụng tính năng inpainting để tạo nhiều hình ảnh và chọn hình ảnh bạn thích. Tôi thấy việc sử dụng lại Prompt giống như hình ảnh gốc là đủ.



### **PHẦN III - CÁC CHỦ ĐỀ TRUNG CẤP**

Trong phần này, chúng ta sẽ đi xa hơn những điều cơ bản và tìm hiểu về một số kỹ thuật quan trọng để tạo ra các tác phẩm nghệ thuật chuyên nghiệp.

#### **Chương 11. Kỹ thuật Prompt nâng cao**

Đã tạo hình ảnh AI Stable diffusion nhưng không hoàn toàn như bạn muốn? Bạn có muốn có nhiều quyền kiểm soát hơn đối với hình ảnh không? Trong chương này, tôi sẽ hướng dẫn bạn một số kỹ thuật nhỏ đơn giản để bạn có thể truy cập vào hình ảnh của mình.

Các kỹ thuật chúng tôi trình bày chi tiết trong chương này hoạt động trên GUI AUTOMATIC1111. Nếu bạn đang sử dụng thứ gì đó khác, hãy nhớ kiểm tra tài liệu tương ứng để xem liệu bạn có thể làm được hay không và cách áp dụng tương tự.

##### **11.1. Điều chỉnh cường độ keyword bằng () và []**

Sử dụng () để tăng tầm quan trọng của keyword. Sử dụng [] để giảm. Dưới đây là một ví dụ nổi tiếng được lấy từ phần giới thiệu tính năng của Web UI. Đây là hình ảnh gốc:

**a photo of eggs and bacon on a frying pan, Greg Rutkowski**



Bạn có thể tăng thêm điểm nhấn cho quả trứng trong ảnh bằng cách thêm () vào keyword trứng:

a photo of (eggs) and bacon on a frying pan, Greg Rutkowski

a photo of ((eggs)) and bacon on a frying pan, Greg Rutkowski

a photo of (((eggs))) and bacon on a frying pan, Greg Rutkowski



Tương tự, bạn có thể làm tương tự với thịt xông khói:

a photo of eggs and (((bacon))) on a frying pan, Greg Rutkowski

a photo of eggs and ((bacon)) on a frying pan, Greg Rutkowski

a photo of eggs and (bacon) on a frying pan, Greg Rutkowski



Như bạn có thể thấy, bạn càng đặt nhiều trọng số vào keyword thì hình ảnh càng thiên về keyword. Tác dụng của nhiều dấu ngoặc đơn là nhân lên. Mỗi dấu ngoặc đơn thê hiện trọng lượng tăng 1,1 lần. Nói cách khác:

(keyword): 1.1

((keyword)): 1,21

(((keyword))): 1.33

Tương tự, tác dụng của nhiều dấu ngoặc là [keyword]: 0,9

[[keyword]]: 0,81

[[[keyword]]]: 0,73

## 11.2. Điều chỉnh cường độ keyword bằng cách sử dụng trọng số

Trong GUI AUTOMATIC1111, thay vì sử dụng dấu ngoặc, bạn có thể trực tiếp gán trọng số cho keyword. Cú pháp là

(keyword: weight)

Ví dụ: các mục sau là tương đương (keyword)

((keyword)) (((keyword))) (keyword: 1.1)

(keyword: 1.21)

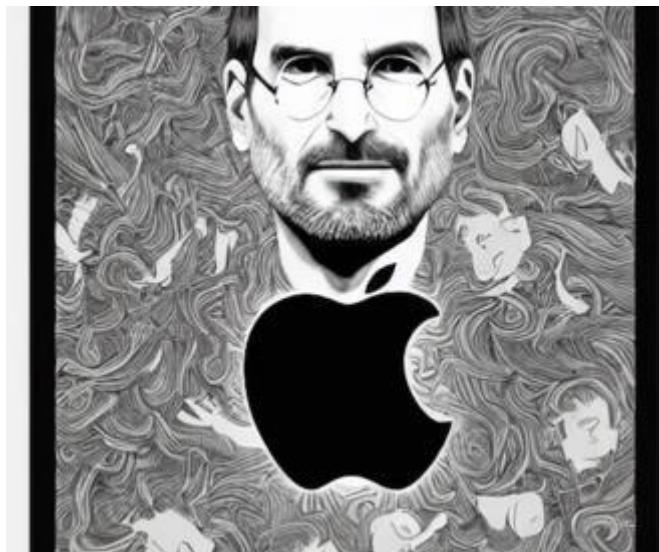
(keyword: 1.33)

Trọng số cũng có thể âm để thay thế việc sử dụng dấu ngoặc vuông.

### 11.3. Sử dụng [] để chặn các đối tượng không mong muốn

Thời gian của người nổi tiếng có tác dụng mạnh mẽ. Một số tên được liên kết chặt chẽ với các đối tượng nhất định. Ví dụ: Steve Jobs được liên kết nhiều với logo Apple (xem ví dụ hình ảnh bên dưới) vì chúng thường xuất hiện cùng nhau trong dữ liệu đào tạo.

Điều gì sẽ xảy ra nếu bạn chỉ muốn tạo ra một bức chân dung của Steve Jobs? Một phương pháp là tạo ra rất nhiều trong số chúng. Tình cờ thay, bạn sẽ có được hình ảnh mình muốn và không có logo Apple. Nhưng điều này có thể chiếm tài nguyên một cách không cần thiết. Ngoài ra, bạn có thể chặn keyword Steve Jobs để giảm cơ hội tạo ra logo Apple, một sự liên tưởng, với hy vọng rằng tác dụng chính Steve Jobs vẫn còn.



### 11.4. Ngăn chặn các đối tượng không mong muốn bằng Negative prompt

Một cách khác để loại bỏ các đối tượng hoặc kiểu mà bạn không muốn là sử dụng các Negative prompt. Trên thực tế, nhiều học viên sử dụng những gợi ý tiêu cực được soạn sẵn. Dưới đây là một ví dụ cho ảnh chân dung.

((((ugly))), ((duplicate))), ((morbid)), ((mutilated)), [out of frame], extra fingers, mutated hands, ((poorly drawn hands)), ((poorly drawn face)), (((mutation))), ((deformed))), ((ugly)), blurry, ((bad anatomy)), (((bad proportions))), ((extra limbs)),

*cloned face, (((disfigured))), out of frame, ugly, extra limbs, (bad anatomy), gross proportions, (malformed limbs), ((missing arms)), ((missing legs)), (((extra arms))), (((extra legs))), mutated hands, (fused fingers), (too many fingers), (((long neck)))*

*((((xáu xí))), (((trùng lắp))), ((bệnh tật)), ((bị cắt xén)), [ngoài khung], ngón tay thừa, bàn tay bị đột biến, ((bàn tay được vẽ kẽm)), ((khuôn mặt được vẽ kẽm)), (((đột biến))), (((biến dạng))), ((xáu)), mờ, ((giải phẫu xáu)), (((tỷ lệ xáu))), ((chân tay thừa)), khuôn mặt nhân bản, ((bị biến dạng))), ngoài khung hình, xáu xí, chân tay thừa, (giải phẫu xáu), tỷ lệ thô, (tay chân dị dạng), ((thiếu tay)), ((thiếu chân)), (((thêm cánh tay))), (((thêm chân))), bàn tay bị đột biến, (ngón tay hợp nhất), (quá nhiều ngón tay), (((cổ dài)))*

Tất cả những gì bạn cần làm là liệt kê *mọi thứ* bạn không muốn nhìn thấy.

### 11.5. Trộn hai keyword

Bạn muốn trộn hai khuôn mặt? Trao đổi keyword được thiết kế cho điều đó. Trên thực tế, đây là một kỹ thuật quan trọng để tạo ra diện mạo mới. Bằng cách sử dụng tên của các diễn viên, bạn có thể dễ dàng tạo ra những khuôn mặt ưa nhìn. Nhưng nhược điểm là khuôn mặt quá dễ nhận biết. Mọi người đều có thể biết họ là ai và bạn mất đi sự mới lạ. Hoán đổi keyword cho phép bạn kết hợp hai khuôn mặt dễ nhận biết lại với nhau để tạo thành một khuôn mặt mới.

Sử dụng cú pháp

*[person1 : person2: amount]*

*[người1 : người2: số tiền]*

để hòa trộn hai khuôn mặt ở mức độ khác nhau. Dưới đây là một ví dụ



Số cuối cùng nằm trong khoảng từ 0 đến 1. Ví dụ: nếu sampling steps là 40, specifying amount = 0,75 sẽ hoán đổi keyword joe biden thành donald trump ở bước  $40 * 0,75 = 30$  steps.

Kỹ thuật này có thể được sử dụng để tạo ra các khuôn mặt mới. Dưới đây là một ví dụ.

*[Emma Watson: Amber heard: 0.85], 40 steps*

[Emma Watson: Amber nghe được: 0,85], 40 bước



Prompt đầy đủ là

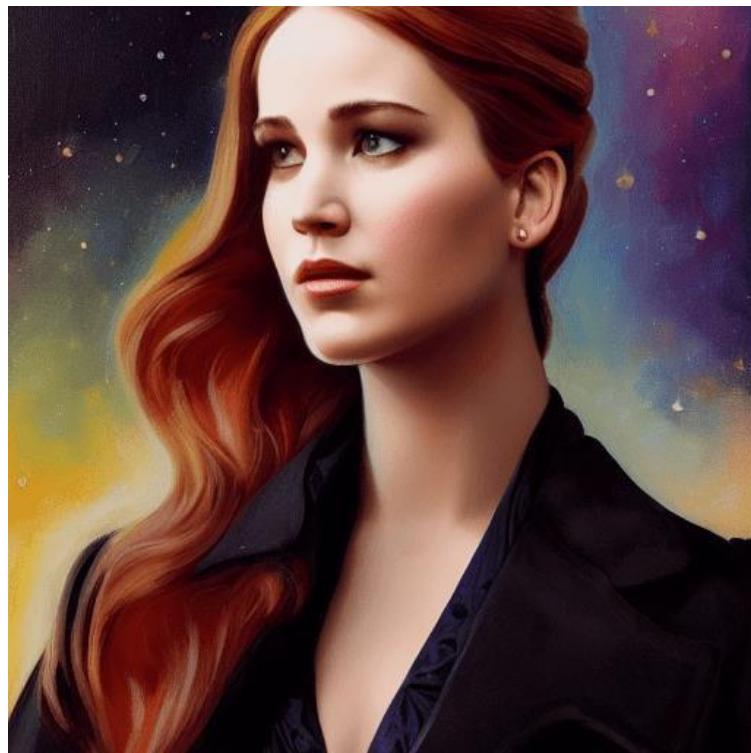
[Emma Watson: Amber heard: 0.85], ((Victorian)), Feminine, ((Perfect Face)), ((arms outstretched above head)), ((Aype Beven)), ((scott williams)) ((jim lee)), ((Leinil Francis Yu)), ((Salva Espin)), ((oil painting)), ((Matteo Lolli)), ((Sophie Anderson)), ((Kris Anka)), (Intricate), (High Detail), (bokeh)

[Emma Watson: Amber đã nghe: 0,85], ((Victoria)), Nữ tính, ((Khuôn mặt hoàn hảo)), ((tay dang rộng trên đầu)), ((Aype Beven)), ((scott williams)) ((jim lee)), ((Leinil Francis Yu)), ((Salva Espin)), ((son đầu)), ((Matteo Lolli)), ((Sophie Anderson)), ((Kris Anka)), (Phúc tạp), (Chi tiết cao), (bokeh)

Bạn có thể tiến thêm một bước nữa là ghép 4 tên lại với nhau:

[Evan Rachel Wood: Jennifer Lawrence: 0.75], [Jennifer Aniston: Jennifer Connelly: 0.85], 40 steps

[Evan Rachel Wood: Jennifer Lawrence: 0,75], [Jennifer Aniston: Jennifer Connelly: 0,85], 40 bước



Prompt đầy đủ:

[Evan Rachel Wood: Jennifer Lawrence: 0.75], [Jennifer Aniston: Jennifer Connelly: 0.85], Feminine,((Perfect Face)), ((arms outstretched above head)), ((Aype Beven)), ((scott williams)) ((jim lee)),((Leinil Francis Yu)), ((Salva Espin)),  
((oil painting)), ((Matteo Lolli)), ((Sophie Anderson)), ((Kris Anka)), (Intricate),(High Detail), (bokeh)

[Evan Rachel Wood: Jennifer Lawrence: 0,75], [Jennifer Aniston: Jennifer Connelly: 0,85], Nữ tính,((Khuôn mặt hoàn hảo)), ((tay dang rộng trên đầu)), ((Aype Beven)), ((scott williams )) ((jim lee)),((Leinil Francis Yu)), ((Salva Espin)),  
((sơn dầu)), ((Matteo Lolli)), ((Sophie Anderson)), ((Kris Anka)), (Phức tạp),(Chi tiết cao), (bokeh)

Kỹ thuật này cho phép bạn tạo ra những khuôn mặt mới lạ với các đặc điểm khuôn mặt chính xác mà bạn muốn. Chỉ cần chọn những người nổi tiếng có ngoại hình nhất định và kết hợp chúng lại với nhau bằng cách hoán đổi keyword!

## Chương 12. Công cụ nâng cấp hình ảnh AI

Công cụ nâng cấp hình ảnh AI như ESRGAN là công cụ không thể thiếu để cải thiện chất lượng hình ảnh được tạo ra bởi Stable Diffusion. Trên thực tế, nó được sử dụng phổ biến đến mức nhiều GUI Stable diffusion đã tích hợp hỗ trợ cho nó. Có nhiều dịch vụ và phần mềm trả phí hứa hẹn sẽ phóng to hình ảnh của bạn đồng thời làm cho nó sắc nét hơn. Bạn không cần phải trả tiền cho những dịch vụ đó. Trong Chương này, chúng ta sẽ tìm hiểu

về công cụ nâng cấp hình ảnh là gì, cách chúng hoạt động và cách sử dụng chúng miễn phí.

### **12.1. Tại sao chúng ta cần một công cụ nâng cấp hình ảnh?**

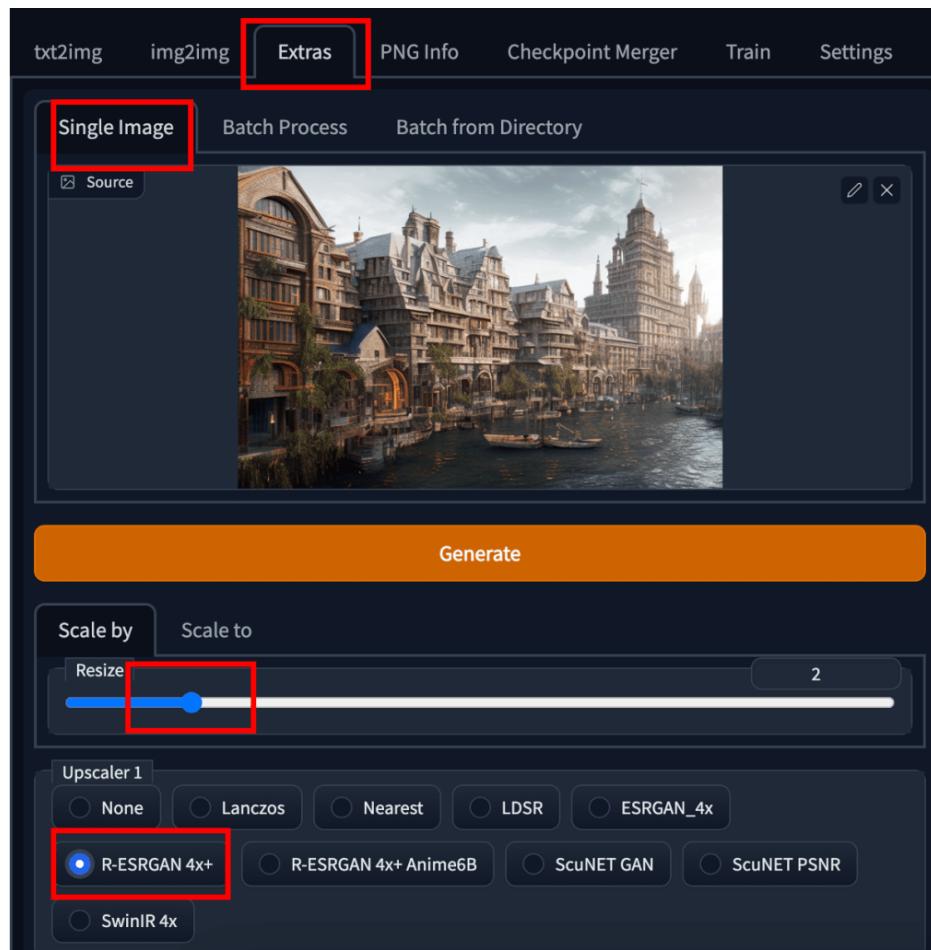
Kích thước hình ảnh mặc định của Stable Diffusion v1 là  $512 \times 512$  pixel. Đây là mức thấp trong tiêu chuẩn ngày nay. Hãy lấy iPhone 12 làm ví dụ. Máy ảnh của nó tạo ra hình ảnh 12MP – tức là  $4.032 \times 3.024$  pixel. Màn hình của nó hiển thị  $2.532 \times 1.170$  pixel nên hình ảnh Stable diffusion không được chia tỷ lệ sẽ cần được phóng to và trông có chất lượng thấp.

Một cảnh phức tạp được tạo ra bởi Stable Diffusion thường không sắc nét như mong đợi. Nó thường gấp khó khăn với các chi tiết tốt.



### **12.2. Làm cách nào để sử dụng Công cụ nâng cấp hình ảnh?**

Chúng ta sẽ tìm hiểu cách sử dụng trình nâng cấp AI bằng GUI AUTOMATIC1111. Đến tab **Extras**, hãy chọn thẻ **Single Image**.



