

KHOA KỸ THUẬT VÀ CÔNG NGHỆ
BỘ MÔN CÔNG NGHỆ THÔNG TIN



ĐỒ ÁN CHUYÊN NGÀNH
HỌC KỲ I, NĂM HỌC 2024 - 2025
**NGHIÊN CỨU GFPGAN
VÀ XÂY DỰNG TRANG WEB
PHỤC CHẾ ẢNH BẰNG AI**

Giáo viên hướng dẫn:
TS. Nguyễn Bảo Ân

Sinh viên thực hiện:
Họ tên: Dương Thành Tân
MSSV: 110121097
Lớp: DA21TTB

Trà Vinh, tháng 01 năm 2025

KHOA KỸ THUẬT VÀ CÔNG NGHỆ
BỘ MÔN CÔNG NGHỆ THÔNG TIN



ĐỒ ÁN CHUYÊN NGÀNH
HỌC KỲ I, NĂM HỌC 2024 - 2025
**NGHIÊN CỨU GFPGAN
VÀ XÂY DỰNG TRANG WEB
PHỤC CHẾ ẢNH BẰNG AI**

Giáo viên hướng dẫn:
TS. Nguyễn Bảo Ân

Sinh viên thực hiện:
Họ tên: Dương Thành Tân
MSSV: 110121097
Lớp: DA21TTB

Trà Vinh, tháng 01 năm 2025

NHẬN XÉT CỦA GIÁO VIÊN HƯỚNG DẪN

[illegible]

Trà Vinh, ngày ... tháng ... năm 2025

Giáo viên hướng dẫn
(Ký tên và ghi rõ họ tên)

NHẬN XÉT CỦA THÀNH VIÊN HỘI ĐỒNG

This image shows a full page of white paper with horizontal dotted lines, typical of primary school writing paper. The lines are evenly spaced and run across the width of the page. There are no margins, text, or other markings on the paper.

Trà Vinh, ngày ... tháng ... năm 2025

Thành viên hội đồng
(Ký tên và ghi rõ họ tên)

LỜI CẢM ƠN

Trước hết, tôi xin gửi lời cảm ơn chân thành đến thầy Nguyễn Bảo Ân, giảng viên hướng dẫn, người đã tận tình hỗ trợ, định hướng và góp ý trong suốt quá trình thực hiện đề tài. Những hướng dẫn chi tiết của thầy, đặc biệt trong việc tiếp cận mô hình GFPGAN và triển khai thuật toán xử lý hình ảnh bằng trí tuệ nhân tạo, đã giúp tôi hiểu rõ hơn về các nguyên lý công nghệ và hoàn thành được các mục tiêu của đề tài.

Tôi cũng xin gửi lời cảm ơn đến Khoa Kỹ thuật và Công Nghệ, Trường Đại học Trà Vinh, đã cung cấp kiến thức nền tảng vững chắc về xử lý ảnh số, trí tuệ nhân tạo và các công nghệ liên quan, góp phần giúp tôi có đủ khả năng tiếp cận và ứng dụng GFPGAN trong việc phục chế ảnh cũ một cách hiệu quả. Những tài liệu học thuật và môi trường học tập mà Khoa cung cấp là nguồn hỗ trợ quan trọng trong suốt quá trình nghiên cứu.

Trong quá trình thực hiện đề tài, dù đã nỗ lực hết mình, tôi nhận thức rõ rằng việc nghiên cứu và ứng dụng công nghệ phục chế ảnh bằng AI vẫn còn nhiều thách thức, đòi hỏi sự cải tiến và hoàn thiện không ngừng. Tôi tin rằng, với sự góp ý và hướng dẫn từ quý thầy cô, tôi sẽ ngày càng hoàn thiện hơn trong những bài báo cáo sắp tới.

Tôi xin chân thành cảm ơn!

MỤC LỤC

LỜI CẢM ƠN	3
MỤC LỤC	4
DANH MỤC HÌNH	7
DANH MỤC BẢNG	8
DANH MỤC TỪ VIẾT TẮT	9
MỞ ĐẦU.....	11
CHƯƠNG 1: TỔNG QUAN.....	14
CHƯƠNG 2: NGHIÊN CỨU LÝ THUYẾT.....	15
2.1. Tổng quan về GFPGAN.....	15
2.1.1. Giới thiệu về GFPGAN	15
2.1.2. Cấu trúc mạng GAN trong GFPGAN	15
2.1.3. Ứng dụng thực tế của GFPGAN	16
2.1.4. Ưu điểm và hạn chế của GFPGAN	17
2.1.5. So sánh GFPGAN với các phương pháp phục chế khác.....	18
2.2. Tổng quan về Real-ESRGAN	19
2.2.1. Giới thiệu về Real-ESRGAN	19
2.2.2. Ứng dụng của Real-ESRGAN trong xử lý ảnh	19
2.2.3. Ưu điểm và hạn chế của Real-ESRGAN.....	20
2.3. Tổng quan về LaMa	21
2.3.1. Giới thiệu về LaMa.....	21
2.3.2. Cơ chế hoạt động Masked Autoencoder.....	22
2.3.3. Ứng dụng của LaMa trong chỉnh sửa ảnh	22
2.3.4. Ưu điểm và hạn chế của LaMa.....	23
2.4. Tổng quan về Stable Diffusion	24
2.4.1. Giới thiệu về Stable Diffusion.....	24
2.4.2. Ứng dụng của Stable Diffusion trong sáng tạo nội dung	24
2.4.3. Ưu điểm và hạn chế của Stable Diffusion	25
2.5. Tổng quan về KLing AI.....	26
2.5.1. Giới thiệu về KLing AI.....	26
2.5.3. Ứng dụng KLing AI trong sáng tạo video	26
2.5.4. Ưu điểm và hạn chế của KLing AI.....	27

2.6. Tổng quan về Flask.....	28
2.6.1. Giới thiệu về Flask	28
2.6.2. Cấu trúc ứng dụng Flask và nguyên lý hoạt động	28
2.6.3. Ứng dụng Flask trong phát triển web	29
2.7. Tổng quan về SQLAlchemy	30
2.7.1. Giới thiệu về SQLAlchemy	30
2.7.2. Ứng dụng SQLAlchemy trong các hệ thống nhỏ gọn	30
2.7.3. Ưu điểm và hạn chế của SQLAlchemy	31
2.8. Tổng quan về PayPal API	32
2.8.1. Giới thiệu PayPal API.....	32
2.8.2. Ứng dụng PayPal API trong các hệ thống thanh toán	32
2.8.3. Ưu điểm và hạn chế của PayPal API	34
CHƯƠNG 3: HIỆN THỰC HÓA NGHIÊN CỨU	35
3.1. Đặt tả ứng dụng	35
3.2. Thiết kế kiến trúc hệ thống	35
3.2.1. Thiết kế kiến trúc tổng quan của hệ thống	35
3.2.2. Thiết kế cơ sở dữ liệu	37
3.3. Xây dựng các Route theo kiến trúc Blueprint	41
3.3.1. Route khởi tạo ứng dụng	41
3.3.2. Route quản lý người dùng	42
3.3.3. Route quản trị viên	43
3.3.4. Route phục chế ảnh chân dung.....	43
3.3.5. Route làm nét và phóng to ảnh.....	44
3.3.6. Route xóa đối tượng trong ảnh	45
3.3.7. Route tạo ảnh từ văn bản	45
3.3.8. Route chuyển đổi ảnh thành video	45
3.3.9. Route giao diện chính và xử lý trang chủ.....	46
3.3.10. Route thanh toán.....	46
CHƯƠNG 4: KẾT QUẢ NGHIÊN CỨU	48
4.1. Dữ liệu thử nghiệm	48
4.1. Dữ liệu thử nghiệm với chức năng phục chế ảnh chân dung	48
4.2. Dữ liệu thử nghiệm với chức năng xóa đối tượng trong ảnh	48