Tải hình ảnh bạn muốn nâng cấp lên **source**. Phần **Resize**, nhiều công cụ nâng cấp ảnh bằng AI được mặc định nâng cấp 4 lần nên 4 là một lựa chọn tốt. Đặt giá trị **Resize** thấp hơn như 2 nếu bạn không muốn hình ảnh quá lớn.

Nếu hình ảnh của bạn là  $512 \times 512$  pixel thì thay đổi kích thước 2x là  $1024 \times 1024$  pixel và 4x là  $2048 \times 2048$  pixel. Chọn R-ESRGAN 4x+, công cụ nâng cấp hình ảnh AI hoạt động với hầu hết hình ảnh. Nhấn Tạo để bắt đầu nâng cấp. Khi hoàn tất, hình ảnh được nâng cấp sẽ xuất hiện trong cửa sổ đầu ra bên phải. Nhấp chuột phải vào hình ảnh để lưu.

### 12.3. Các tùy chọn mô hình nâng cấp hình ảnh

#### 12.3.1. LDSR

**Latent Diffusion Super Resolution** (LDSR) công cụ nâng cấp ban đầu được phát hành cùng với Stable Diffusion 1.4. Đó là một mô hình được đào tạo để thực hiện các nhiệm vụ nâng cấp. Mặc dù mang lại chất lượng vượt trội nhưng nó cực kỳ chậm.

#### 12.3.2. ESRGAN 4x

**Enhanced Super-Resolution Generative Adversarial Networks** (ESRGAN) là một cải tiến cho mô hình SRGAN trước đó. Nó có xu hướng giữ lại các chi tiết tốt và tạo ra hình ảnh sắc nét.

### 12.3.3. R-ESRGAN 4x

Các **ESRGAN thực sự** (R-ESRGAN) là một cải tiến của ESRGAN và được thiết kế để khôi phục nhiều loại hình ảnh *thế giới thực*. Nó mô hình hóa các mức độ biến dạng khác nhau từ ống kính máy ảnh và nén kỹ thuật số. So với ESRGAN, nó có xu hướng tạo ra hình ảnh mượt mà hơn. R-ESRGAN hoạt động tốt nhất với hình ảnh chân thực.

### 12.4. Những mô hình nâng cấp khác

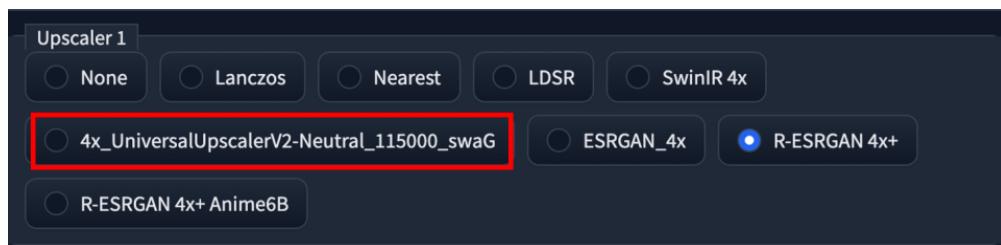
Hình ảnh anime yêu cầu những người nâng cấp được đào tạo đặc biệt để khôi phục chúng. Hãy tìm từ *phim hoạt hình* nhân danh người cao cấp hơn. Có một kho lưu trữ công cộng để tìm kiếm và tải xuống các bản nâng cấp mới. Truy cập cơ sở dữ liệu mô hình Upscaler (URL bên dưới) để tải xuống các trình nâng cấp khác.

[https://upscale.wiki/wiki/Model\\_Database](https://upscale.wiki/wiki/Model_Database)

### 12.5. Cài đặt bộ nâng cấp mới

Để cài đặt trình nâng cấp hình ảnh trong GUI AUTOMATIC1111, hãy tải xuống mô hình từ cơ sở dữ liệu mô hình nâng cấp và đặt nó vào thư mục  
stable-diffusion-webui/models/ESRGAN

Khởi động lại GUI. Trình nâng cấp của bạn bây giờ sẽ có sẵn để lựa chọn. Dưới đây là những gì bạn sẽ thấy sau khi cài đặt Universal Upscaler V2.



## Chương 13. Quy trình làm việc hoàn chỉnh từ đầu đến cuối

Một tác phẩm nghệ thuật Stable diffusion tuyệt đẹp không được tạo ra bởi một Prompt đơn giản. Quy trình làm việc là một quá trình gồm nhiều bước. Trong chương này, tôi sẽ thực hiện từng bước một quy trình làm việc hoàn chỉnh.

Các bước trong quy trình công việc này là:

1. Xây dựng một prompt cơ sở.
2. Chọn một mô hình.
3. Hiệu chỉnh Prompt để tạo ra hình ảnh có bố cục tốt.
4. Sửa lỗi bằng inpainting.
5. Nâng chất lượng hình ảnh bằng Upscale.
6. Điều chỉnh lần cuối bằng phần mềm chỉnh sửa ảnh.

### 13.1. Phần mềm

Tôi sẽ sử dụng GUI AUTOMATIC1111 trong quy trình làm việc này vì tôi muốn khai thác một số kỹ thuật nhắc nhở nâng cao và thử nghiệm nhiều mô hình.

Trong giai đoạn xử lý hậu kỳ, tôi sẽ sử dụng GIMP, một trình chỉnh sửa ảnh nguồn mở và miễn phí, để thực hiện những điều chỉnh nhỏ cuối cùng.

### 13.2. Xây dựng Prompt cơ bản

Nhiệm vụ đầu tiên là tìm kiếm Prompt gần giống với những gì bạn muốn. Điều này bao gồm hai khía cạnh: (1) **Chủ đề (Subject)** và 2) **Phong cách**. Bạn nên mô tả **chủ đề** càng chi tiết càng tốt. Tiếp theo, bao gồm nhiều keyword có ảnh hưởng đến **phong cách và thẩm mỹ**.

Giả sử tôi muốn tạo một bức hình vẽ một người phụ nữ. Sử dụng Prompt sau đây:

*A digital painting of [blake lively:Ana de Armas:0.8], full-body, contemporary white top, red dress, by Artgerm, Guangjian, artstation, soft eyes, extremely detailed face, stunningly beautiful, highly detailed, sharp focus, radiant light rays*

*Một bức tranh kỹ thuật số của [blake sống động:Ana de Armas:0.8], toàn thân, áo trắng đương đại, váy đỏ, của Artgerm, Quảng Kiện, artstation, đôi mắt dịu dàng, khuôn mặt cực kỳ chi tiết, đẹp tuyệt vời, độ chi tiết cao, tiêu điểm sắc nét, rạng rỡ các tia sáng*

Đưa ra một Negative prompt tiêu chuẩn chỉ trong trường hợp...

*ugly, disfigured, deformed, cropped*

*xấu xí, biến dạng, biến dạng, bị cắt xén*

tôi đã sử dụng *trộn keyword* diễn tả ở Chương 11 để kiểm soát mô hình của tôi trông như thế nào. Đây chỉ là Prompt nháp. Chúng tôi sẽ thực hiện các thay đổi đối với nó trong quy trình làm việc. Hãy chuyển sang bước tiếp theo.

### 13.3. Chọn một mô hình

Kiểm tra Prompt cơ sở cũng là thời điểm tốt để chọn mô hình. Đối với ảnh chân dung kỹ thuật số, tôi sẽ thử nghiệm ba mẫu sau:

- Stable diffusion 1.5: Mô hình cơ sở
- F222: Chuyên dành cho nữ (Chú ý: đây là model NSFW)
- OpenJourney: Phong cách MidJourney v4

Dưới đây là các mẫu từ mỗi mô hình. Chúng tôi sẽ chọn một mô hình dựa trên phong cách. Đừng lo lắng về việc có được một hình ảnh hoàn hảo. Những khiếm khuyết có thể

được sửa chữa sau này. Tôi sẽ sử dụng kích thước chân dung  $512 \times 832$  pixel để tăng cơ hội tạo ra một bức chân dung toàn thân. Các thông số còn lại là tiêu chuẩn: 25 sampling steps của Euler, CFG Scale = 7.

**Stable diffusion v1.5** tạo ra sự kết hợp giữa phong cách kỹ thuật số và nhiếp ảnh. Một số thì đẹp nhưng nhiều trong số chúng có giải phẫu xấu nên sẽ khó sửa.



*Stable diffusion v1.5*

Đó là một sự đặt cược an toàn để sử dụng **F222** để tạo ra hình ảnh theo phong cách chân dung. Được tinh chỉnh với số lượng lớn hình ảnh phụ nữ, các bộ phận cơ thể thường được tạo ra chính xác. Nó tạo ra một phong cách nghệ thuật kỹ thuật số khá chuẩn với Prompt này.



*F222*

OpenJourney được đào tạo bằng các hình ảnh được tạo bởi MidJourney v4, một dịch vụ tạo hình ảnh AI trả phí. Những hình ảnh có một cảm giác và cái nhìn khác biệt.



*OpenJourney*

Tôi thích cả phong cách của F222 và OpenJourney. Nhưng **hãy chọn OpenJourney để tiếp tục quy trình làm việc.**

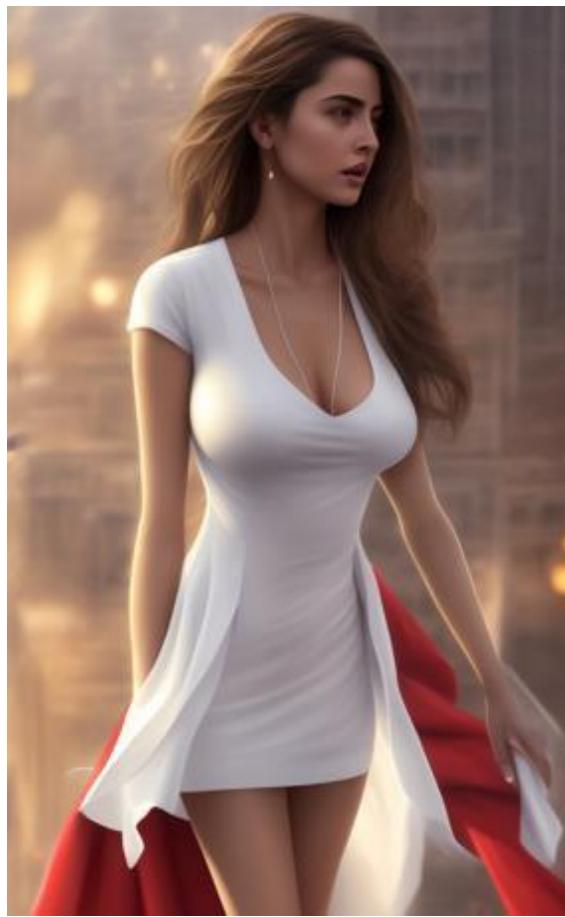
#### 13.4. Tinh chỉnh Prompt và có được bối cảnh tốt

Bây giờ chúng ta đã chọn được một mô hình, hãy tinh chỉnh Prompt bằng cách thêm hoặc xóa keyword cho đến khi bạn có được hình ảnh có thể sử dụng cho giai đoạn tiếp theo.

Khi thử nghiệm các Prompt mới, nên tạo ra ít nhất **4 hình ảnh cùng một lúc**.

Mục tiêu ở giai đoạn này không phải là chọn một hình ảnh hoàn hảo mà là chọn một hình ảnh có **thành phần tốt**. Những lỗi nhỏ có thể được sửa chữa sau này bằng cách sử dụng inpainting. Nhưng đừng chọn một cái có quá nhiều khuyết điểm. Bạn sẽ tốn rất nhiều thời gian để chỉnh sửa nó.

Tôi giữ batch size là 4 và Seed là -1 (ngẫu nhiên). Tôi tiếp tục sửa đổi Prompt, tạo ra một loạt 4 hình ảnh trong khi vẫn chú ý đến bối cảnh. Hãy sử dụng hình ảnh sau đây cho phần còn lại của hướng dẫn.



Và đây là Prompt cuối cùng:

*A digital painting of [blake lively:Ana de Armas:0.8] in street cityscape background, happy, full-body, contemporary white top, red dress, (stocking:1.2), by Artgerm, Guangjian, artstation, soft eyes, extremely detailed face, stunningly beautiful, highly detailed, sharp focus, radiant light rays, cinematic lighting, colorful, volumetric light*

*Một bức tranh kỹ thuật số của [blake sống động:Ana de Armas:0.8] trên nền cảnh quan thành phố đường phố, hạnh phúc, toàn thân, áo trắng đương đại, váy đỏ, (stocking:1.2), của Artgerm, Quảng Kiện, artstation, đôi mắt dịu dàng, cực kỳ chi tiết khuôn mặt, đẹp tuyệt vời, độ chi tiết cao, tiêu điểm sắc nét, tia sáng rạng rỡ, ánh sáng điện ảnh, nhiều màu sắc, ánh sáng thể tích*

Negative prompt vẫn như cũ:

*ugly, disfigured, deformed, cropped*

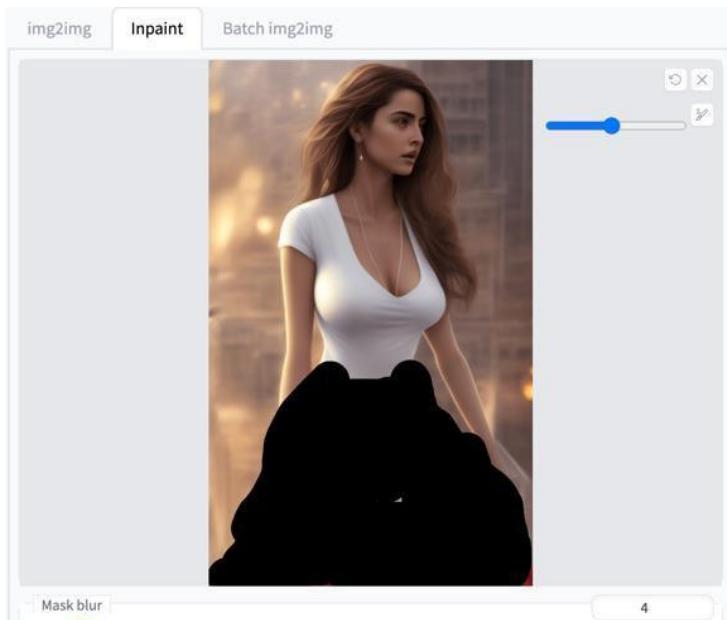
*xấu xí, biến dạng, biến dạng, bị cắt xén*

### **13.5. Sửa lỗi bằng Inpainting**

Nửa trên của hình ảnh này có hình thức tốt, nhưng nửa dưới không hợp lý lắm. Hãy sử

dụng inpainting để khắc phục nó. Nếu bạn vừa tạo hình ảnh này trong **txt2img** tab với AUTOMATIC1111, sử dụng **Send to Inpaint** để gửi hình ảnh và thông số tới **Inpaint**.

Nếu bạn đã lưu hình ảnh vào thư mục trên máy tính, hãy truy cập **PNG info** tab, kéo và thả hình ảnh vào khung hình ảnh. Các tham số tạo ảnh sẽ được điền vào ô text. Nhấn **Send to inpaint**. Chuyển đến **img2img** tab. Chọn tab phụ **Inpaint**. Sử dụng công cụ cọ vẽ để tô vùng bạn muốn tái tạo.

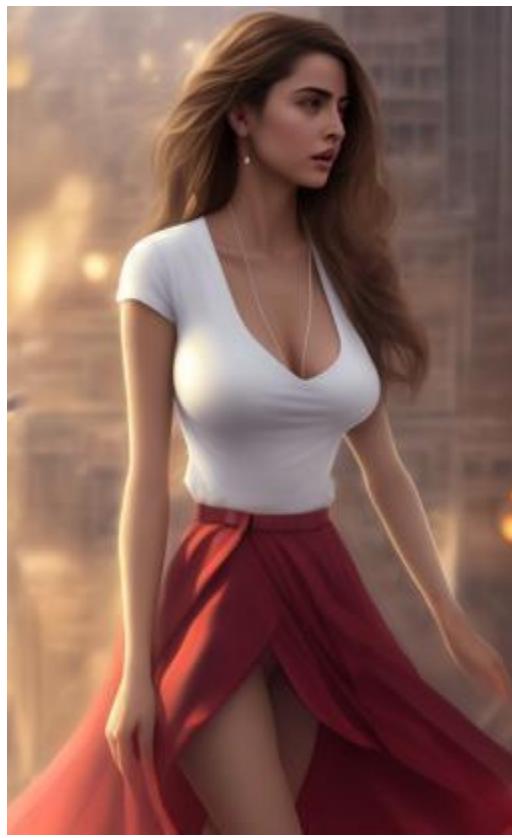


Chỉnh **mask content = original** và **Denoising strength = 0,66**. Tôi muốn một chiếc váy màu đỏ nên tôi xóa keyword “màu trắng” và tăng điểm nhấn của chiếc váy đỏ lên 1,3 trong Prompt:

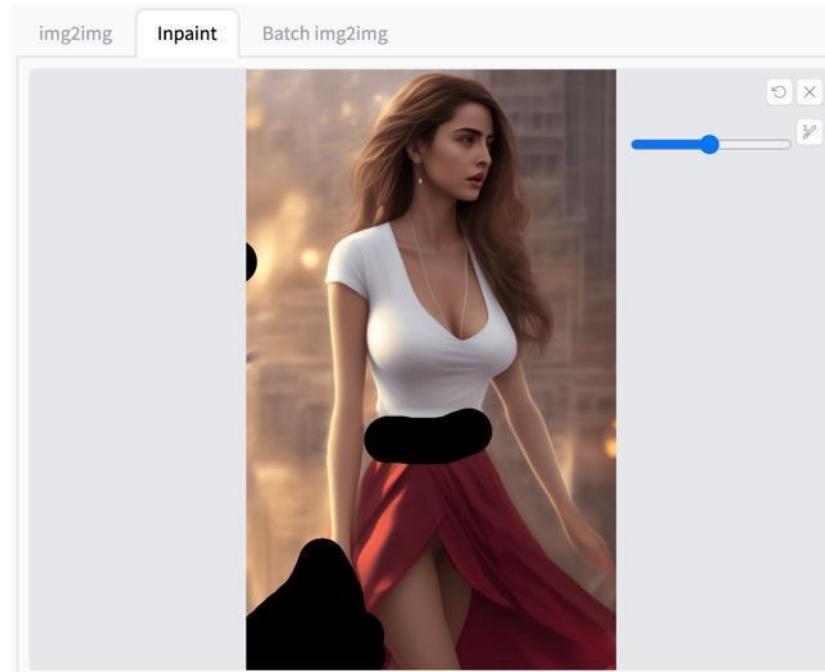
*A digital painting of [blade lively:Ana de Armas:0.8] in street cityscape background, happy, full-body, contemporary top, (red dress: 1.3), (stocking:1.2), by Artgerm, Guangjian, artstation, soft eyes, extremely detailed face, stunningly beautiful, highly detailed, sharp focus, radiant light rays, cinematic lighting, colorful, volumetric light*

*Một bức tranh kỹ thuật số về [blade sống động:Ana de Armas:0.8] trong nền cảnh quan thành phố đường phố, hạnh phúc, toàn thân, áo đương đại, (váy đỏ: 1.3), (stocking:1.2), bởi Artgerm, Quảng Kiện, artstation, đôi mắt mềm mại, khuôn mặt cực kỳ chi tiết, đẹp tuyệt vời, độ chi tiết cao, tiêu cự sắc nét, tia sáng rạng rỡ, ánh sáng điện ảnh, nhiều màu sắc, ánh sáng thể tích*

Đây là kết quả inpainting mà tôi đã chọn:



Hãy sửa lại thắt lưng và váy thêm một chút nhé. Prompt có thể giữ nguyên. Cường độ khử nhiễu có thể được điều chỉnh tùy thuộc vào mức độ thay đổi mà bạn muốn. Đối với phần inpainting này, nó vẫn giữ nguyên ở mức 0,66.

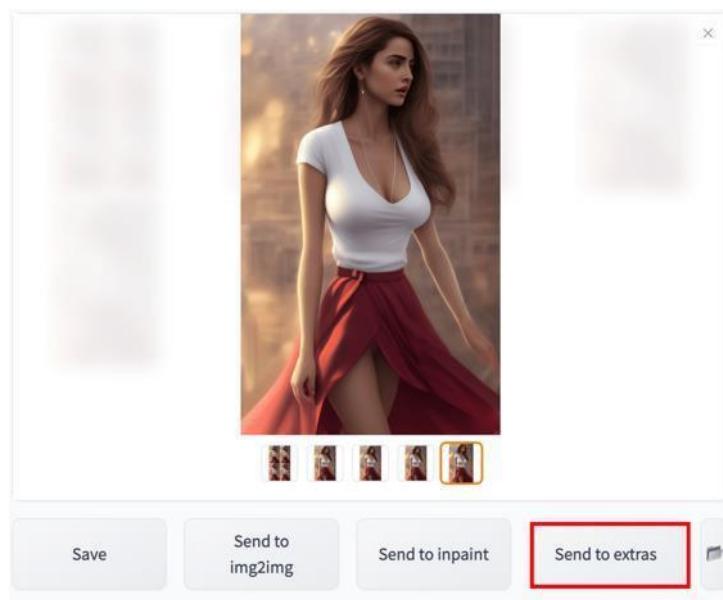


Bây giờ chúng ta nhận được:



Duy trì **mask content = original** nhưng điều chỉnh **denoising strength** tăng lên và giảm xuống để có kết quả tốt nhất.

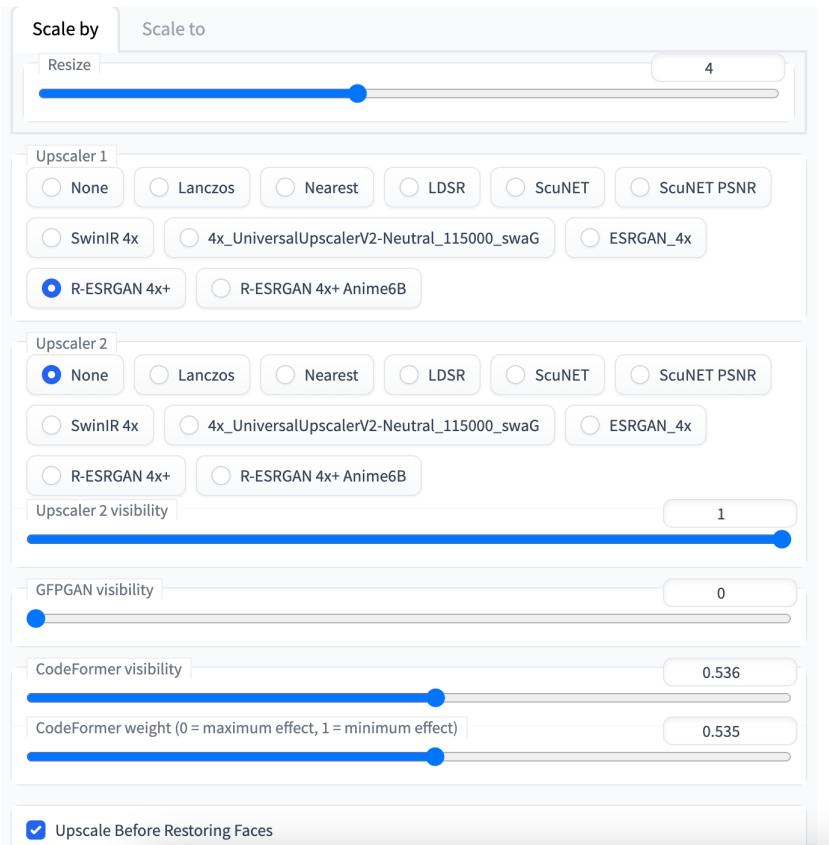
Cuối cùng, chỉnh sửa nền để loại bỏ mọi chi tiết không hợp lý.



Nhấn vào **Send to Extras** để nâng cấp chất lượng ảnh. Bạn có thể sử dụng thông số sau:

- Resize: 4
- Upscaler: R-ESRGAN 4x+
- CodeFormer visibility: ~0,5

- CodeFormer weight: ~0,5
- Upscale before restoring faces: Có



### 13.6. Điều chỉnh cuối cùng

Bước cuối cùng là thực hiện một số **điều chỉnh nhỏ về độ sáng và độ tương phản** sử dụng [GIMP](#), hoặc bất kỳ phần mềm chỉnh sửa ảnh nào bạn chọn (Photopea, Photoshop...). Bạn có thể làm nhiều việc khác để chỉnh sửa hình ảnh. Tôi sẽ chỉ đề cập đến những điều cơ bản nhất như: **levels, curves and crop**.

## Chương 14. Sử dụng VAE để cải thiện đôi mắt và khuôn mặt

### 14.1. VAE là gì?

VAE là viết tắt của **bộ mã hóa tự động biến thiên**. Nó là một phần của mô hình mang nơ ron, nó mã hóa và giải mã các hình ảnh đến và đi từ latent space nhỏ hơn, để việc tính toán có thể nhanh hơn.

Sau khi chạy n steps thì quá trình lấy mẫu hoàn tất, khi này chúng ta nhận được hình ảnh dạng Latent. Để hoàn thiện bức ảnh trở về định dạng bình thường với đầy đủ thông tin, một quá trình decoder được thực hiện. Quá trình này yêu cầu một loại model thành phần để có thể hoạt động, chúng ta gọi nó là VAE. Trong trường hợp tác vụ img-img, VAE sẽ làm nhiệm vụ encoder ảnh đầu vào từ định dạng thông thường sang định dạng latent và đưa vào quá trình UNET. Hiện nay đã có các checkpoint đã được nhúng VAE sẵn, tuy nhiên trường hợp thiếu VAE, phổ biến là ở các phiên bản v1.5, chúng ta có thể tải thêm VAE.

Bạn không cần cài đặt tệp VAE để chạy Stable diffusion. Bất kỳ kiểu máy nào bạn sử dụng, dù là v1, v2 hay tùy chỉnh, đều đã có VAE mặc định. Khi mọi người nói tải xuống và sử dụng VAE, đó là sử dụng **Phiên bản cải tiến** của nó.

#### 14.2. Sử dụng VAE có tác dụng gì?

VAE sẽ giải mã hình ảnh tốt hơn, các chi tiết được phục hồi tốt hơn. Ngược lại, sử dụng không đúng VAE sẽ làm cho hình ảnh bị vỡ, hoặc xấu.

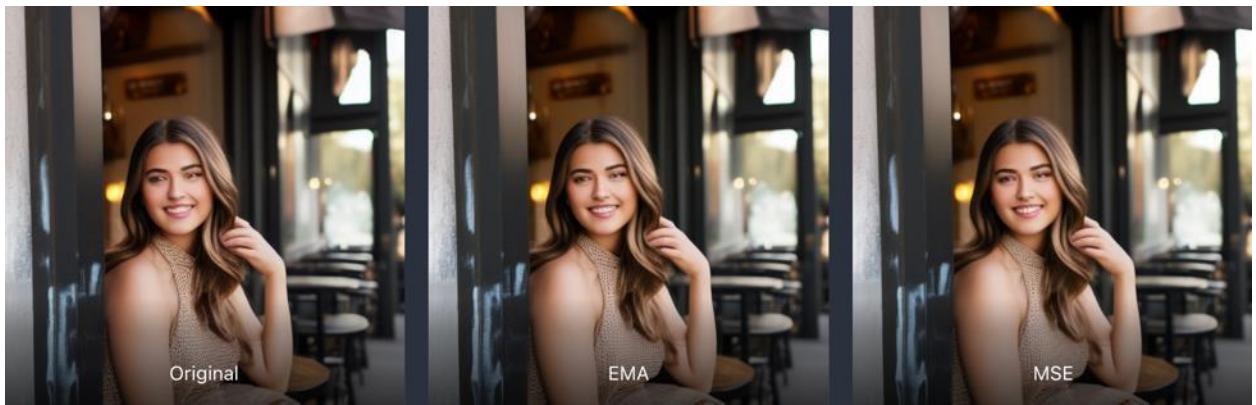
Stability AI đã phát hành hai biến thể của bộ giải mã VAE được tinh chỉnh là EMA và MSE. Hãy xem phần so sánh của chúng được mô tả bên dưới.



So sánh độ ổn định của AI giữa EMA, MSE và bộ giải mã gốc. (hình ảnh  $256 \times 256$ )

Đánh giá về độ ổn định với hình ảnh  $256 \times 256$  là EMA cho hình ảnh sắc nét hơn khi hình ảnh của MSE mượt mà hơn.

Trong mọi trường hợp, VAE mới hoạt động kém hơn. Dưới đây là so sánh giữa bản gốc, EMA và MSE sử dụng mô hình Stable Diffusion v1.5. (Có thể tìm thấy Prompt ở đây)



So sánh VAE giữa bản gốc, EMA và MSE bằng SD v1.5. Lưu ý đôi mắt bị cắt xén trong ảnh gốc đã được phục hồi.

Các cải tiến đói với việc tạo văn bản không rõ ràng (Đã thêm “holding a sign said Stable Diffusion” vào Prompt).



So sánh VAE giữa bản gốc, EMA và MSE. Cải thiện hiển thị văn bản là không rõ ràng.

Bạn cũng có thể sử dụng các VAE này với một mô hình tùy chỉnh. Thử nghiệm với một số mẫu anime nhưng không thấy cải thiện gì cả. Lưu ý cuối cùng, EMA và MSE tương thích với Stable Diffusion v2.0. Bạn có thể sử dụng chúng, nhưng hiệu quả là không nhiều phiên bản 2.0 đã có khả năng hiển thị mắt rất tốt.

### 14.3. Bạn có nên sử dụng VAE?

Bạn không cần sử dụng VAE nếu bạn hài lòng với kết quả mình nhận được. Ví dụ: bạn đang sử dụng tính năng phục hồi khuôn mặt như CodeFormer để sửa mắt. Bạn nên sử dụng VAE nếu bạn đang cần thực hiện tất cả những cải tiến nhỏ mà bạn có thể nhận được.

### 14.4. Cài đặt VAE

Hiện tại có hai phiên bản cải tiến của VAE do Stability phát hành. Dưới đây là các liên kết tải xuống trực tiếp.

EMA: <https://huggingface.co/stabilityai/sd-vae-ft-ema-origen/resolve/main/vae-ft-ema-560000-ema-pruned.ckpt>

MSE: <https://huggingface.co/stabilityai/sd-vae-ft-mse-origen/resolve/main/vae-ft-mse-840000-ema-pruned.ckpt>

Để cài đặt trên GUI AUTOMATIC1111, hãy lưu các tệp VAE đã tải xuống vào thư mục stable-diffusion-webui/models/VAE

## Chương 15. Embeddings

### 15.1. Embeddings là gì?

Embedding là kết quả của phương pháp tinh chỉnh tài liệu gọi là textual inversion. Giống như hypernetwork, textual inversion không thay đổi mô hình. Nó chỉ xác định các từ khóa mới để đạt được các phong cách nhất định. Phương pháp này đã thu hút được sự chú ý vì nó có khả năng đưa các kiểu hoặc đối tượng mới vào một mô hình chỉ với 3 - 5 hình ảnh mẫu.

### 15.2. Ví dụ về nhúng

Nhúng có thể được sử dụng cho các đối tượng mới. Dưới đây là một ví dụ về việc tiêm thuốc cho một con mèo đồ chơi. Lưu ý rằng khái niệm mới (mèo đồ chơi) có thể được sử dụng cùng với các khái niệm hiện có khác (thuyền, ba lô, v.v.) trong mô hình.



Ví dụ về nhúng một đối tượng

Nhúng cũng có thể là một phong cách mới. Ví dụ dưới đây cho thấy việc nhúng một kiểu mới và chuyển kiểu đó sang ngữ cảnh khác.



Ví dụ về việc nhúng một style.

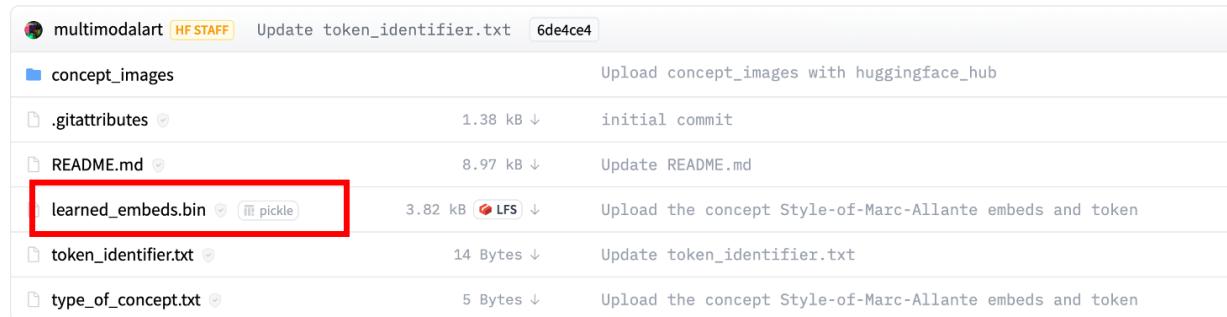
### 15.3. Tìm Embeddings ở đâu?

Hugging Face *Stable Diffusion Concept Library*, là kho lưu trữ số lượng lớn các phần nhúng tùy chỉnh. Civitai là một trang web tuyệt vời khác mà bạn có thể tải xuống các mô hình. Đảm bảo lọc bằng "đảo ngược văn bản" để chỉ duyệt các phần nhúng.

Thư viện khái niệm Stable diffusion: <https://huggingface.co/sd-concepts-library> Civtai: <https://civitai.com/>

#### 15.4. Sử dụng phần nhúng với AUTOMATIC1111

Sử dụng tính năng nhúng trong AUTOMATIC1111 rất dễ dàng. Đầu tiên, tải xuống tệp nhúng từ <https://huggingface.co/sd-concepts-library>. Tìm tập tin có tên *learned\_embedds.bin*, nhập vào tên tệp và nhấp vào nút tải xuống ở trang tiếp theo.



Kết tiếp, **đổi tên tệp thành keyword bạn muốn sử dụng**, đổi tên thành từ khóa mà bạn muốn sử dụng ví dụ: `marc_allante.bin`

Đặt nó vào thư mục nhúng trong thư mục làm việc của GUI:

`stable-diffusion-webui/embeddings`

Khởi động lại GUI. Trong khi khởi động, bạn sẽ thấy một thông báo như:

Loaded a total of 1 textual inversion embeddings. Embeddings: `marc_allante`

Sử dụng tên tệp như một phần của Prompt để

Ví dụ: Prompt sau sẽ hoạt động trên AUTOMATIC1111.

`(marc_allante:1.2) a dog`

Chúng ta có được hình ảnh với phong cách như mong đợi.



### 15.5. Lưu ý khi sử dụng phần nhúng trong AUTOMATIC1111

Nếu bạn chú ý đến Prompt, bạn sẽ nhận thấy tôi đã tăng độ mạnh của keyword kích hoạt `marc_allante`. Tôi thấy cần phải điều chỉnh độ mạnh của keyword. Có thể bạn sẽ phải thử **nhiều lần keyword strength** để có được hiệu quả như mong muốn. Dưới đây là một ví dụ về việc thay đổi cường độ trong khi vẫn giữ nguyên Seed và mọi thứ khác.