4.3. Dữ liệu thử nghiệm với chức năng tạo ảnh từ văn bản	48
4.2. Kết quả thực nghiệm hệ thống	49
4.2.1. Chức năng đăng nhập	49
4.2.2. Chức năng đăng ký	49
4.2.3. Chức năng chức năng phục chế ảnh chân dung	49
4.2.4. Chức năng làm nét và phóng to ảnh	50
4.2.5. Chức năng xóa đối tượng trong ảnh	50
4.2.6. Chức năng tạo ảnh từ văn bản	50
CHƯƠNG 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN	51
5.1. Kết luận	51
5.2. Hạn chế	51
5.3. Phương hướng phát triển.....	52
DANH MỤC TÀI LIỆU THAM KHẢO	53

DANH MỤC HÌNH

Hình 2. 1. Minh họa ảnh phục chế thông qua GFPGAN [1].....	15
Hình 2. 2. Minh họa ảnh phục chế thông qua Real-ESRGAN [2]	19
Hình 2. 3. Minh họa luồng hoạt động của Lama [4]	22
Hình 2. 4. Minh họa hình ảnh do Stable Diffusion tạo [5].....	24
Hình 2. 5. Minh họa quy trình chuyển động hình ảnh thông qua KlingAI [6].....	26
Hình 3. 1. Minh họa mô hình ERD	37
Hình 4. 1. Minh họa kết quả phục chế ảnh chân dung bằng GFPGAN	48
Hình 4. 2. Minh họa kết quả chức năng xóa đối tượng trong ảnh.....	48
Hình 4. 3. Minh họa kết quả tính năng tạo ảnh từ văn bản	48
Hình 4. 4. Minh họa giao diện đăng nhập	49
Hình 4. 5. Minh họa giao diện đăng ký	49
Hình 4. 6. Minh họa giao diện chức năng phục chế ảnh	49
Hình 4. 7. Minh họa giao diện chức năng làm nét ảnh.....	50
Hình 4. 8. Minh họa giao diện chức năng xóa đối tượng trong ảnh.....	50
Hình 4. 9. Minh họa giao diện chức năng tạo ảnh từ văn bản.....	50

DANH MỤC BẢNG

Bảng 2. 1. Bảng so sánh giữa GFPGAN với Photoshop thủ công	18
Bảng 3. 1. Bảng cơ sở dữ liệu thông tin người dùng.....	37
Bảng 3. 2. Bảng cơ sở dữ liệu thông tin lịch sử sử dụng.....	38
Bảng 3. 3. Bảng cơ sở dữ liệu ảnh xử lý	38
Bảng 3. 4. Bảng cơ sở dữ liệu chuyển ảnh thành video	39
Bảng 3. 5. Bảng cơ sở dữ liệu thanh toán.....	40
Bảng 3. 6. Bảng cơ sở dữ liệu gói thanh toán.....	40
Bảng 3. 7. Bảng cơ sở dữ liệu liên kết.....	40

DANH MỤC TỪ VIẾT TẮT

STT	KÝ HIỆU VIẾT TẮT	NỘI DUNG VIẾT TẮT
1	AI	Artificial Intelligence
2	API	Application Programming Interface
3	CSDL	Cơ sở dữ liệu
4	CDN	Content Delivery Network
5	ESRGAN	Enhanced Super-Resolution GAN
6	GFPGAN	Generative Facial Prior GAN
7	LaMa	Large Mask Inpainting
8	Real-ESRGAN	Real-Enhanced Super-Resolution GAN
9	SQL	Structured Query Language

TÓM TẮT ĐỒ ÁN CHUYÊN NGÀNH

Trong bối cảnh công nghệ trí tuệ nhân tạo (AI) ngày càng phát triển và chứng minh vai trò quan trọng trong xử lý dữ liệu trực quan, việc phục chế ảnh cũ, cải thiện chất lượng hình ảnh và tạo nội dung trực quan đã trở thành một nhu cầu thực tế. Đồ án chuyên ngành này tập trung vào việc nghiên cứu và ứng dụng các công nghệ AI tiên tiến, cụ thể là GFPGAN, Real-ESRGAN, LaMa, Stable Diffusion và K Ling AI, để xây dựng một nền tảng web tích hợp giúp phục chế ảnh chân dung, làm nét và phóng to ảnh, xóa đối tượng không mong muốn, tạo ảnh từ văn bản và chuyển đổi ảnh thành video. Nền tảng này hướng tới đối tượng người dùng phổ thông, với giao diện thân thiện, dễ sử dụng, đồng thời cho phép truy cập các chức năng nâng cao qua hệ thống thanh toán PayPal API.

Quá trình xây dựng hệ thống được thực hiện theo mô hình Client-Server với framework Flask, trong đó backend đảm nhiệm xử lý logic nghiệp vụ và tương tác với cơ sở dữ liệu SQLite, còn frontend đảm bảo hiển thị giao diện và xử lý các yêu cầu từ phía người dùng. Hệ thống sử dụng kiến trúc modular design với các route (Blueprint) riêng biệt cho từng chức năng, giúp dễ dàng bảo trì và mở rộng. Các thuật toán AI như GFPGAN và Real-ESRGAN được tích hợp trực tiếp trong backend để đảm bảo tốc độ xử lý tối ưu, trong khi các mô hình như Stable Diffusion và K Ling AI được kết nối qua API bên thứ ba để thực hiện các tác vụ chuyên sâu hơn.

Kết quả của đề tài là một nền tảng web hoàn chỉnh, có khả năng xử lý và cải thiện chất lượng hình ảnh dựa trên AI một cách tự động, nhanh chóng và chính xác. Đề tài không chỉ dừng lại ở việc cung cấp giải pháp xử lý hình ảnh mà còn hướng tới khả năng thương mại hóa với hệ thống quản lý người dùng, gói tin dụng và thanh toán trực tuyến. Tiềm năng mở rộng của dự án bao gồm việc triển khai trên môi trường đám mây, cải tiến mô hình xử lý và hỗ trợ thêm nhiều tính năng nâng cao khác trong tương lai.