(`marc_allante: 1.0`)

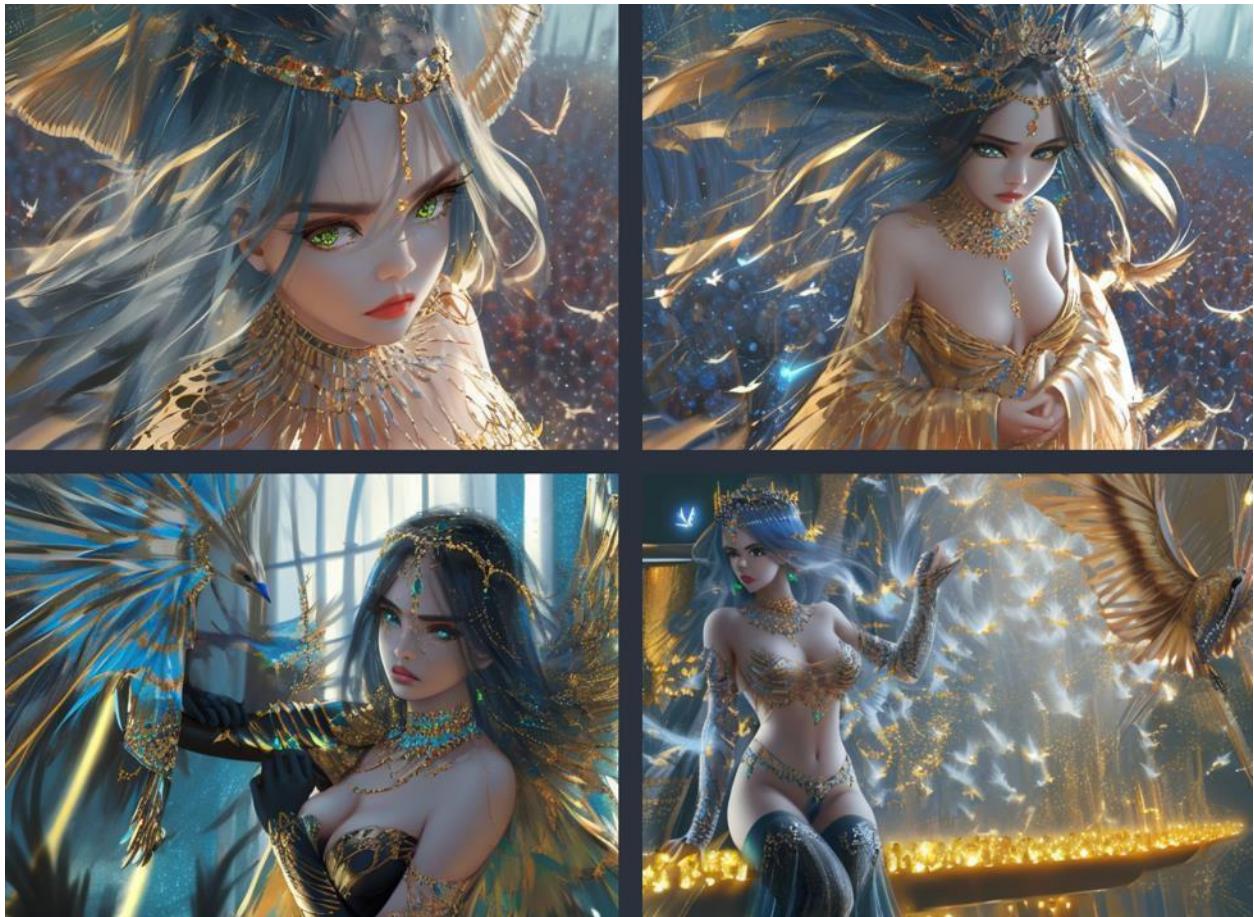
(`marc_allante: 1.2`)

(`marc_allante: 1.4`)

### 15.6. Một số nhúng hay sử dụng

Có rất nhiều phần nhúng có sẵn có thể thử, các bạn có thể tham khảo các tệp nhúng dưới đây.

### 15.6.1. Wlop\_style



Nếu bạn đã làm việc với các mô hình cơ sở Stable diffusion, bạn sẽ thấy không thể tạo ra phong cách **Wlop\_style** (một họa sĩ minh họa trên trang web artstation) cho dù bạn có cố gắng thế nào đi chăng nữa. Việc nhúng cùng với một mô hình tùy chỉnh cuối cùng cũng có thể thực hiện được điều này.

Các *wlop\_style* việc nhúng có thể thể hiện một số phong cách minh họa đẹp mắt của nghệ sĩ. Nó nên được sử dụng với **SirVeggie's wlop-any** mô hình tùy chỉnh. (Xem Chương 6 để cài đặt các mô hình tùy chỉnh.)

Liên kết tải xuống:

Nhúng: [https://huggingface.co/datasets/Nerfgun3/wlop\\_style/resolve/main/wlop\\_style.pt](https://huggingface.co/datasets/Nerfgun3/wlop_style/resolve/main/wlop_style.pt)

Mô hình Wlop-bất kỳ: <https://huggingface.co/SirVeggie/wlop/resolve/main/wlop-any.ckpt>

Prompt hoạt động cho AUTOMATIC1111 là

*(wlop\_style :0.6) (m\_wlop:1.4) woman wearing dress, perfect face, beautiful detailed eyes, long hair, birds*

*(wlop\_style :0.6) (m\_wlop:1.4) người phụ nữ mặc váy, khuôn mặt hoàn hảo, đôi mắt đẹp*

*chi tiết, mái tóc dài, hình chim*

Negative prompt:

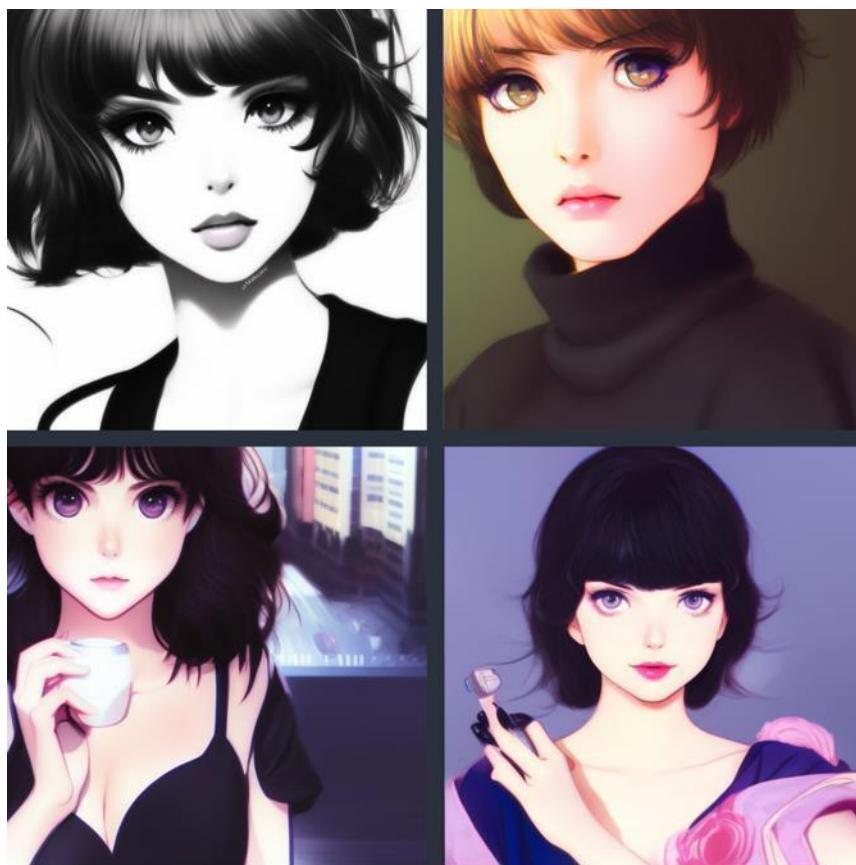
*closed eyes, disfigured, deformed*

*nhắm mắt, biến dạng, biến dạng*

*wlop\_style* là keyword để nhúng, *m\_wlop* là keyword cho mô hình.

Đừng nản lòng nếu bạn không có phong cách. Hãy thử thay đổi độ mạnh của hai keyword. Một số đối tượng có thể không hoạt động với tính năng nhúng. Hãy thử một số đồ vật phổ biến trong tác phẩm nghệ thuật của wlop.

### 15.6.2. Kuvshinov



Kuvshinov là một minh họa của Nga. Bạn có thể dùng *kuvshinov* nhúng với Stable diffusion v1.4.

Liên kết tải xuống: [https://huggingface.co/sd-concepts-library/kuvshinov/resolve/main/learned\\_embeddings.bin](https://huggingface.co/sd-concepts-library/kuvshinov/resolve/main/learned_embeddings.bin)

Prompt:

*(\_kuvshinov:1), a woman with beautiful detailed eyes, highlight hair*

(*\_kuvshinov:1*), một người phụ nữ có đôi mắt đẹp chi tiết, mái tóc highlight

Negative prompt:

*disfigured, deformed*

*bị biến dạng, bị biến dạng*

(Lưu ý tôi đã đổi tên phần nhúng thành *\_kuvshinov.bin*)

### **15.7. Sự khác biệt giữa Embedding, Dreambooth và Hypernetwork**

Có ba phương pháp phổ biến để tinh chỉnh các mô hình Stable diffusion: embedding, dreambooth và hypernetwork.

**Embedding** định nghĩa keyword mới để mô tả một khái niệm mới mà không thay đổi mô hình. Các vectơ nhúng được lưu trữ trong tệp .bin hoặc .pt. Kích thước tệp của nó rất nhỏ, thường dưới 100 kB.

**Dreambooth** đưa ra một khái niệm mới bằng cách tinh chỉnh toàn bộ mô hình. Kích thước tệp là điển hình của Stable diffusion, khoảng 2 – 4 GB. Phần mở rộng của tệp giống như các mô hình khác, ckpt.

**Hypernetwork** là một mạng bổ sung được gắn vào mô hình UNet khử nhiễu của mô hình Stable diffusion. Mục đích là để tinh chỉnh mô hình mà không thay đổi mô hình. Kích thước tập tin thường là khoảng 100 MB.

### **15.8. Ưu và nhược điểm của việc sử dụng nhúng**

Một trong những lợi thế của việc sử dụng tính năng nhúng là kích thước nhỏ của nó. Với kích thước tệp từ 100 KB trở xuống, thật đơn giản để lưu trữ nhiều tệp trong máy tính của bạn. Vì nội dung nhúng chỉ là keyword mới nên chúng có thể được sử dụng cùng nhau trong cùng một hình ảnh.

Hạn chế của việc sử dụng tính năng nhúng là đôi khi không rõ nó sẽ được sử dụng với mô hình nào. Nếu không nói, bạn có thể bắt đầu với v1.4 hoặc v1.5. Bạn cũng có thể muốn bao gồm VAE để xem điều đó có tạo ra sự khác biệt nào không. Đối với phong cách anime.

Nói chung, tôi thấy việc sử dụng tính năng nhúng khó hơn một chút so với việc sử dụng các mô hình tùy chỉnh.

## **Chương 16. Mô hình V2**

Stable Diffusion v2.0 được phát hành lần đầu tiên vào tháng 11 năm 2022, tiếp theo là v2.1 một tháng sau đó. Nói chung, chúng ta gọi loạt mô hình v2.x là **mô hình v2**.

### **16.1. Sự khác biệt giữa mô hình v1 và v2**

Stable Diffusion v2 khác với v1 ở một số điểm

- Mô hình nhúng văn bản lớn hơn, giúp cải thiện chất lượng hình ảnh.
- Độ phân giải hình ảnh cao hơn. Hiện nay các mẫu có sẵn ở độ phân giải 512 và 768 pixel.
- Bạn không thể tạo hình ảnh có yếu tố khóa tân (NSFW) nữa vì chúng đã bị xóa khỏi dữ liệu huấn luyện.
- Đã có mô hình độ sâu cho hình ảnh mới. (Thêm về điều này trong Chương 18)

## 16.2. Các biến thể mẫu V2

Có hai mô hình chuyển văn bản thành hình ảnh:

- Mô hình cơ sở 2.1: Kích thước hình ảnh mặc định là  $512 \times 512$  pixel
- Model 2.1: Kích thước hình ảnh mặc định là  $768 \times 768$  pixel

Model 768 có thể tạo ra hình ảnh lớn hơn. Bạn có thể đặt kích thước hình ảnh thành  $768 \times 768$  mà không phải lo lắng về vấn đề lỗi hình ảnh tạo ra có hai đầu. Điều này đặc biệt hữu ích khi tạo cảnh lớn hơn với các ký tự nhỏ. Các khuôn mặt có thể được tạo ra rõ ràng hơn một chút so với mô hình 512.

Nhược điểm của mẫu 768 là mất nhiều thời gian hơn để tạo ra hình ảnh. Hình ảnh lớn hơn có thể giới hạn kích thước lô, tùy thuộc vào lượng VRAM mà GPU của bạn có

## 16.3. Tiếp nhận các mô hình v2

Cộng đồng Stable diffusion thường nhận thấy mô hình 2.0 tạo ra kết quả kém hơn mô hình v1. Nhìn chung, tính thẩm mỹ của hình ảnh được cải thiện. Do đó, chúng ta sẽ chỉ đề cập đến việc cài đặt mô hình v2.1 trong Chương này.

Một tính năng thú vị của mô hình v2 là Negative prompt trở nên rất quan trọng. Bạn không thể tạo ra hình ảnh đẹp nếu không sử dụng nó. Chúng tôi sẽ đề cập đến điều đó trong Chương 17.

## 16.4. Cài đặt Model v2.1 trong AUTOMATIC1111

Để cài đặt **512 base model**:

1. Tải xuống tệp mô hình (v2-1\_512-ema-pruned.ckpt) bằng URL bên dưới.

[https://huggingface.co/stabilityai/stable-diffusion-2-1-base/resolve/main/v2-1\\_512-ema-cắt tia.ckpt](https://huggingface.co/stabilityai/stable-diffusion-2-1-base/resolve/main/v2-1_512-ema-cắt tia.ckpt)

2. Tải xuống tệp cấu hình bằng URL bên dưới. Đổi tên nó thành **v2-1\_512-ema-pruned.yaml**.

<https://raw.githubusercontent.com/Stability-AI/stablediffusion/main/configs/stable-diffusion/v2-suy luân.yaml>

3. Đặt cả hai vào thư mục models:

stable-diffusion-webui/models/Stable-diffusion

Để cài đặt **768 base model**:

1. Tải xuống tệp mô hình (v2-1\_512-ema-pruned.ckpt) bằng URL bên dưới  
[https://huggingface.co/stabilityai/stable-diffusion-2-1/resolve/main/v2-1\\_768-ema-pruned.ckpt](https://huggingface.co/stabilityai/stable-diffusion-2-1/resolve/main/v2-1_768-ema-pruned.ckpt)

2. Tải xuống tệp cấu hình bằng URL bên dưới. Đổi tên nó thành **v2-1\_512-ema-pruned.yaml**.

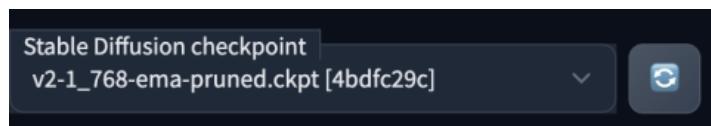
<https://raw.githubusercontent.com/Stability-AI/stablediffusion/main/configs/stable-diffusion/v2-suy-luan-v.yaml>

3. Đặt cả hai vào thư mục models:

stable-diffusion-webui/models/Stable-diffusion

### 16.5. Cách sử dụng mô hình 2.1

Để sử dụng phiên bản 768 của mô hình Stable Diffusion 2.1, hãy chọn *v2-1\_768-ema-pruned.ckpt* trong menu thả xuống Checkpoint Stable diffusion ở trên cùng bên trái.



Mô hình được thiết kế để tạo ra hình ảnh  $768 \times 768$ . Vì vậy, hãy đặt chiều rộng và/hoặc chiều cao của hình ảnh thành 768 để có kết quả tốt nhất.

Để sử dụng mô hình cơ sở, hãy chọn *v2-1\_512-ema-pruned.ckpt* thay vì.

### 16.6. Xử lý sự cố

Hãy thử các cách sau nếu quá trình cài đặt không thành công.

- Xem liệu GUI AUTOMATIC1111 của bạn có lỗi thời hay không. Trong terminal, sử dụng lệnh git pull trong thư mục stable-diffusion-webui và khởi động lại GUI.
- Kiểm tra xem tệp yaml đã được tải xuống chính xác chưa. Nội dung của nó phải là một tệp văn bản đơn giản, không có thẻ HTML.
- Kiểm tra xem tệp yaml có được đổi tên chính xác như mô tả trong phần trước hay không.
- Nếu 2.0 hoặc 2.1 đang tạo ra hình ảnh màu đen, hãy bật độ chính xác hoàn toàn với tối ưu hóa các đối số khởi động --no-half hoặc --xformers.



Một hình ảnh 2.1.

### 16.7. Lời khuyên khi sử dụng mô hình 2.1

**1. Viết chi tiết Prompt.** Tương tự như 2.0, Prompt cần phải thật cụ thể và chi tiết để có được hình ảnh như ý muốn.

**2. Sử dụng Negative prompt.** Negative prompt rất quan trọng đối với các mẫu v2. Tôi khuyên bạn nên giữ một Negative prompt có sẵn cho những bức chân dung, ví dụ:

*cropped, lowres, poorly drawn face, out of frame, poorly drawn hands, blurry, bad art, blurred, text, watermark, disfigured, deformed, closed eyes*

bị cắt, độ phân giải thấp, khuôn mặt được vẽ kém, ngoài khung, bàn tay được vẽ kém, mờ, nghệ thuật xấu, bị mờ, văn bản, hình mờ, bị biến dạng, bị biến dạng, nhầm mắt

Trong Chương 17 sẽ nói kỹ hơn về nội dung này.

**3. Sử dụng kích thước hình ảnh chính xác.** Đặt ít nhất một cạnh là 512 px cho mô hình cơ sở 512 và 768 px cho mô hình 768.

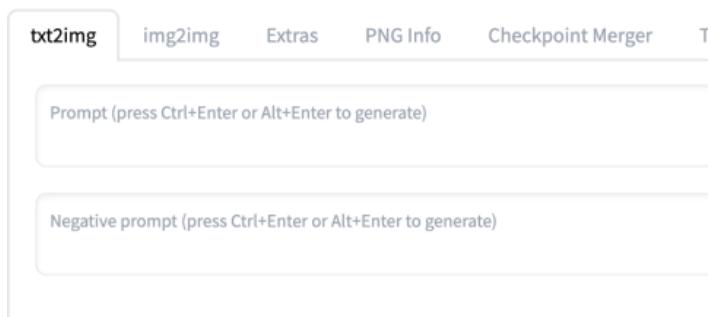
### Chương 17. Negative prompt

Negative prompt cung cấp cho bạn một cách bổ sung để kiểm soát việc tạo văn bản thành hình ảnh. Nhiều người coi nó như một tính năng tùy chọn trong các mô hình Stable diffusion 1.4 hoặc 1.5. Mọi thứ đã thay đổi khi phát hành Stable Diffusion v2. Negative prompt trở nên không thể thiếu.

Tôi sẽ hướng dẫn một số trường hợp sử dụng Negative prompt bao gồm sửa đổi nội dung và sửa đổi phong cách. Sau đó tôi sẽ chứng minh tầm quan trọng của những Negative

prompt trong mô hình v2. Tôi sẽ trình bày cách tìm kiếm một *Negative prompt* phổ quát.

## 17.1. Cách nhập Negative prompt



Nhiều GUI hoặc dịch vụ web Stable diffusion đưa ra những Negative prompt. Trong AUTOMATIC1111, bạn nhập Negative Prompt ngay bên dưới nơi bạn đặt Prompt.

Tiếp theo, tôi sẽ điểm qua một số trường hợp sử dụng Negative prompt.

## 17.2. Xóa đối tượng

Cách sử dụng rõ ràng đầu tiên là xóa bất cứ thứ gì bạn không muốn thấy trong ảnh. Giả sử bạn đã tạo ra một bức tranh về Paris trong một ngày mưa. Hãy xem xét hình ảnh sau đây và Prompt.



*Prompt: autumn in paris, ornate, beautiful, atmosphere, vibe, mist, smoke, fire, chimney, rain, wet, pristine, puddles, melting, dripping, snow, creek, lush, ice, bridge, forest, roses, flowers, by stanley artgerm lau, greg rutkowski, thomas kinkade, alphonse mucha, loish, norman rockwell.*

*Prompt: mùa thu ở Paris, trang trí công phu, đẹp, bầu không khí, rung cảm, sương mù, khói, lửa, ống khói, mưa, ẩm ướt, nguyên sơ, vũng nước, tan chảy, nhỏ giọt, tuyết, lạch, tui tốt, băng, cây cầu, rừng, hoa hồng, hoa, của stanley artgerm lau, greg rutkowski, thomas kinkade, alphonse mucha, loish, Norman rockwell.*

Bạn muốn tạo một cái khác nhưng là một con đường vắng. Những gì bạn có thể làm là sử dụng cùng một giá trị Seed chỉ định hình ảnh và thêm Negative prompt “people”. Bạn

nhận được một hình ảnh với hầu hết mọi người bị loại bỏ.



Lưu ý rằng cảnh rất giống nhưng không hoàn toàn giống với cảnh gốc. Nếu bạn cần bản gốc, bạn sẽ cần sử dụng tính năng inpainting để loại bỏ người một cách tỉ mỉ trong khi vẫn giữ cho khung cảnh mạch lạc.

Bạn có thể nhận thấy rằng vẫn còn một người trong hình trên. Bạn có thể yêu cầu Stable Diffusion cố gắng hơn nữa bằng cách nhấn mạnh vào Negative prompt (*people:1.3*). Điều đó cho Stable Diffusion biết rằng keyword *people* bây giờ quan trọng hơn 30%.



### 17.3. Sửa đổi hình ảnh

Bạn có thể thúc đẩy Stable diffusion để thực hiện những thay đổi tinh tế bằng những Negative prompt. Chính xác là bạn không muốn xóa bất cứ thứ gì ngoài việc thực hiện những thay đổi nhỏ đối với chủ đề. Hãy làm việc với hình ảnh cơ bản này:



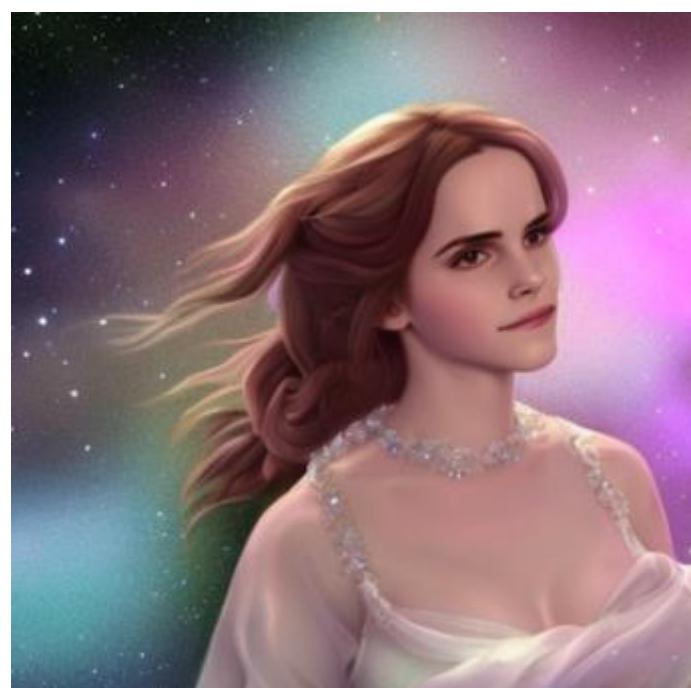
*emma watson as nature magic celestial, top down pose, long hair, soft pink and white transparent cloth, space, D&D, shiny background, intricate, elegant, highly detailed, digital painting, artstation, concept art, smooth, sharp focus, illustration, artgerm, bouguereau.*

*emma watson trong vai thiên nhiên kỳ diệu, tư thế từ trên xuống, tóc dài, vải trong suốt màu trắng và hồng mềm mại, không gian, D&D, nền sáng bóng, phức tạp, thanh lịch, có độ chi tiết cao, vẽ tranh kỹ thuật số, artstation, nghệ thuật khái niệm, mịn màng, tiêu điểm sắc nét, hình minh họa, Artgerm, bouguereau.*

Hình như trời có gió và tóc bay bay. Hãy sử dụng Prompt phủ định “windy” để giữ cho tóc xõa xuống.



Emma trong ảnh gốc trông non trẻ. Sử dụng **Negative prompt** “*underage*” khiến cô ấy trông trưởng thành hơn.



Nếu chúng ta thuận gió nhưng muôn tóc che tai thì sao? Hãy thêm vào **Negative prompt** “*tai*” với các yếu tố nhán mạnh khác nhau. Dưới đây là ba điểm nhán tăng dần 1.3, 1.6 và 1.9.



Tai bị tóc che phủ nhiều hơn ở tất cả các yếu tố nhấn mạnh nhưng khi hệ số đạt 1,9, bối cảnh của hình ảnh đã thay đổi. Negative prompt có thể ảnh hưởng mạnh mẽ đến quá trình khuếch tán.

#### 17.4. Negative prompt khi chuyển đổi keyword

Bây giờ điều gì sẽ xảy ra nếu bạn thực sự muốn sử dụng điểm nhấn cao? (ear:1.9)? Tôi không biết tai của bạn có vấn đề gì nhưng tôi có một mẹo nhỏ dành cho bạn. Bạn có thể sử dụng tính năng chuyển đổi keyword để trước tiên sử dụng một từ vô nghĩa làm Prompt phủ định, sau đó chuyển sang (ear:1.9) ở bước lấy mẫu sau.

Hãy chọn *các* như một Negative prompt vô nghĩa, ngu ngốc. Bạn có thể xác minh sự vô dụng của nó bằng cách đặt nó vào Prompt phủ định. Bạn sẽ nhận được hình ảnh giống như khi bạn không đặt bất cứ thứ gì. Bây giờ hãy sử dụng điều này làm Prompt phủ định:

[the: (ear:1.9): 0.5]

Vì tôi đang sử dụng 20 bước lấy mẫu, điều này có nghĩa là sử dụng Prompt phủ định ở bước 1 – 10 và (ear:1.9) ở bước 11-20. Lý do đằng sau điều này là quá trình khuếch tán quan trọng nhất ở những bước đầu. Các bước sau chỉ là tinh chỉnh chi tiết hơn, chẳng hạn như phần lông che tai.



Bây giờ những gì chúng ta đã đạt được không có gì đáng kinh ngạc.

- Nay giờ chúng ta có thể sử dụng điểm nhấn mạnh hơn nhiều (ear:1.9) mà không thay đổi bối cảnh.
- Chúng ta có được một hình ảnh gần giống với hình ảnh gốc hơn nhiều.
- Tai được che lại.

### 17.5. Sửa đổi kiểu

Negative prompt không chỉ hữu ích trong việc sửa đổi nội dung mà còn sửa đổi phong cách. Tại sao sử dụng Prompt phủ định để thay đổi phong cách? Đôi khi việc thêm quá nhiều vào Positive prompt chỉ khiến bộ khuếch tán bối rối. Hãy tưởng tượng ai đó bảo bạn đi đến 77 địa điểm (giới hạn mã thông báo) cùng một lúc. Sẽ rất hữu ích nếu họ cho bạn biết những lĩnh vực nào cần **tránh xa**.

### 17.6. Làm sắc nét

Thay vì sử dụng keyword “sharp”, “focused” trong Prompt, bạn có thể sử dụng “**blurry**” trong Negative prompt. Hình ảnh trở nên sắc nét hơn.

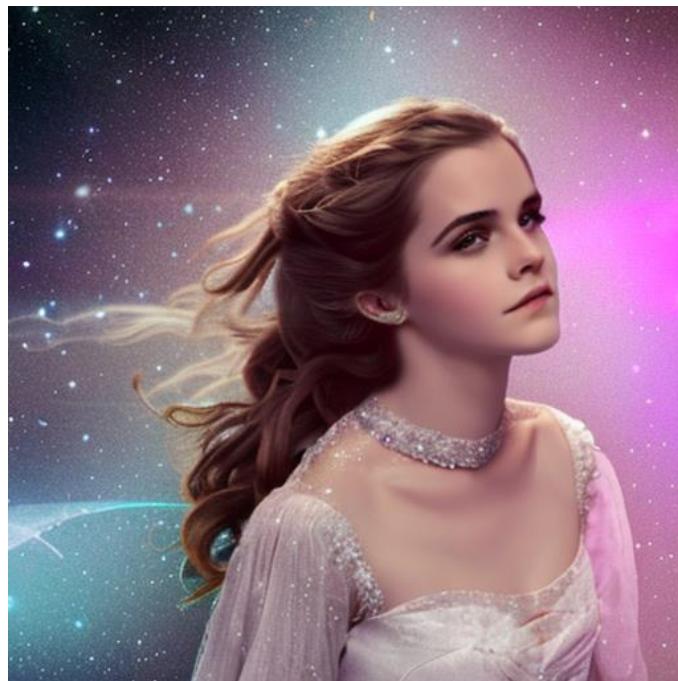


### 17.7. Ảnh thực tế

Sử dụng **negative prompt painting, cartoon** làm cho nó giống ảnh hơn.



Nếu bạn muốn giữ nguyên bối cảnh ban đầu, bạn có thể thử nghiệm chuyển đổi keyword mà tôi đã đề cập trước đó. Sử dụng [the: (painting animation:1.9): 0.3] ta có:



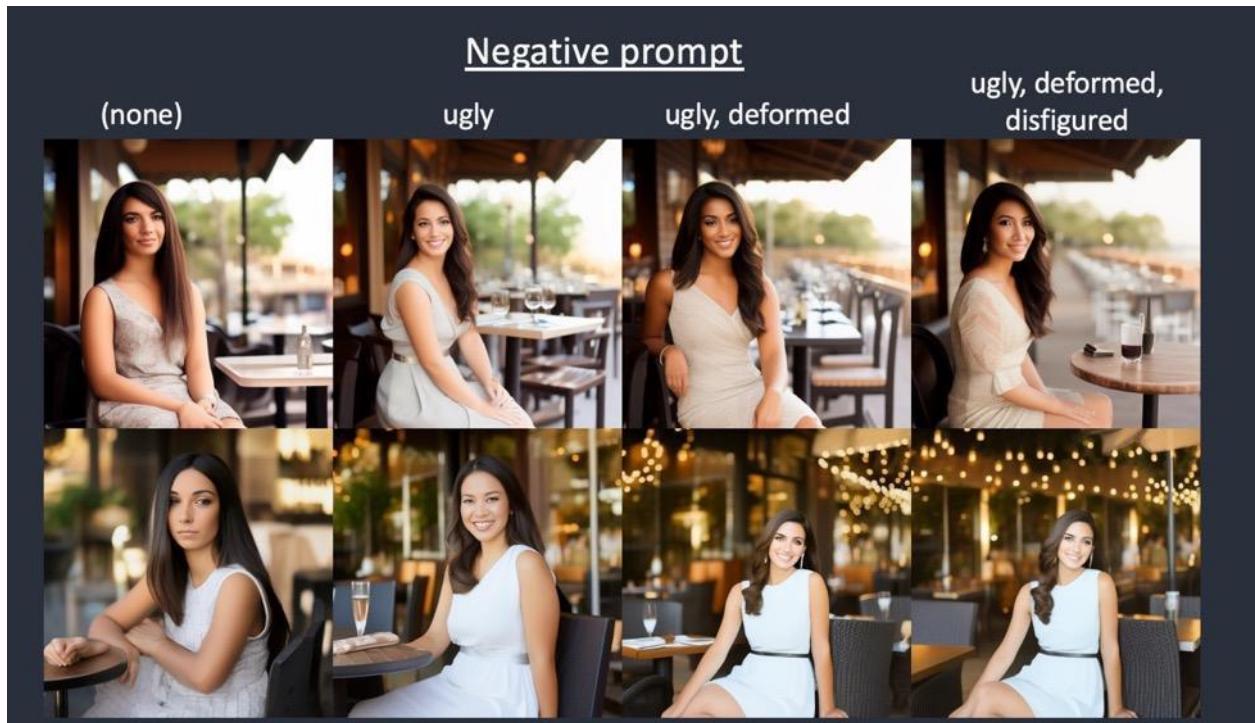
Nó gần giống với bản gốc hơn nhiều nhưng có thêm phong cách mới.

### 17.8. Negative prompt rất quan trọng đối với các mẫu v2

Kinh nghiệm của riêng tôi là Negative prompt rất quan trọng đối với các mẫu v2. Dưới đây tôi đã sử dụng Positive prompt để tạo ra con người thực tế nhưng với mô hình Stable diffusion 2.1.

*a young female, highlights in hair, sitting outside restaurant, brown eyes, wearing a dress, side light*

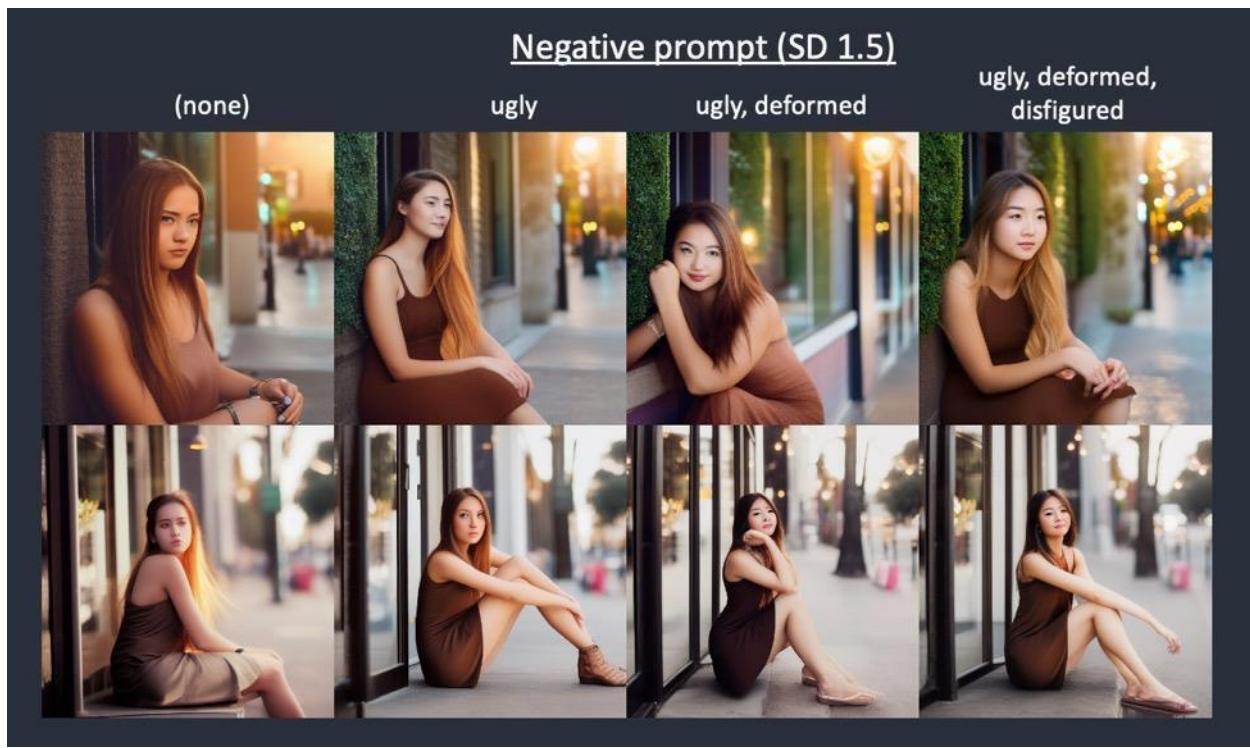
một phụ nữ trẻ, tóc highlight, ngoài nhà hàng, mắt nâu, mặc váy, đèn chiếu sáng bên



Chỉ thêm hai đến ba Negative prompt sẽ dần dần cải thiện tính thẩm mỹ của hình ảnh. Tôi có thể nói rằng điều này khá gần với chất lượng của các mẫu v1.