MỞ ĐẦU

1. Lý do chọn đề tài

Trong thời đại công nghệ phát triển mạnh mẽ, trí tuệ nhân tạo (AI) đang dần thay đổi cách con người tương tác với dữ liệu, đặc biệt trong lĩnh vực xử lý hình ảnh. Nhu cầu phục chế, cải thiện chất lượng ảnh cũ hay tạo nội dung hình ảnh chất lượng cao ngày càng tăng, không chỉ phục vụ cho mục đích cá nhân mà còn có giá trị quan trọng trong việc bảo tồn tư liệu lịch sử, nghệ thuật và lưu trữ kỹ thuật số. Các công nghệ AI tiên tiến như GFPGAN giúp phục chế ảnh chân dung cũ, Real-ESRGAN làm nét ảnh, LaMa xóa đối tượng không mong muốn và Stable Diffusion tạo ảnh từ văn bản đã chứng minh được khả năng vượt trội trong lĩnh vực này.

Xuất phát từ tiềm năng ứng dụng thực tiễn đó, đề tài "Nghiên cứu GFPGAN và xây dựng trang web phục chế ảnh bằng AI" được lựa chọn với mục tiêu không chỉ tập trung vào việc nghiên cứu chi tiết thuật toán GFPGAN, mà còn ứng dụng tổng hợp các công nghệ AI trong việc cải thiện chất lượng hình ảnh. Đề tài hướng đến việc phát triển một nền tảng web tích hợp các chức năng như phục chế ảnh chân dung, làm nét ảnh, xóa đối tượng, tạo ảnh từ văn bản và chuyển ảnh thành video. Đồng thời, hệ thống sẽ kết hợp giải pháp thanh toán tín dụng thông qua PayPal API, cho phép người dùng nạp tín dụng và sử dụng các tính năng nâng cao. Việc lựa chọn đề tài này không chỉ mang tính học thuật trong việc tìm hiểu và triển khai các mô hình AI, mà còn có tính ứng dụng cao, phù hợp với các yêu cầu thực tế về lưu giữ kỷ niệm, bảo tồn văn hóa số và sáng tạo nội dung hiện đại.

2. Mục tiêu nghiên cứu

Mục tiêu chính của đề tài là nghiên cứu và ứng dụng công nghệ trí tuệ nhân tạo để xây dựng một hệ thống web cho phép phục chế ảnh chân dung và xử lý ảnh chất lượng cao một cách tự động, hiệu quả. Cụ thể, các mục tiêu nghiên cứu bao gồm:

- **Nghiên cứu chuyên sâu về GFPGAN:** Tìm hiểu nguyên lý hoạt động của GFPGAN trong việc phục chế ảnh chân dung cũ.
- **Xây dựng hệ thống web:** Thiết kế và phát triển một nền tảng web hoàn chỉnh, thân thiện với người dùng, hỗ trợ tải ảnh lên và xử lý trực tiếp trên nền tảng.
- **Tích hợp đa chức năng AI:** Kết hợp các công cụ GFPGAN, Real-ESRGAN, LaMa, Stable Diffusion, và Kling AI API để cung cấp các tính năng:
 - Phục chế ảnh chân dung.

- Làm nét và phóng to ảnh.
- Xóa đối tượng trong ảnh.
- Tạo ảnh từ văn bản.
- Chuyển đổi ảnh thành video.
- **Xây dựng hệ thống quản lý người dùng:** Triển khai chức năng đăng ký, đăng nhập và quản lý tài khoản người dùng.
- **Tích hợp thanh toán trực tuyến:** Sử dụng PayPal API cho phép nạp tín dụng và thanh toán khi sử dụng các dịch vụ cao cấp.
- **Thiết kế và tối ưu cơ sở dữ liệu:** Xây dựng CSDL SQLite theo chuẩn hóa, hỗ trợ lưu trữ thông tin người dùng, lịch sử sử dụng và quản lý giao dịch.

3. Phương pháp nghiên cứu

Nghiên cứu lý thuyết: Tìm hiểu các thuật toán và mô hình trí tuệ nhân tạo trong xử lý ảnh, tập trung vào GFPGAN, Real-ESRGAN, LaMa, Stable Diffusion và cách thức hoạt động của mạng đối sinh GAN (Generative Adversarial Network).

Phân tích và tổng hợp tài liệu: Tham khảo các tài liệu chuyên ngành, bài báo khoa học, mã nguồn mở và tài liệu chính thức từ các thư viện AI để nắm rõ nguyên lý hoạt động và khả năng ứng dụng.

Thiết kế và cài đặt hệ thống: Xây dựng kiến trúc hệ thống web bao gồm cả frontend và backend, áp dụng Flask cho backend, tích hợp các mô hình AI và triển khai cơ sở dữ liệu SQLite để lưu trữ thông tin người dùng và lịch sử sử dụng.

Thực nghiệm và kiểm thử: Tiến hành chạy thử các mô hình AI với hình ảnh thực tế, đánh giá chất lượng đầu ra dựa trên độ sắc nét, độ chi tiết và thời gian xử lý.

Thanh toán: Tích hợp PayPal API cho chức năng thanh toán, đảm bảo tính bảo mật và hiệu quả khi người dùng thực hiện giao dịch trên hệ thống.

4. Đối tượng nghiên cứu

Đối tượng nghiên cứu của đề tài tập trung vào các công nghệ và giải pháp phục vụ cho việc xử lý và cải thiện chất lượng hình ảnh thông qua các công cụ trí tuệ nhân tạo (AI). Cụ thể bao gồm:

- **Các mô hình xử lý ảnh AI:** Tìm hiểu các mô hình như GFPGAN, Real-ESRGAN, LaMa, Stable Diffusion và cách chúng được áp dụng để phục chế, làm nét, xóa đối tượng và tạo ảnh từ văn bản.

- **Hệ thống web:** Phân tích và nghiên cứu cách xây dựng một nền tảng web tích hợp đầy đủ tính năng, từ việc tải ảnh, xử lý trực tiếp trên server đến hiển thị kết quả cho người dùng.
- **Quản lý người dùng và giao dịch:** Nghiên cứu các phương pháp quản lý tài khoản người dùng, hệ thống tín dụng và thanh toán trực tuyến thông qua PayPal API.
- **Cơ sở dữ liệu:** Tìm hiểu về SQLite và phương pháp thiết kế cơ sở dữ liệu theo mô hình chuẩn hóa để lưu trữ thông tin người dùng, lịch sử sử dụng và giao dịch.

Đối tượng nghiên cứu được lựa chọn nhằm đảm bảo hệ thống có khả năng ứng dụng thực tiễn, vận hành ổn định, bảo mật và thân thiện với người dùng.

5. Phạm vi nghiên cứu

Đề tài tập trung vào việc nghiên cứu và ứng dụng các mô hình xử lý ảnh AI như GFPGAN, Real-ESRGAN, LaMa và Stable Diffusion để thực hiện các tác vụ phục chế, làm nét, xóa đối tượng và tạo ảnh từ văn bản trên ảnh tĩnh. Hệ thống được xây dựng dưới dạng web application sử dụng Flask, lưu trữ dữ liệu bằng SQLite và triển khai các giao dịch nạp tín dụng qua PayPal API. Quá trình thử nghiệm được thực hiện trên môi trường cục bộ, tập trung vào dữ liệu ảnh số định dạng phổ biến như .jpg, .png.