### 17.9. Negative prompt với v1.5

Hãy lặp lại bài tập trên mô hình v1.5. Hình ảnh hiển thị khá tốt mà không có bất kỳ Negative prompt nào trong phiên bản 1.5. Thêm Negative prompt *ugly, deformed and disfigured* có thể đã cải thiện nhiều thứ nhưng nó không rõ ràng như trong v2.1. Như thế mô hình v1.5 không hiểu những từ này.



### 17.10. Tại sao Negative prompt trở nên quan trọng hơn trong v2?

Hai thay đổi trong v2 là:

- Sử dụng mô hình ngôn ngữ OpenCLIP lớn hơn.
- Lọc nội dung NSFW trong dữ liệu đào tạo.

Nguyên nhân đầu tiên chuyển từ mô hình CLIP của Open AI sang OpenCLIP. Điều này ảnh hưởng đến việc nhúng của mô hình. AI mờ đã đào tạo mô hình CLIP bằng dữ liệu độc quyền. Nếu dữ liệu được quản lý chặt chẽ đến mức mỗi người đều có ngoại hình trên mức trung bình, thì việc nhắc đến “woman” sẽ giống như nhắc nhở “beautiful woman” (đó là ý của người dùng).

Thứ hai là những gì được coi là NSFW cũng có thể có tính thẩm mỹ cao. Đó có thể là lỗi của bộ lọc hoặc đó chỉ là bản chất của hình ảnh NSFW. Việc loại trừ hình ảnh NSFW cũng vô tình làm sai lệch dữ liệu theo hướng xấu và xấu.

### 17.11. Negative prompt phổ quát cho V2

Chúng ta đã đề cập đến tầm quan trọng của Negative prompt trong v2. Nay chúng ta hãy tìm một gợi ý tiêu cực phổ quát hay. Tôi sẽ sử dụng mô hình 2.1 (512 pixel) cho thử nghiệm này. Các bản gốc không có Prompt phủ định là



*Nguyên bản.*

Không tệ nhưng có thể được cải thiện. Sử dụng Negative prompt tối giản của chúng tôi “*ugly, disfigured, deformed*”, chúng tôi thấy ngay những cải tiến:



*Negative prompt: xấu xí, biến dạng, biến dạng*

Thêm *underexposed* và *overexposed* (*phơi sáng quá mức*) giúp hình ảnh bớt phẳng hơn.



*Negative prompt: underexposed, overexposed, ugly, disfigured, deformed.*

*Negative prompt: thiếu sáng, phơi sáng quá mức, xấu xí, biến dạng, biến dạng.*

Sẽ không hại gì khi thêm *độ tương phản thấp*.



*Negative prompt: độ tương phản thấp, thiếu sáng, phơi sáng quá mức, xấu, biến dạng, biến dạng*

Tiếp theo, hãy kiểm tra Negative prompt phổ biến này về v2 trôi nổi trên internet:

*ugly, tiling, poorly drawn hands, poorly drawn feet, poorly drawn face, out of frame, extra limbs, disfigured, deformed, body out of frame, blurry, bad anatomy, blurred, watermark, grainy, signature, cut off, draft*

xấu, óp lát, tay vẽ kém, chân vẽ kém, mặt vẽ kém, ngoài khung, chân tay thừa, biến dạng, biến dạng, thân ngoài khung, mờ, giải phẫu xấu, mờ, hình mờ, sân sùi, chữ ký, cắt đứt, nháp



Tôi nghĩ nó đang hoạt động tốt mặc dù nó có thể đã sửa đổi phong cách một chút. Điều này có thể do keyword phủ định gây ra *blurry, blurred, grainy, draft*. Một số phong cách có thể trông giống như vậy. Việc xóa những keyword này dường như sẽ trở lại gần hơn với phong cách ban đầu.

*ugly, tiling, poorly drawn hands, poorly drawn feet, poorly drawn face, out of frame, extra limbs, disfigured, deformed, body out of frame, bad anatomy, watermark, signature, cut off*

xấu, óp lát, tay vẽ kém, chân vẽ kém, mặt vẽ kém, ngoài khung, chân tay thừa, biến dạng, biến dạng, cơ thể ngoài khung, giải phẫu xấu, hình mờ, chữ ký, bị cắt bỏ



Tiếp theo, thêm keyword chiếu sáng mà chúng ta vừa sử dụng (*độ tương phản thấp, thiếu sáng, phơi sáng quá mức*). Nó giúp ích cho độ tương phản và dải động.

*ugly, tiling, poorly drawn hands, poorly drawn feet, poorly drawn face, out of frame, extra limbs, disfigured, deformed, body out of frame, bad anatomy, watermark, signature, cut off, low contrast, underexposed, overexposed*

*xấu, óp lát, tay vẽ kém, chân vẽ kém, mặt vẽ kém, ngoài khung, chân tay thừa, biến dạng, biến dạng, cơ thể nằm ngoài khung, giải phẫu xấu, hình mờ, chữ ký, bị cắt, độ tương phản thấp, thiếu sáng, phơi sáng quá mức*



Bây giờ, việc thêm một số keyword phủ định khác để tránh lấy mẫu tác phẩm nghệ thuật xấu hoặc các bức vẽ của người mới, chúng tôi đi đến Negative prompt cuối cùng bên dưới. Đây là một **Negative prompt phô quát mà không ảnh hưởng đến phong cách**.

*ugly, tiling, poorly drawn hands, poorly drawn feet, poorly drawn face, out of frame, extra limbs, disfigured, deformed, body out of frame, bad anatomy, watermark, signature, cut off, low contrast, underexposed, overexposed, bad art, beginner, amateur, distorted face*

xấu, ốp lát, tay vẽ kém, chân vẽ kém, mặt vẽ kém, ngoài khung, chân tay thừa, biến dạng, biến dạng, cơ thể nằm ngoài khung, giải phẫu xấu, hình mờ, chữ ký, bị cắt, độ tương phản thấp, thiếu sáng, phơi sáng quá mức, xấu nghệ thuật, người mới bắt đầu, nghiệp dư, khuôn mặt méo mó



Đây là một cải tiến lớn so với việc không có Negative prompt. Bạn có thể muôn loại bô độ tương phản thấp, thiếu sáng hoặc phơi sáng quá mức nếu đó là phong cách.

### 17.12. Kiểm tra Negative prompt phổ quát

Hãy thử nghiệm Negative prompt chung mà chúng ta vừa phát hiện ra.

*ugly, tiling, poorly drawn hands, poorly drawn feet, poorly drawn face, out of frame, extra limbs, disfigured, deformed, body out of frame, bad anatomy, watermark, signature, cut off, low contrast, underexposed, overexposed, bad art, beginner, amateur, distorted face*

xấu, ốp lát, tay vẽ kém, chân vẽ kém, mặt vẽ kém, ngoài khung, chân tay thừa, biến dạng, biến dạng, cơ thể nằm ngoài khung, giải phẫu xấu, hình mờ, chữ ký, bị cắt, độ tương phản thấp, thiếu sáng, phơi sáng quá mức, xấu nghệ thuật, người mới bắt đầu, nghiệp dư, khuôn mặt méo mó

Hãy thử nghiệm đầu tiên **phong cách chụp ảnh** với Prompt sau.

*A man walking around her neighborhood, highlight hair, detailed eyes, sharp focus, young face, perfect symmetric face, pupil reflecting surroundings, realistic skin, soft healthy skin*

*Một người đàn ông đi dạo quanh khu phố của cô, mái tóc nổi bật, đôi mắt chi tiết, tiêu*

*điểm sắc nét, khuôn mặt trẻ, khuôn mặt cân đối hoàn hảo, đồng tử phản chiếu khung cảnh xung quanh, làn da chân thực, làn da mềm mại khỏe mạnh*



*Bên trái: Không có Negative prompt. Bên phải: Có Negative prompt.*

Negative prompt chung hoạt động tốt với hình ảnh theo phong cách ảnh. Anh chàng trông cao cấp hơn rất nhiều và đã dành nhiều thời gian hơn cho mái tóc của mình vào buổi sáng...

Điều gì sẽ xảy ra nếu chúng ta sử dụng Prompt phủ định với **phong cách hoạt hình?**

Prompt:

*anime style girl on battleground, holding a ninja sword, detailed eyes, perfect face*

*Cô gái phong cách anime trên chiến trường, cầm kiếm ninja, đôi mắt chi tiết, khuôn mặt hoàn hảo*



*Bên trái: Không có Negative prompt. Bên phải: Có Negative prompt.*

Negative prompt phổ quát đã giúp các nhân vật hoạt động tốt như nhau trong phong cách

anime. Đối tượng trông đẹp hơn, đẹp trai hơn và sẵn sàng chiến đấu hơn. Thanh kiếm ninja được dựng thẳng lên trông nguy hiểm hơn.

Cuối cùng, hãy thử nghiệm với phong cách vẽ tranh sơn dầu.

*impressionist oil painting of a young man standing right next to a red tesla roadster by john sargent*

*Bức tranh sơn dầu theo trường phái ấn tượng vẽ một chàng trai trẻ đứng cạnh chiếc xe Tesla Roadster màu đỏ của John Sargent*



Bên trái: Không có Prompt phủ định. Đúng: VỚI Prompt phủ định.

Negative prompt chung sẽ giúp ích cho cả Tesla và anh chàng. Thay vì trưng bày một chiếc ô tô cũ nát với một cậu thiếu niên rắc rối, giờ đây nó lại trưng bày một chiếc ô tô mới sáng bóng với một chàng trai trẻ trông có vẻ triệu đô.

## PHẦN IV - CÁC CHỦ ĐỀ NÂNG CAO

### Chương 18. Độ sâu của hình ảnh

Độ sâu trên hình ảnh (Depth2img) là mô hình được đánh giá thấp trong Stable Diffusion v2. Đây là một cải tiến đối với tính năng chuyển image-to-image (img2img), tận dụng thông tin có chiều sâu khi tạo hình ảnh mới. Trong Chương này, chúng ta sẽ xem xét kỹ hơn để xem nó là gì, cách cài đặt và sử dụng nó cũng như nó có thể làm gì cho bạn.

#### 18.1. Độ sâu của hình ảnh có thể làm gì

Với độ sâu của hình ảnh, bạn có thể kiểm soát tốt hơn việc tổng hợp chủ đề và hậu cảnh riêng biệt. Giả sử tôi muốn biến khung cảnh lãng mạn này thành *La La Land* vào một trận đấu vật...



*Ảnh gốc.*

Chúng ta sẽ đi vào chi tiết hơn sau, nhưng hiện tại chỉ coi độ sâu trên hình ảnh là phiên bản nâng cao của image-to-image. Chúng có thể được sử dụng theo cách tương tự - với một hình ảnh và Prompt văn bản, nó sẽ tạo ra một hình ảnh mới. Giả sử tôi sử dụng Prompt

*photo of two men wrestling*

*hình ảnh hai người đàn ông đang vật lộn*

cho cả hình ảnh với hình ảnh và độ sâu của hình ảnh. Dưới đây là kết quả với Denoising strength thay đổi từ 0,4 đến 1,0. (Hãy nhớ rằng Denoising strength càng cao thì hình ảnh sẽ càng thay đổi nhiều.)



*So sánh hình ảnh với hình ảnh và độ sâu với hình ảnh.*

Chúng ta hãy xem các thế hệ từ hình ảnh đến hình ảnh (hàng trên cùng). Chúng tôi gặp phải một vấn đề: Ở Denoising strength thấp, hình ảnh không đủ thay đổi. Ở Denoising

strength cao, chúng ta thấy hai đồ vật nhưng thành phần ban đầu bị mất. Độ sâu của hình ảnh giải quyết vấn đề này. Bạn có thể tăng Denoising strength lên tới 1 (tối đa) mà không làm mất bối cảnh ban đầu!

Bây giờ bạn đã biết độ sâu của hình ảnh có thể làm được gì, hãy xem nó hoạt động như thế nào.

### 18.2. Độ sâu của hình ảnh là gì?

Trong tính năng chuyển đổi image-to-image, Stable diffusion lấy hình ảnh và Prompt làm đầu vào. Việc tạo hình ảnh dựa trên CẢ HAI hình ảnh và Prompt. Hình ảnh cuối cùng giống với hình ảnh đầu vào về màu sắc và hình dạng.

Trong chiều sâu của hình ảnh, Stable diffusion tương tự lấy hình ảnh và Prompt làm đầu vào. Trước tiên, mô hình này ước tính bản đồ độ sâu của hình ảnh đầu vào bằng MIDaS, một mô hình AI được phát triển vào năm 2019 để ước tính khả năng nhận biết độ sâu bằng một mắt (tức là ước tính độ sâu từ một góc nhìn). Bản đồ độ sâu sau đó được Stable Diffusion sử dụng như một điều kiện bổ sung cho việc tạo hình ảnh.

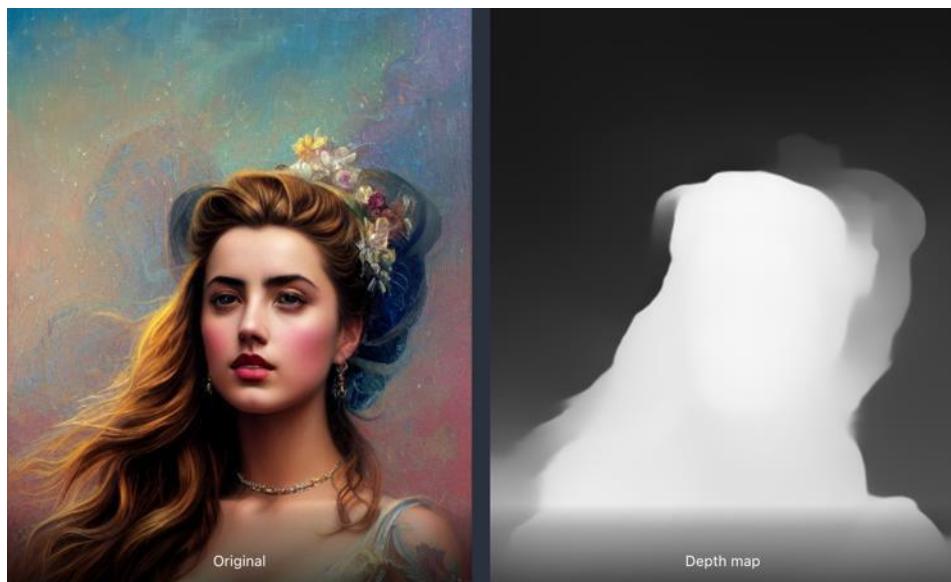
Nói cách khác, độ sâu của hình ảnh sử dụng ba điều kiện để tạo ra một hình ảnh mới: (1) Prompt văn bản, (2) hình ảnh gốc và (3) bản đồ độ sâu.

Được trang bị bản đồ độ sâu, mô hình có *một số* kiến thức về bối cảnh ba chiều của cảnh. **Các thế hệ hình ảnh của các đối tượng nền trước và nền có thể được tách riêng.**

### 18.3. Bản đồ độ sâu (Depth map)

Bạn không cần cung cấp bản đồ độ sâu để sử dụng độ sâu cho hình ảnh. Nó được tạo tự động bởi mô hình.

Bản đồ độ sâu là một hình ảnh thang màu xám đơn giản có cùng kích thước với hình ảnh gốc mã hóa thông tin độ sâu. Màu trắng hoàn toàn có nghĩa là vật thể ở gần bạn nhất. Nhiều màu đen hơn có nghĩa là xa hơn. Dưới đây là ví dụ về hình ảnh và bản đồ độ sâu của nó được MIDaS ước tính.



#### 18.4. Cài đặt mô hình độ sâu cho hình ảnh

Để cài đặt mô hình độ sâu trên hình ảnh trong GUI AUTOMATIC1111:

1. Tải về **tập tin mô hình** (512-độ sâu-ema.ckpt)

(URL:<https://huggingface.co/stabilityai/stable-diffusion-2-deep/resolve/main/512-depid-me.ckpt> )

2. Tải về **tập tin cấu hình**, đổi tên nó thành 512-deep-ema.yaml

(URL:[https://raw.githubusercontent.com/Stability-AI/stablediffusion/main/configs/stable\\_khuếch tán/v2-midas-inference.yaml](https://raw.githubusercontent.com/Stability-AI/stablediffusion/main/configs/stable_khuếch tán/v2-midas-inference.yaml) )

Đặt cả hai vào thư mục mô hình:

stable-diffusion-webui/models/Stable-diffusion

#### 18.5. Sử dụng mô hình độ sâu trên hình ảnh

Để sử dụng mô hình, nhấn nút làm mới bên cạnh hộp chọn Checkpoint ở trên cùng bên trái. Chọn 512-deep-ema.ckpt để tải mô hình.