CHƯƠNG 1: TỔNG QUAN

Trong bối cảnh công nghệ số phát triển mạnh mẽ, nhu cầu phục chế, cải thiện chất lượng ảnh cũ và tạo nội dung hình ảnh chất lượng cao ngày càng tăng. Những bức ảnh cũ, mờ nhòe hoặc ảnh tư liệu lịch sử khi phục chế không chỉ mang giá trị thẩm mỹ mà còn góp phần bảo tồn di sản và lưu trữ những khoảnh khắc quan trọng. Tuy nhiên, việc phục chế thủ công đòi hỏi nhiều kỹ thuật phức tạp, tốn thời gian và không phải ai cũng có thể thực hiện. Điều này đặt ra yêu cầu cần có một giải pháp tự động hóa, ứng dụng các công nghệ AI để xử lý ảnh một cách nhanh chóng và hiệu quả.

Để giải quyết vấn đề trên, đề tài lựa chọn sử dụng các công nghệ trí tuệ nhân tạo và xây dựng một hệ thống web phục chế ảnh bằng AI. Trong đó, GFPGAN là mô hình chính để phục chế ảnh chân dung, kết hợp với các công cụ như Real-ESRGAN để làm nét ảnh, LaMa để xóa đối tượng và Stable Diffusion để tạo ảnh từ văn bản. Hệ thống web được phát triển bằng Flask, kết hợp với cơ sở dữ liệu SQLite để lưu trữ thông tin người dùng và lịch sử sử dụng. Bên cạnh đó, chức năng thanh toán tín dụng được tích hợp thông qua PayPal API, cho phép người dùng nạp tín dụng để sử dụng các dịch vụ cao cấp.

Sau quá trình nghiên cứu và thử nghiệm, hệ thống đã hoàn thành các chức năng cốt lõi bao gồm phục chế ảnh chân dung, làm nét ảnh, xóa đối tượng không mong muốn, tạo ảnh từ văn bản và chuyển đổi ảnh thành video. Hệ thống web có thể triển khai trên các máy chủ cục bộ hoặc đám mây, đảm bảo tính ổn định, tốc độ xử lý hợp lý và giao diện thân thiện với người dùng. Giải pháp này không chỉ phục vụ nhu cầu cá nhân mà còn có thể ứng dụng trong các bảo tàng, lưu trữ di sản số, cũng như hỗ trợ các nhà thiết kế và nhiếp ảnh gia trong sáng tạo nội dung hình ảnh chất lượng cao.

CHƯƠNG 2: NGHIÊN CỨU LÝ THUYẾT

2.1. Tổng quan về GFPGAN

2.1.1. Giới thiệu về GFPGAN

GFPGAN (Generative Facial Prior-Generative Adversarial Network) là một mô hình trí tuệ nhân tạo được thiết kế chuyên biệt cho nhiệm vụ phục chế ảnh chân dung cũ hoặc bị hỏng. Được phát triển bởi nhóm nghiên cứu Tencent ARC, GFPGAN sử dụng mạng đối sinh GAN (Generative Adversarial Network) để khôi phục chi tiết khuôn mặt từ ảnh chân dung có chất lượng thấp, làm mờ hoặc mất nét [1].



Hình 2. 1. Minh họa ảnh phục chế thông qua GFPGAN [1]

Cơ chế cốt lõi của GFPGAN là sử dụng mô hình GAN cải tiến, trong đó kết hợp một mạng Encoder và một bộ giải mã (Decoder) cùng với một Facial Prior Module để khôi phục chi tiết khuôn mặt tự nhiên và chân thực. Đặc biệt, GFPGAN có khả năng tái tạo chi tiết khuôn mặt như mắt, mũi, miệng mà không làm biến dạng khuôn mặt gốc, một bước tiến vượt bậc so với các kỹ thuật phục chế truyền thống.

GFPGAN được giới thiệu vào năm 2021 và nhanh chóng trở thành một công cụ mã nguồn mở phổ biến trong cộng đồng AI và xử lý ảnh. Nó đã được ứng dụng rộng rãi trong các dự án phục chế ảnh di sản, khôi phục ảnh tư liệu lịch sử và cả các ứng dụng sáng tạo nghệ thuật [1].

2.1.2. Cấu trúc mạng GAN trong GFPGAN

GFPGAN (Generative Facial Prior - Generative Adversarial Network) được xây dựng dựa trên cấu trúc Generative Adversarial Network (GAN) – một mô hình học sâu đặc biệt hiệu quả trong việc tạo ra dữ liệu mới từ dữ liệu huấn luyện. GAN bao gồm hai thành phần chính:

- **Bộ sinh (Generator):** Đảm nhận nhiệm vụ tạo ra ảnh phục chế từ ảnh đầu vào chất lượng kém.
- **Bộ phân biệt (Discriminator):** Đóng vai trò như một “giám khảo”, nhằm xác định xem ảnh được tạo ra từ bộ sinh có chân thực và giống với dữ liệu gốc hay không.

Cấu trúc cụ thể trong GFPGAN: GFPGAN đã cải tiến kiến trúc GAN truyền thống bằng việc kết hợp thêm một thành phần mới gọi là Facial Prior Module (FPM), giúp cải thiện đáng kể chất lượng phục chế ảnh chân dung.

Encoder (Bộ mã hóa): Tiếp nhận hình ảnh chất lượng thấp và trích xuất các đặc trưng chính từ ảnh đầu vào. Sau đó, mã hóa các đặc điểm quan trọng như cấu trúc khuôn mặt, đường nét chính và tỷ lệ khuôn mặt.

Facial Prior Module (FPM): Là thành phần đặc biệt trong GFPGAN giúp cải thiện chất lượng phục chế. Được huấn luyện trước (pre-trained) với dữ liệu khuôn mặt lớn để bổ sung thông tin chi tiết bị mất trong ảnh gốc. Thay vì chỉ dựa vào ảnh đầu vào, FPM mang lại khả năng khôi phục chân thực bằng cách thêm thông tin đặc trưng khuôn mặt từ dữ liệu học sâu.

Decoder (Bộ giải mã): Tái tạo lại ảnh chất lượng cao từ các đặc trưng trích xuất bởi Encoder và FPM. Đầu ra là ảnh chân dung phục chế với chi tiết và độ phân giải cao hơn ảnh gốc.

Discriminator (Bộ phân biệt): Được huấn luyện song song với Generator để đảm bảo ảnh tạo ra trông tự nhiên và chân thực nhất. Dùng tổn thất Adversarial Loss để hướng dẫn Generator cải tiến liên tục [1].

2.1.3. Ứng dụng thực tế của GFPGAN

GFPGAN không chỉ là một công cụ phục chế ảnh chân dung đơn thuần mà còn có tính ứng dụng rộng rãi trong nhiều lĩnh vực thực tế, từ việc bảo tồn di sản, sáng tạo nghệ thuật đến hỗ trợ các ngành công nghiệp kỹ thuật số.

Phục chế ảnh cũ và ảnh lịch sử

- GFPGAN đặc biệt hữu ích trong việc phục hồi ảnh chân dung cũ, mờ, hư hỏng do thời gian.
- Được sử dụng trong các bảo tàng, viện lưu trữ để khôi phục hình ảnh tư liệu lịch sử, mang lại độ chi tiết cao và giữ được đặc trưng gốc của ảnh.

Nâng cấp chất lượng ảnh chân dung trong nhiếp ảnh

- Hỗ trợ nhiếp ảnh gia cải thiện chất lượng ảnh chân dung cũ hoặc bị mờ khi chụp thiếu sáng.
- Ứng dụng trong việc cải thiện chất lượng ảnh chân dung kỹ thuật số sau khi chụp bằng điện thoại hoặc camera chất lượng thấp.

Sử dụng trong ngành công nghiệp sáng tạo nội dung và giải trí

- Phục vụ cho các nhà thiết kế đồ họa, nhà làm phim khi cần phục chế hình ảnh tư liệu cũ.
- Ứng dụng trong chỉnh sửa hậu kỳ phim ảnh khi cần tái tạo các cảnh quay hoặc nhân vật cũ.

Tích hợp trong các nền tảng chỉnh sửa ảnh trực tuyến

- Được sử dụng trong các nền tảng như Let's Enhance, Remini, giúp người dùng phục chế ảnh trực tuyến một cách dễ dàng mà không cần kiến thức chuyên môn về AI.
- Khả năng xử lý thời gian thực và dễ triển khai trên các ứng dụng web.

Ứng dụng trong nghiên cứu khoa học và pháp y

- Sử dụng trong khoa học pháp y để phục hồi và tái tạo khuôn mặt từ ảnh cũ hoặc ảnh nhận diện.
- Hỗ trợ các chuyên gia y tế trong việc xử lý ảnh y khoa có độ phân giải thấp.

2.1.4. Ưu điểm và hạn chế của GFPGAN

GFPGAN là một trong những mô hình AI hàng đầu trong việc phục chế ảnh chân dung, nhưng vẫn tồn tại một số điểm cần cải tiến khi so sánh với các giải pháp khác. Dưới đây là phân tích chi tiết về **ưu điểm** và **hạn chế** của GFPGAN.

Ưu điểm:

Phục chế ảnh chân dung với chất lượng cao: GFPGAN có khả năng phục hồi chi tiết khuôn mặt rõ ràng và sắc nét ngay cả khi ảnh gốc có chất lượng rất kém. Tái tạo chính xác các đặc điểm khuôn mặt như mắt, mũi, miệng mà không làm biến dạng tổng thể khuôn mặt.

Tốc độ xử lý nhanh và tự động hóa cao: Quá trình xử lý ảnh diễn ra tự động hoàn toàn mà không cần chỉnh sửa thủ công. Thời gian xử lý nhanh, phù hợp cho các ứng dụng xử lý thời gian thực.

Tính ứng dụng rộng rãi: Áp dụng hiệu quả trong bảo tồn di sản, phục chế ảnh cũ, ảnh lịch sử. Dễ dàng tích hợp vào các nền tảng web và phần mềm chỉnh sửa ảnh như Remini, Let's Enhance.

Khả năng mã nguồn mở và dễ triển khai: GFPGAN là một dự án mã nguồn mở (open-source) trên GitHub, giúp các nhà phát triển dễ dàng tích hợp và tùy chỉnh theo nhu cầu. Không yêu cầu phần cứng cao cấp để triển khai cục bộ, có thể chạy trên GPU thông thường.

Hạn chế:

Chỉ tối ưu cho ảnh chân dung: GFPGAN chỉ tập trung vào phục chế khuôn mặt, không phù hợp cho ảnh phong cảnh, ảnh vật thể hoặc các loại ảnh phi chân dung.

Hạn chế trong ảnh cực mờ hoặc mất nét nghiêm trọng: Đối với ảnh bị hư hỏng nặng hoặc quá mờ, mô hình có thể dự đoán sai các đặc điểm khuôn mặt, dẫn đến kết quả không chính xác.

Không thể xử lý nền và chi tiết phụ xung quanh: GFPGAN chỉ tập trung vào phần khuôn mặt, không xử lý phần nền hoặc chi tiết phụ như quần áo, đồ vật trong ảnh.

Cần điều chỉnh tham số khi xử lý ảnh phức tạp: Khi sử dụng cho các ảnh có nhiều người hoặc khuôn mặt bị che khuất, kết quả có thể cần được tinh chỉnh thủ công.

2.1.5. So sánh GFPGAN với các phương pháp phục chế khác

GFPGAN là một trong những mô hình nổi bật trong lĩnh vực phục chế ảnh chân dung. Tuy nhiên, để hiểu rõ hơn về tính hiệu quả, cần so sánh với một số phương pháp phục chế ảnh phổ biến khác như ESRGAN, DFDNet, và Photoshop (thủ công). Dưới đây là bảng so sánh dựa trên các tiêu chí quan trọng.

Bảng 2. 1. Bảng so sánh giữa GFPGAN với Photoshop thủ công

Tiêu chí	GFPGAN	Photoshop (Thủ công)
Mục đích sử dụng	Phục chế ảnh chân dung	Phục chế từng chi tiết
Loại dữ liệu phù hợp	Ảnh chân dung cũ, ảnh mờ	Tất cả các loại ảnh
Công nghệ sử dụng	GAN với Facial Prior Module	Chỉnh sửa thủ công
Tự động hóa	Hoàn toàn tự động	Chỉnh tay hoàn toàn
Tốc độ xử lý	Nhanh (thời gian thực)	Chậm
Độ chính xác chi tiết	Cao đối với ảnh chân dung	Phụ thuộc tay nghề

Khả năng phục hồi	Chỉ tập trung vào khuôn mặt	Phục hồi theo yêu cầu
Khả năng tùy chỉnh	Giới hạn trong mô hình	Tùy chỉnh hoàn toàn
Khả năng triển khai Web	Dễ dàng triển khai web	Không triển khai web

2.2. Tổng quan về Real-ESRGAN

2.2.1. Giới thiệu về Real-ESRGAN

Real-ESRGAN (Enhanced Super-Resolution Generative Adversarial Network) là một mô hình học sâu dựa trên ESRGAN, được cải tiến và mở rộng cho nhiệm vụ nâng cấp chất lượng ảnh (Image Super-Resolution). Được phát triển bởi Xintao Wang và cộng sự, Real-ESRGAN được thiết kế để làm nét và phóng to ảnh kỹ thuật số mà vẫn giữ nguyên chi tiết, thậm chí khôi phục các phần thông tin bị mờ hoặc mất nét [2].



Hình 2. 2. Minh họa ảnh phục chế thông qua Real-ESRGAN [2]

Real-ESRGAN sử dụng một biến thể của GAN, trong đó:

- **Bộ sinh (Generator):** Tái tạo ảnh có độ phân giải cao từ ảnh đầu vào có chất lượng thấp.
- **Bộ phân biệt (Discriminator):** Đánh giá xem ảnh tái tạo có chân thực như ảnh gốc hay không.
- **Feature Extraction Network:** Dựa trên VGG19 để trích xuất đặc trưng và so sánh chi tiết.

Điểm nổi bật của Real-ESRGAN:

- Hỗ trợ ảnh chất lượng thấp hơn so với ESRGAN truyền thống.
- Khôi phục chi tiết chân thực hơn nhờ bộ lọc đa tần số.
- Được huấn luyện với bộ dữ liệu đa dạng, đảm bảo phục chế được nhiều loại ảnh, đồng thời xử lý và làm nét được nhiều khuôn mặt có trong ảnh [3].