Lưu ý rằng mô hình độ sâu có thể được sử dụng cho hình ảnh và ảnh động, nhưng không thể sử dụng văn bản thành hình ảnh. Bạn sẽ thấy một lỗi nếu bạn cố gắng làm điều đó. Để sử dụng mô hình, hãy chuyển đến tab img2img. Làm theo hướng dẫn để sử dụng img2img và inpainting. Rõ ràng từ tên mẫu, đây là mẫu 512. Điều đó có nghĩa là nó hoạt động tốt nhất khi có ít nhất một cạnh của hình ảnh mới là 512 pixel.

Bây giờ bạn đã hiểu nó hoạt động như thế nào, cách cài đặt và sử dụng nó. Chúng ta hãy đi qua một vài trường hợp sử dụng.

#### 18.6. Trường hợp sử dụng 1: Thay thế cho Img2img

Giả sử bạn có một bức ảnh chân dung như thế này.



*Nguyên bản.*

Bạn muốn có sự đa dạng bằng cách thêm một phụ nữ châu Á vào. Nhưng bạn đã thiết kế các yếu tố bô sung xung quanh anh ta nên bạn không muốn hình dạng của người đó thay đổi. Bạn có thể làm điều đó với img2img, nhưng bạn không thể đặt Denoising strength quá cao vì bạn sẽ mất hình dạng ban đầu. Với img2img, đây là cách tốt nhất bạn có thể làm:



*Image2Image. Gợi ý: một phụ nữ châu Á xinh đẹp hạnh phúc với đôi mắt chi tiết hoàn hảo, đường nét khuôn mặt chi tiết, làn da chi tiết, ánh sáng tự nhiên, mái tóc dài. Độ khử nhiễu: 0,58*

Có một số thay đổi về hình dạng nhưng không quá tệ. Tuy nhiên, Denoising strength vẫn quá thấp để có thể làm sai lệch màu da của người đàn ông ban đầu. Ngoài ra, Prompt về mái tóc dài cũng không được tuân theo. Nay giờ vấn đề nan giải: Tăng Denoising strength lên 1 sẽ mang lại kết quả như chúng ta muốn nhưng chúng ta sẽ mất hình dạng ban đầu:



*Chuyển đổi image-to-image với cùng một Prompt nhưng Denoising strength được đặt thành 1.*

Sử dụng mô hình độ sâu trên hình ảnh sẽ cho phép đặt Denoising strength thành 1 mà không làm mất hình dạng ban đầu:



*Độ sâu của hình ảnh có cùng Prompt và Denoising strength được đặt thành 1.*

Lưu ý rằng hình dạng của người đàn ông là hoàn toàn được bảo tồn và Stable Diffusion

bằng cách nào đó tìm ra cách tạo ra mái tóc dài.

### 18.7. Trường hợp sử dụng 2: Inpainting

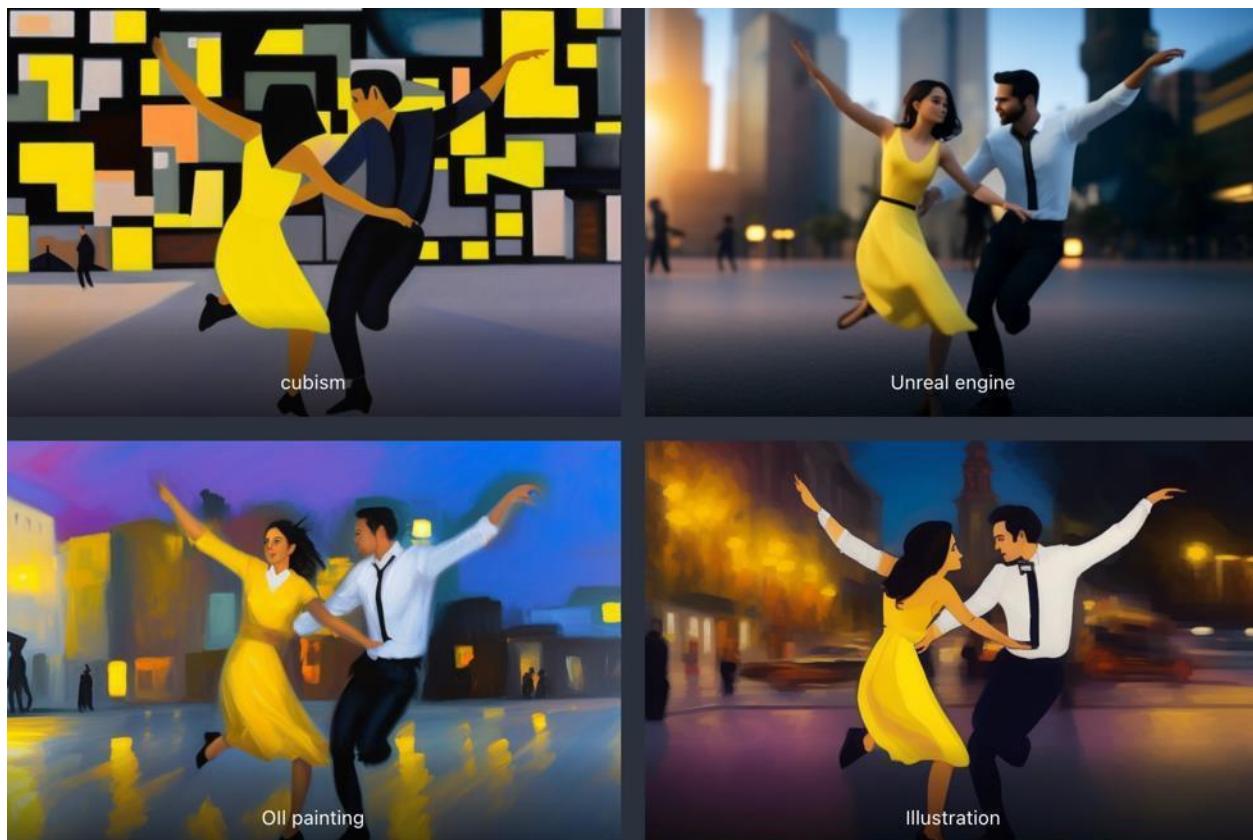
Tương tự, bạn có thể sử dụng độ sâu của hình ảnh trong việc vẽ nội dung, để sửa lỗi hoặc tạo nội dung mới. Bạn muốn sử dụng độ sâu của hình ảnh nếu bạn quan tâm đến việc giữ nguyên bối cảnh ban đầu.



Inpainting bằng mô hình có chiều sâu trên hình ảnh.

### 18.8. Chuyển phong cách

Lợi thế của độ sâu trên hình ảnh là bạn có thể điều chỉnh Denoising strength lên tối 1 mà không làm mất bối cảnh ban đầu. Điều đó làm cho việc chuyển đổi một cảnh sang một phong cách khác trở nên dễ dàng. Dưới đây là một số ví dụ:



Chuyển kiểu với Độ sâu cho hình ảnh.

### 18.9. Sao chép một tư thế có sẵn

Thật khó để tạo ra một tư thế cụ thể của con người với Stable diffusion. Với độ sâu của hình ảnh, bạn có thể sử dụng ảnh có tư thế bạn muốn làm ảnh cơ bản. Đặt Denoising strength thành 1! Tư thế sẽ được giữ nguyên hoàn toàn với độ sâu của hình ảnh. Bức ảnh có thể là cảnh phim, bức tranh hoặc ảnh bạn chụp bằng điện thoại.

Không còn chân tay thừa, bàn tay kỳ lạ và vô số bức vẽ để sửa tư thế!

## Chương 19. Sử dụng Dreambooth để đưa mọi thứ vào sự Stable diffusion



Dreambooth là một cách để đưa bất cứ thứ gì - người thân yêu, con chó của bạn, đồ chơi yêu thích của bạn - vào mô hình Stable diffusion. Chúng tôi sẽ giới thiệu Dreambooth là gì, nó hoạt động như thế nào và thực hiện đào tạo như thế nào.

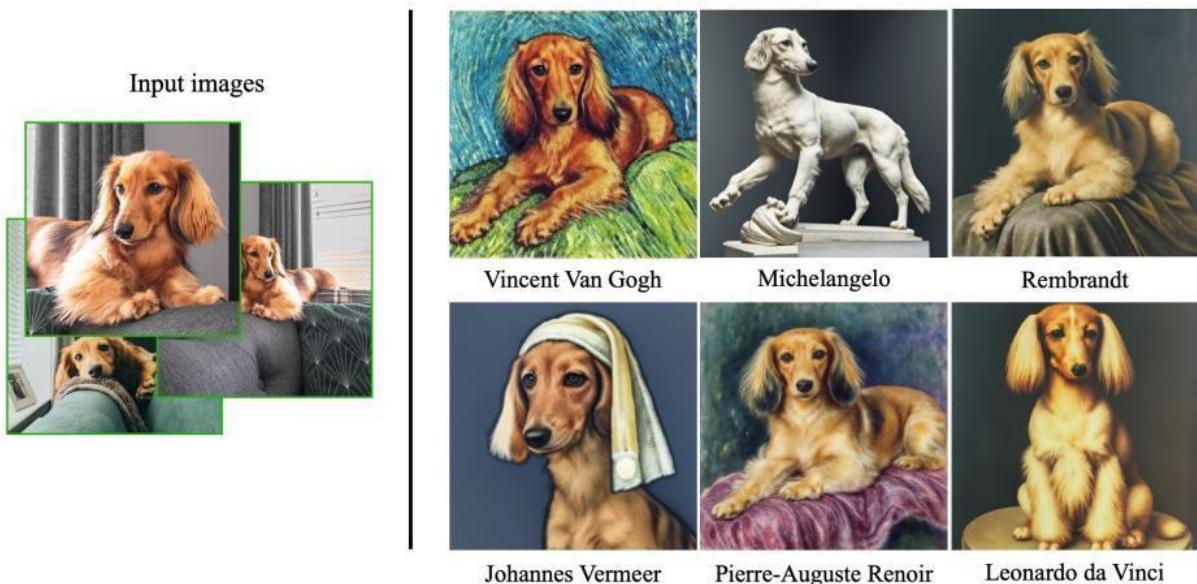
Bạn có biết nhiều Models tùy chỉnh được đào tạo bằng Dreambooth không? Sau khi hoàn thành hướng dẫn này, bạn sẽ biết cách tự làm.

### 19.1. Dreambooth là gì?

Được nhóm nghiên cứu của Google xuất bản vào năm 2022, Dreambooth là một kỹ thuật để tinh chỉnh các mô hình khuếch tán (như Stable diffusion) bằng cách đưa một chủ đề tùy chỉnh vào mô hình. Tại sao nó lại gọi là Dreambooth? Theo nhóm nghiên cứu của Google,

*Nó giống như một buồng chụp ảnh, nhưng một khi chủ đề được chụp, nó có thể được tổng hợp ở bất cứ nơi nào bạn mơ ước.*

Nghe thật là tuyệt! Nhưng nó hoạt động tốt như thế nào? Dưới đây là một ví dụ trong bài viết nghiên cứu. Chỉ sử dụng 3 hình ảnh của một con chó cụ thể làm đầu vào (Hãy gọi tên con chó này là Devora), mô hình dreamboothed có thể tạo ra hình ảnh của Devora trong các bối cảnh khác nhau.



*Chỉ với 3 hình ảnh đào tạo, Dreambooth sẽ đưa chủ đề tùy chỉnh vào mô hình khuếch tán một cách liền mạch.*

## 19.2. Những gì bạn cần để đào tạo Dreambooth

Bạn sẽ cần ba thứ

- Một số hình ảnh tùy chỉnh
- Mã định danh duy nhất
- Một tên lớp

Trong ví dụ trên. Mã định danh duy nhất là **Devora**. Tên lớp là **dog**. Sau đó, bạn sẽ cần xây dựng Prompt cá thể của mình:

*a photo of [unique identifier] [class name]*

*ảnh của [mã định danh duy nhất] [tên lớp]*

Và một Prompt của lớp:

*a photo of [class name]*

*một bức ảnh của [tên lớp]*

Trong ví dụ trên, **Prompt** là

*a photo of Devora dog*

*hình ảnh chú chó Devora*

Vì Devora là một con chó nên **Prompt của lớp** là

*a photo of a dog*

*một bức ảnh của một con chó*

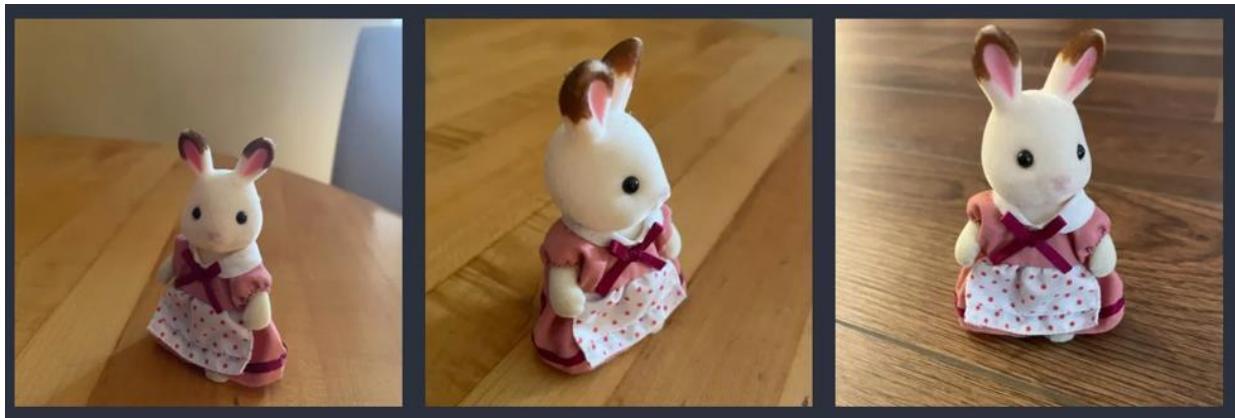
Bây giờ bạn đã hiểu mình cần gì, hãy bắt tay vào đào tạo!

URL của sổ tay Colab:

[https://colab.research.google.com/github/sagiodev/stablediffusion\\_webui/blob/master/DreamBooth\\_Stable\\_Diffusion\\_SDA.ipynb](https://colab.research.google.com/github/sagiodev/stablediffusion_webui/blob/master/DreamBooth_Stable_Diffusion_SDA.ipynb)

### **Bước 1: Lấy hình ảnh đào tạo**

Giống như bất kỳ nhiệm vụ học máy nào, dữ liệu đào tạo chất lượng cao là yếu tố quan trọng nhất dẫn đến thành công của bạn. Chụp 3-10 bức ảnh về chủ đề tùy chỉnh của bạn. Hình ảnh nên được chụp từ nhiều góc độ khác nhau. Chủ thể cũng phải có nhiều phông nền khác nhau để Models có thể phân biệt chủ đề với phông nền. Tôi sẽ sử dụng công cụ này trong phần hướng dẫn.

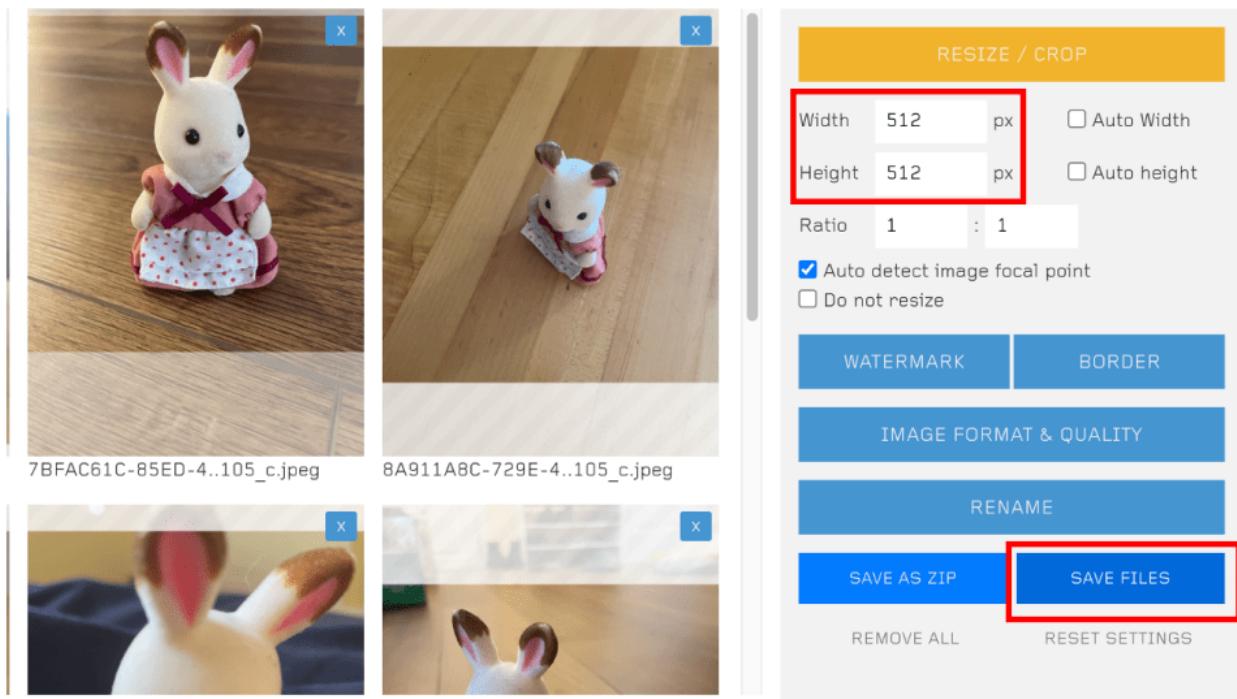


### **Bước 2: Thay đổi kích thước hình ảnh của bạn**

Để sử dụng hình ảnh trong quá trình đào tạo, trước tiên bạn cần thay đổi kích thước của chúng thành  $512 \times 512$  pixel để đào tạo với các mô hình v1.