2.2.2. Ứng dụng của Real-ESRGAN trong xử lý ảnh

Real-ESRGAN không chỉ dừng lại ở việc làm nét ảnh đơn thuần mà còn có tính ứng dụng rộng rãi trong các lĩnh vực như:

Nâng cấp ảnh chân dung và ảnh cũ:

- Cải thiện độ sắc nét cho ảnh chân dung bị mờ hoặc chất lượng thấp.
- Ứng dụng trong phục chế ảnh gia đình cũ, ảnh thẻ, ảnh kỷ niệm.

Nâng cấp ảnh trong sản phẩm in ấn và thiết kế:

- Tăng độ phân giải ảnh phục vụ cho in ấn poster, banner mà không làm vỡ ảnh.
- Hữu ích cho nhiếp ảnh gia và nhà thiết kế đồ họa trong việc xử lý ảnh thương mại.

Xử lý ảnh trong game và video:

- Dùng để tăng độ phân giải cho các texture cũ trong game và phim.
- Ứng dụng trong việc nâng cấp đồ họa game cũ cho các bản remaster.

Phân tích và phục hồi dữ liệu khoa học:

- Nâng cấp chất lượng ảnh trong hình ảnh y khoa, ảnh vệ tinh.
- Phục vụ cho phân tích pháp y khi cần khôi phục ảnh có chất lượng kém.

2.2.3. Ưu điểm và hạn chế của Real-ESRGAN

Real-ESRGAN là một trong những mô hình học sâu tiên tiến nhất trong lĩnh vực nâng cấp chất lượng ảnh (Super-Resolution). Được xây dựng trên nền tảng của ESRGAN (Enhanced Super-Resolution Generative Adversarial Network) với những cải tiến đáng kể, mô hình này đã chứng minh hiệu quả vượt trội trong việc tăng cường độ phân giải và cải thiện chi tiết hình ảnh. Tuy nhiên, như bất kỳ công nghệ nào, Real-ESRGAN cũng tồn tại những ưu điểm nổi bật và hạn chế nhất định cần cân nhắc khi ứng dụng vào thực tế.

Ưu điểm:

Phục hồi chi tiết hình ảnh ấn tượng: Real-ESRGAN có thể tái tạo lại các chi tiết phức tạp trong ảnh như kết cấu da, vải và các hoa văn nhỏ một cách chân thực. Nhờ sự cải tiến từ mạng GAN kết hợp với Feature Matching Loss, hệ thống này có thể giữ nguyên các chi tiết tự nhiên trong khi loại bỏ hiện tượng nhòe và răng cưa khi phóng to ảnh gấp 2 đến 3 lần.

Khả năng xử lý ảnh đa dạng: Không giống các công cụ làm nét truyền thống, Real-ESRGAN không chỉ giới hạn ở ảnh chân dung mà có thể xử lý nhiều loại ảnh khác nhau, từ ảnh phong cảnh, hình minh họa, cho đến texture trong game và dữ liệu y tế. Điều này mở rộng đáng kể tính ứng dụng thực tế của mô hình.

Tự động hóa và dễ triển khai: Real-ESRGAN được thiết kế để hoạt động tự động hoàn toàn, không yêu cầu can thiệp thủ công từ người dùng. Điều này giúp giảm thời gian chỉnh sửa và phù hợp với các hệ thống cần xử lý hàng loạt ảnh. Ngoài ra, mã nguồn mở trên GitHub giúp các nhà phát triển dễ dàng tích hợp vào các ứng dụng web và phần mềm chỉnh sửa ảnh.

Tối ưu cho xử lý thời gian thực: Với sự hỗ trợ tốt trên các card đồ họa hiện đại (GPU), Real-ESRGAN có thể xử lý ảnh trong thời gian thực. Điều này đặc biệt phù hợp với các ứng dụng cần xử lý nhanh, chẳng hạn như phục chế ảnh trực tiếp trên nền tảng web hoặc chỉnh sửa ảnh tự động.

Nhược điểm:

Giới hạn khi xử lý ảnh cực mờ hoặc hư hỏng nặng: Real-ESRGAN hoạt động hiệu quả nhất với ảnh chất lượng trung bình đến thấp. Tuy nhiên, khi đối mặt với những ảnh quá mờ, mất chi tiết nghiêm trọng, kết quả phục hồi có thể không đạt được độ chính xác như mong đợi. Trong một số trường hợp, mô hình có thể phục chế sai lệch chi tiết hoặc thậm chí tạo ra các chi tiết giả mạo.

Không tối ưu cho ảnh chân dung chuyên biệt: Real-ESRGAN được thiết kế cho nâng cấp ảnh tổng quát, do đó nó không tối ưu bằng GFPGAN khi xử lý các ảnh chân dung.

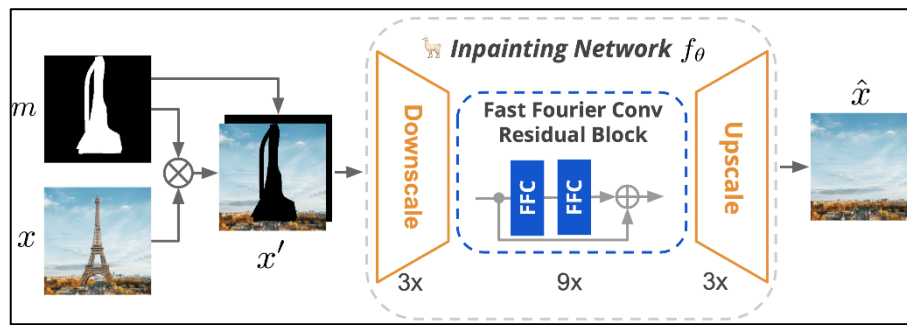
Hạn chế xử lý phong nền và chi tiết phức tạp: Real-ESRGAN chủ yếu tập trung vào phục chế đối tượng chính, nên khi xử lý các chi tiết nền phức tạp hoặc ảnh có nhiều lớp đối tượng chồng chéo, kết quả có thể chưa đạt độ chính xác cao.

Phụ thuộc vào sức mạnh phần cứng: Mặc dù đã được tối ưu hóa, nhưng để đạt hiệu suất xử lý thời gian thực, Real-ESRGAN yêu cầu GPU mạnh mẽ. Điều này có thể là rào cản đối với người dùng phổ thông hoặc các hệ thống không có các card đồ họa chuyên dụng.

2.3. Tổng quan về LaMa

2.3.1. Giới thiệu về LaMa

LaMa (Large Mask Inpainting Model) là một mô hình trí tuệ nhân tạo tiên tiến được thiết kế cho nhiệm vụ xóa đối tượng không mong muốn trong ảnh (Image Inpainting). Được phát triển bởi Samsung AI Center Moscow, LaMa ứng dụng công nghệ Masked Autoencoder (MAE) để phục hồi các vùng bị xóa trong ảnh một cách chân thực và tự nhiên [4].