[BIRMS](https://www.birme.net/?target_width=512&target_height=512) ([https://www.birme.net/?target\\_width=512&target\\_height=512](https://www.birme.net/?target_width=512&target_height=512)) là một trang web thuận tiện cho việc thay đổi kích thước hình ảnh.

1. Thả hình ảnh của bạn vào trang BIRME.
2. Điều chỉnh khung vẽ của mỗi hình ảnh để nó hiển thị đầy đủ chủ đề.
3. Đảm bảo chiều rộng và chiều cao đều là 512 px.
4. Nhấn **SAVE FILES** để lưu hình ảnh đã thay đổi kích thước vào máy tính của bạn.



### 19.3. Đào tạo (Training)

Bạn nên sử dụng Google Colab để đào tạo vì nó giúp bạn tránh khỏi rắc rối khi thiết lập. Toàn bộ quá trình đào tạo mất khoảng 30 phút. Nếu không sử dụng Google Colab Pro, bạn có thể khó hoàn thành Training mô hình mà không bị ngắt kết nối. Mua một số tín dụng điện toán để tránh vấn đề bị ngắt kết nối khi đào tạo mô hình. Kể từ tháng 12 năm 2022, 10 USD sẽ giúp bạn có được 50 giờ, vì vậy đây không phải là một khoản chi phí quá lớn. Sổ ghi chép sẽ lưu mô hình vào Google Drive của bạn. Đảm bảo bạn có ít nhất 2GB nếu chọn fp16 (được khuyến nghị) và 4GB nếu không.

1. Mở sổ ghi chép Colab (URL: [https://colab.research.google.com/github/sagiodev/stablediffusion\\_webui/blob/master/DreamBooth\\_Stable\\_Diffusion\\_SDA.ipynb](https://colab.research.google.com/github/sagiodev/stablediffusion_webui/blob/master/DreamBooth_Stable_Diffusion_SDA.ipynb) )
2. Bạn không cần thay đổi MODEL\_NAME nếu muốn huấn luyện từ mô hình Stable diffusion v1.5 (Được khuyến nghị).
3. Đặt vào **instance prompt** Và **class prompt**. Đối với hình ảnh của mình, tôi đặt tên cho con thỏ đồ chơi là zwx để Prompt mẫu của tôi là “photo of zwx toy” và Prompt lớp học của tôi là “photo of a toy”.

▶ DreamBooth

● Name/Path of the initial model.

**MODEL\_NAME:** "runwayml/stable-diffusion-v1-5"

Enter instance prompt and class prompt.  
 Example 1: photo of zwx person, photo of a person  
 Example 2: photo of zwx toy, photo of a toy

**instance\_prompt:** photo of zwx toy

**class\_prompt:** photo of a toy

**training\_steps:** 800

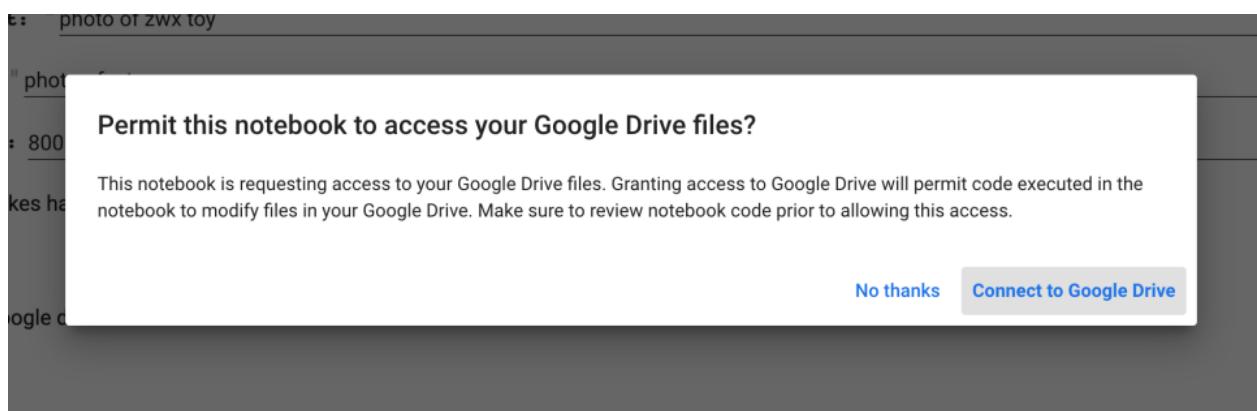
Convert to fp16? (takes half the space (2GB)).  
 **fp16:**

Save to model to Google drive?

Clear log after run?  
 **CLEAR\_LOG:**

4. Bấm vào nút Play ở bên trái ô để bắt đầu xử lý.

5. Cấp quyền truy cập Google Drive. Hiện tại không có cách nào dễ dàng để tải xuống tệp mô hình ngoại trừ việc lưu vào Google Drive.



6. Nhấn **Chọn tập tin** để tải lên các hình ảnh đã thay đổi kích thước.

Convert to fp16? (takes half the space (2GB)).

fp16:

Save to model to Google drive?

Clear log after run?

CLEAR\_LOG:

Show code

```
... Mounted at /content/drive
Tesla T4, 15109 MiB, 15109 MiB
[*] Weights will be saved at /content/stable_diffusion_weights/output
uploading instance images for `photo of zwx toy`
```

No file chosen

7. Sẽ mất khoảng 30 phút để hoàn thành đào tạo mô hình. Khi hoàn tất, bạn sẽ thấy một vài hình ảnh mẫu được tạo từ mô hình mới.



8. Mô hình tùy chỉnh của bạn sẽ được lưu trong Google Drive, trong thư mục *Dreambooth\_model*. Tải xuống tệp Checkpoint mô hình và cài đặt nó trong GUI yêu thích của bạn.

### Thử nghiệm mô hình

Bạn cũng có thể sử dụng ô thứ hai của sổ ghi chép để kiểm tra bằng mô hình.

▶ Test image generation from model

▶ **prompt:** "oil painting of zwx in style of van gogh  
**negative\_prompt:** "Insert text here  
**num\_samples:** 2  
**guidance\_scale:** 7.5  
**num\_inference\_steps:** 30  
**height:** 512  
**width:** 512  
**seed:** -1

[Show code](#)

Sử dụng Prompt

*oil painting of zwx in style of van gogh*

*tranh sơn dầu zwx theo phong cách van gogh*

với mô hình mới được đào tạo của mình, tôi hài lòng với những gì mình có:



## PHẦN IV – HƯỚNG DẪN THỰC HÀNH

### 19.4. Khôi phục ảnh cũ với ComfyUI

Các bước thực hiện như sau:

**B1:** Download bộ cài

[https://github.com/comfyanonymous/ComfyUI/releases/download/latest/new\\_ComfyUI\\_windows\\_portable\\_nvidia\\_cu121\\_or\\_cpu.7z](https://github.com/comfyanonymous/ComfyUI/releases/download/latest/new_ComfyUI_windows_portable_nvidia_cu121_or_cpu.7z)

**B2:** Cài đặt UI manager

<https://github.com/ltdrdata/ComfyUI-Manager.git>

**B3:** Download workflow

[https://drive.google.com/file/d/1xDI7sa\\_o4tYN3ICEBJ6zvgwj-PMI25g0/view](https://drive.google.com/file/d/1xDI7sa_o4tYN3ICEBJ6zvgwj-PMI25g0/view)

**B4:** Fix lỗi các Custom nodes

**B5:** Download các thư viện cần thiết (checkpoint, Vae, Controlnet...)

**B6:** Upload ảnh và sử dụng

Hướng dẫn chi tiết tại video: <https://youtu.be/XymTZGN-UiM>

### 19.5. Làm đẹp, mịn da cho ảnh



Hướng dẫn chi tiết tại video: <https://youtu.be/DUfhbKVQ5lk>

### 19.6. Tạo ảnh quảng cáo sản phẩm với ComfyUI

**B1. Download Workflow:** <https://drive.google.com/drive/folders/1oBTN5UZ7-yaBhkdyYa8F3VTO8OXF2JP3?usp=sharing>

**B2. Cài đặt custom node, fix lỗi:**

- Fix lỗi trên ComfyUI Manager

**B3. Download các file checkpoint, thư viện cần thiết vào thư mục**

- ComfyUI\models\Checkpoint:
  - **juggernautXL\_v8Rundiffusion.safetensors**
- ComfyUI\models\controlnet:
  - **control-lora-depth-rank256.safetensors**
- ComfyUI\models\upscale\_models:
  - **4x-UltraSharp.pth**

**B4. Sử dụng:** Tải ảnh có sẵn lên, nhập prompt và tạo ảnh  
Hướng dẫn chi tiết tại video: <https://youtu.be/fgQrf1JtiXw>

### 19.7. Ghép ảnh cho bé

Ví dụ: Ghép hình em bé vào trong vỏ trứng



### B1. Chuẩn bị tài nguyên

1. Hình mẫu em bé:



2. Hình nền bối cảnh: Có thể tải hình có sẵn trên Internet, hoặc tạo ra từ Stable Diffusion



3. Checkpoint: majicmixRealistic\_v7.safetensor  
([https://huggingface.co/GreenGrape/231209/blob/045ebfc504c47ba8ccc424f1869c65a223d1f5cc/majicmixRealistic\\_v7.safetensors](https://huggingface.co/GreenGrape/231209/blob/045ebfc504c47ba8ccc424f1869c65a223d1f5cc/majicmixRealistic_v7.safetensors))
4. Lora: **HAOTU\_20240508045444**  
<https://drive.google.com/drive/folders/1SDfkGlFAtuzG6QO9zbEq-VVK5ISabB3x?usp=sharing>
5. Controlnet: control\_v11p\_sd15\_softedge => Lưu vào thư mục ~\stable-diffusion-webui-forge\models\ControlNet\models\ControlNet\

[https://huggingface.co/llyasviel/ControlNet-v1-1/blob/main/control\\_v11p\\_sd15\\_softedge.pth](https://huggingface.co/llyasviel/ControlNet-v1-1/blob/main/control_v11p_sd15_softedge.pth)

6. IP-Adapter: ip-adapter-plus\_sd15

[https://huggingface.co/h94/IP-Adapter/blob/main/models/ip-adapter-plus\\_sd15.safetensors](https://huggingface.co/h94/IP-Adapter/blob/main/models/ip-adapter-plus_sd15.safetensors)

## B2. Tạo khuôn hình, tạo Mask



## B3. Xử lý trên Stable Diffusion

Sử dụng Img2Img > Inpaint Upload + Control net: SoftEdge + IP Adapter

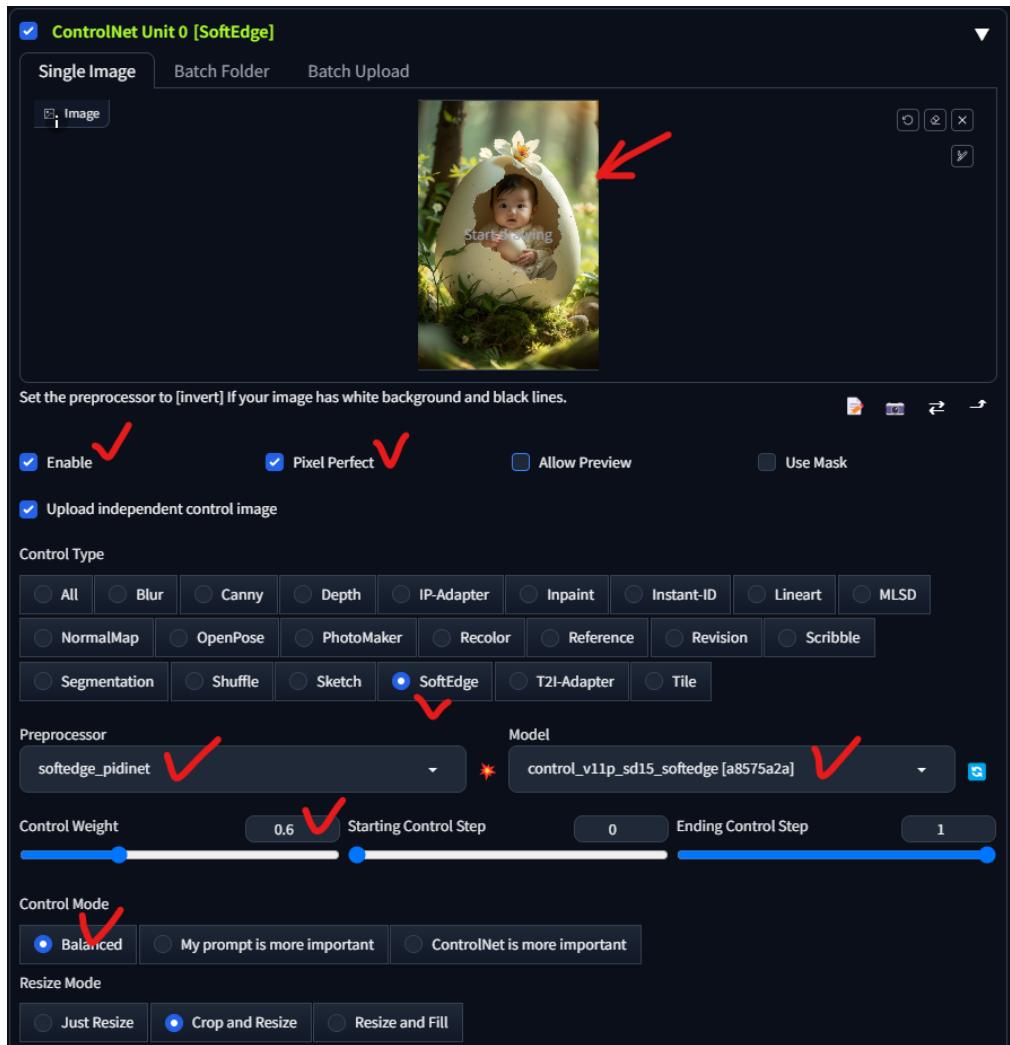
**Prompt:**

HAOTU, flower,panorama, a baby that breaks through an egg, forest, masterpiece, best quality, finely detail, highres, 8k,beautiful and aesthetic <lora:HAOTU\_20240508045444:0.8>

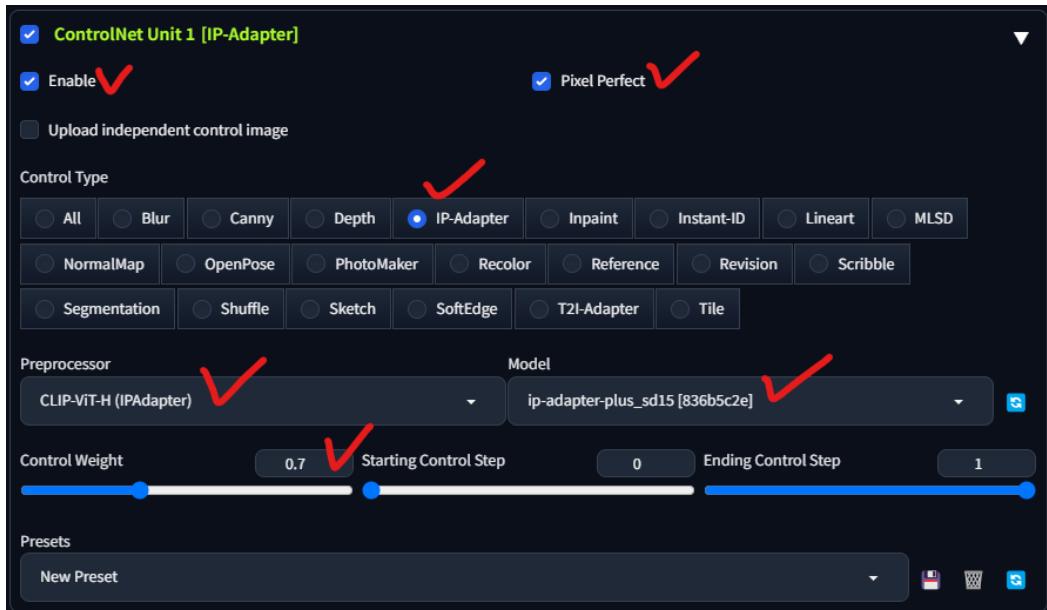
**Negative prompt:**

sketches, (worst quality:2), (low quality:2), (normal quality:2), lowers, normal quality, ((monochrome)), ((grayscale)), facing away, looking away, text, error ,extra digit, fewer digits, cropped,jpeg artifacts,signature, watermark, username,blurry, skin spots, acnes, skin blemishes, bad anatomy,fat,bad feet,cropped,poorly drawn hands,poorly drawn face,mutation,deformed, tilted head, bad anatomy,bad hands,extra fingers,fewer digits,extra limbs,extra arms,extra legs,malformed limbs,fused fingers, too many fingers,long neck,cross-eyed,mutated hands,bad body,bad proportions,gross proportions,text,error,missing fingers, missing arms,missing legs,extra digit, extra arms, extra leg, extra foot,missing fingers

## Cầu hình Controlnet 1



## Cấu hình Controlnet 2



Hướng dẫn chi tiết tại video: <https://youtu.be/TVZb3mb-iJ8>

## **THÔNG TIN LIÊN HỆ**

### **TÀI LIỆU THAM KHẢO**

<https://stable-diffusion-art.com> (nguồn chính)

<https://huggingface.co/hollowstrawberry/stable-diffusion-guide/blob/main/README.md>

<https://civitai.com/articles>