Hình 2. 3. Minh họa luồng hoạt động của Lama [4]

Khác với các phương pháp truyền thống chỉ dựa trên phép nội suy pixel hoặc sao chép từ các vùng lân cận, LaMa có khả năng học và dự đoán các chi tiết bị thiếu trong ảnh dựa trên ngữ cảnh xung quanh. Điều này giúp hệ thống lấp đầy khoảng trống một cách thông minh, tạo ra các chi tiết mới phù hợp thay vì chỉ làm mờ hoặc nhân bản vùng lân cận [4].

2.3.2. Cơ chế hoạt động Masked Autoencoder

Cơ chế Masked Autoencoder (MAE) là nền tảng cốt lõi của LaMa, hoạt động theo nguyên lý mã hóa và giải mã các phần dữ liệu bị thiếu trong ảnh. MAE là một phiên bản cải tiến từ mạng autoencoder truyền thống với cách tiếp cận đặc biệt trong việc ẩn (masking) dữ liệu.

Quy trình hoạt động:

- **Input (Ảnh đầu vào):** Người dùng tải lên một bức ảnh và đánh dấu vùng cần xóa bằng mặt nạ (mask).
- **Masking (Ẩn vùng cần xóa):** Một phần dữ liệu trong ảnh (vùng bị đánh dấu) sẽ bị che khuất hoàn toàn trong quá trình huấn luyện.
- **Encoding (Mã hóa):** Bộ mã hóa (encoder) trích xuất các đặc trưng từ phần ảnh còn lại.
- **Decoding (Giải mã):** Bộ giải mã (decoder) dựa trên các đặc trưng trích xuất và phần dữ liệu còn lại để dự đoán và tái tạo phần bị che khuất.
- **Output (Ảnh đầu ra):** Tạo ra một phiên bản ảnh mới với các chi tiết đã được tái tạo hợp lý và tự nhiên.

2.3.3. Ứng dụng của LaMa trong chỉnh sửa ảnh

LaMa có tính ứng dụng thực tế cao, đặc biệt trong các nhiệm vụ liên quan đến chỉnh sửa và phục hồi ảnh kỹ thuật số. Một số ứng dụng tiêu biểu bao gồm:

Xóa đối tượng không mong muốn:

- Xóa người thừa trong ảnh chụp.

- Loại bỏ watermark, logo hoặc chữ trong ảnh.
- Xóa các chi tiết thừa trong ảnh nghệ thuật.

Phục hồi ảnh bị hỏng:

- Lấp đầy các vùng ảnh bị mất dữ liệu do hỏng hóc.
- Phục hồi ảnh kỹ thuật số bị lỗi pixel hoặc tệp ảnh hỏng.

Chỉnh sửa hậu kỳ trong nhiếp ảnh và video:

- Hỗ trợ xóa dây điện, biển báo trong ảnh phong cảnh.
- Xóa vật thể gây nhiễu trong video, hỗ trợ các công việc chỉnh sửa hậu kỳ trong phim ảnh.

Tái tạo nội dung sáng tạo:

- Được sử dụng trong sáng tạo nội dung số, giúp mở rộng ảnh, tạo các hiệu ứng nghệ thuật độc đáo.

2.3.4. Ưu điểm và hạn chế của LaMa

Ưu điểm:

- **Khả năng tái tạo chi tiết tốt:** Có thể xóa đối tượng lớn mà vẫn tái tạo chi tiết tự nhiên.
- **Xử lý tự động:** Hệ thống hoàn toàn tự động, không cần chỉnh sửa thủ công.
- **Dễ triển khai:** Mã nguồn mở, dễ dàng tích hợp vào các hệ thống web.
- **Hiệu suất cao:** Áp dụng công nghệ Masked Autoencoder giúp phục hồi các vùng ảnh phức tạp.

Hạn chế:

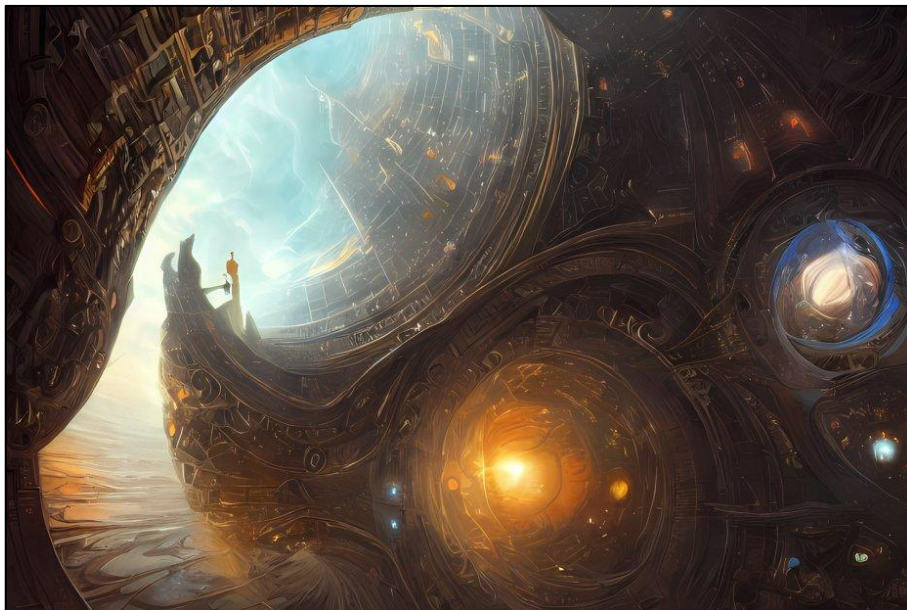
- **Giới hạn đối với chi tiết phức tạp:** Đối với các ảnh có nhiều lớp đối tượng chồng chéo, LaMa có thể tái tạo chưa chính xác.
- **Không tối ưu cho chi tiết nhỏ:** Khi xóa các đối tượng nhỏ như mắt kính, hoa văn, hệ thống có thể không tạo lại đúng chi tiết.
- **Phụ thuộc vào dữ liệu huấn luyện:** Chất lượng ảnh đầu ra phụ thuộc vào mức độ đa dạng của tập dữ liệu huấn luyện.

LaMa là một công cụ mạnh mẽ trong việc xóa đối tượng không mong muốn và phục hồi hình ảnh nhờ công nghệ Masked Autoencoder. Với khả năng tự động hóa cao và dễ triển khai, LaMa phù hợp cho các ứng dụng chỉnh sửa ảnh nhanh, phục vụ cho nhiếp ảnh gia, nhà thiết kế và người dùng phổ thông. Tuy nhiên, để tối ưu với các chi tiết nhỏ hoặc ảnh phức tạp, LaMa cần được kết hợp với các phương pháp xử lý khác.

2.4. Tổng quan về Stable Diffusion

2.4.1. Giới thiệu về Stable Diffusion

Stable Diffusion là một mô hình trí tuệ nhân tạo tiên tiến chuyên sử dụng cho tạo ảnh từ văn bản (Text-to-Image). Được phát triển bởi nhóm CompVis và Stability AI, mô hình này dựa trên nguyên lý Diffusion Models – một phương pháp học sâu tiên tiến cho phép tạo ra hình ảnh từ dữ liệu nhiễu ngẫu nhiên dựa trên một chuỗi các bước giảm nhiễu dần dần [5].



Hình 2. 4. Minh họa hình ảnh do Stable Diffusion tạo [5]

Khác với các mô hình GAN truyền thống, Stable Diffusion hoạt động theo nguyên lý mô hình khuếch tán (Diffusion Models), trong đó một ảnh ban đầu được biểu diễn dưới dạng nhiễu Gaussian ngẫu nhiên, sau đó qua các bước học ngược dần để tái tạo lại hình ảnh chi tiết. Điều này cho phép mô hình sinh ảnh mới từ mô tả văn bản một cách chính xác và đa dạng [5].

Stable Diffusion nổi bật nhờ vào việc có thể chạy cục bộ trên các thiết bị phần cứng tiêu chuẩn mà không cần phụ thuộc vào tài nguyên đám mây quá mạnh mẽ. Đây là một bước tiến lớn so với các đối thủ như DALL·E hay Imagen.

2.4.2. Ứng dụng của Stable Diffusion trong sáng tạo nội dung

Stable Diffusion không chỉ là một công cụ sáng tạo hình ảnh mà còn có ứng dụng đa dạng trong nhiều lĩnh vực thực tế, đặc biệt là trong sáng tạo nội dung và truyền thông.

Sáng tạo nghệ thuật số: Tạo tranh minh họa, concept art, được các nghệ sĩ kỹ thuật số sử dụng để tạo các tác phẩm nghệ thuật sáng tạo.

Tạo nội dung marketing và truyền thông: Thiết kế hình ảnh quảng cáo, tạo các banner, poster và hình ảnh sáng tạo phục vụ chiến dịch marketing, tạo hình ảnh cho truyền thông xã hội.

Hỗ trợ thiết kế và ý tưởng sáng tạo: Thiết kế nhân vật, bối cảnh game, được sử dụng trong ngành công nghiệp game để tạo mẫu nhân vật, bối cảnh và ý tưởng thiết kế. Giúp các nhà thiết kế thử nghiệm các ý tưởng đồ họa nhanh chóng và tiết kiệm thời gian.

Tạo hình ảnh y khoa và dữ liệu khoa học: Mô phỏng hình ảnh y khoa và tạo ra các hình ảnh giả lập để hỗ trợ đào tạo y khoa.

Tùy chỉnh hình ảnh cá nhân: Tạo hình đại diện (AI Avatars), tùy chỉnh ảnh chân dung với phong cách nghệ thuật khác nhau.

2.4.3. Ưu điểm và hạn chế của Stable Diffusion

Ưu điểm:

- **Tạo nội dung sáng tạo từ văn bản:** Biến các mô tả văn bản thành ảnh chi tiết một cách sáng tạo và độc đáo.
- **Khả năng tùy chỉnh cao:** Có thể huấn luyện thêm trên tập dữ liệu riêng để tạo ra các phong cách nghệ thuật độc quyền.
- **Chạy cục bộ:** Không cần GPU quá mạnh, có thể hoạt động độc lập trên máy tính cá nhân.
- **Mã nguồn mở:** Được phát hành mã nguồn mở, giúp cộng đồng nhà phát triển dễ dàng tùy biến và mở rộng.
- **Ứng dụng đa lĩnh vực:** Từ nghệ thuật số đến nghiên cứu khoa học.

Hạn chế:

- **Phụ thuộc vào dữ liệu huấn luyện:** Chất lượng ảnh sinh ra phụ thuộc vào độ phong phú của dữ liệu huấn luyện.
- **Thiếu kiểm soát chi tiết:** Đôi khi các chi tiết nhỏ như tay, mắt có thể bị méo mó hoặc không chính xác.
- **Khó tái tạo kết quả giống nhau:** Do bản chất ngẫu nhiên của Diffusion Models, việc tái tạo chính xác một hình ảnh là rất khó.
- **Giới hạn về độ phân giải:** Phiên bản Stable Diffusion tiêu chuẩn có giới hạn độ phân giải đầu ra, cần tích hợp Real-ESRGAN để nâng cấp hình ảnh.

2.5. Tổng quan về KLing AI

2.5.1. Giới thiệu về KLing AI

KLing AI là một công nghệ tiên tiến trong lĩnh vực chuyển đổi ảnh tĩnh thành video động, ứng dụng thuật toán AI tổng hợp chuyển động để tạo ra các đoạn video từ một hoặc nhiều bức ảnh đầu vào. KLing AI có thể tái hiện chuyển động một cách mượt mà và chân thực, từ đó mang lại hiệu ứng động cho các bức ảnh tĩnh mà không cần đến các nguồn video thực tế [6].



Hình 2. 5. Minh họa quy trình chuyển động hình ảnh thông qua KLingAI [6]

Công nghệ này sử dụng các mạng nơ-ron học sâu (Deep Neural Networks), đặc biệt là các kỹ thuật như Optical Flow Estimation và Frame Interpolation để phân tích và dự đoán chuyển động giữa các khung hình. KLing AI đã nhanh chóng trở nên phổ biến trong sản xuất nội dung số, nghệ thuật kỹ thuật số và phục hồi video lịch sử [6].

2.5.3. Ứng dụng KLing AI trong sáng tạo video

KLing AI không chỉ là một công cụ xử lý ảnh đơn thuần mà còn có nhiều ứng dụng thực tế trong các lĩnh vực sáng tạo nội dung và truyền thông:

Nghệ thuật và sáng tạo nội dung số:

- Tạo các video motion art từ tranh tĩnh.
- Sáng tạo các video visual storytelling cho các dự án nghệ thuật.

Phục hồi ảnh và video lịch sử:

- Biến các ảnh đen trắng cũ thành video có chuyển động mượt mà.
- Tạo các video tài liệu mang tính lịch sử, tái hiện các nhân vật nổi tiếng trong các bảo tàng số hóa.

Quảng cáo và tiếp thị số:

- Tạo video quảng cáo động từ banner tĩnh.

- Tăng cường hiệu ứng thị giác trong các chiến dịch truyền thông trực tuyến.

Giải trí và truyền thông:

- Ứng dụng trong sản xuất video âm nhạc, tạo các hiệu ứng hình ảnh trừu tượng.
- Tạo video mở đầu cho các chương trình truyền hình hoặc video giới thiệu sản phẩm nổi bật.

2.5.4. Ưu điểm và hạn chế của KLint AI

Ưu điểm:

Tự động hóa hoàn toàn: KLint AI được thiết kế với giao diện thân thiện, hoàn toàn tự động hóa quá trình chuyển đổi ảnh thành video mà không yêu cầu người dùng có kiến thức chuyên sâu về chỉnh sửa video. Người dùng chỉ cần tải lên ảnh, chọn một số tùy chỉnh cơ bản (như tốc độ chuyển động, độ phân giải) và hệ thống sẽ thực hiện toàn bộ quá trình còn lại. Điều này đặc biệt hữu ích cho người dùng không chuyên, chẳng hạn như giáo viên, nhà thiết kế nội dung trực tuyến hoặc các nghệ sĩ kỹ thuật số.

Tạo chuyển động mượt mà: Sử dụng AI Motion Estimation và Optical Flow, KLint AI có thể mô phỏng chuyển động một cách mượt mà và chân thực. Công nghệ này giúp hệ thống tạo ra các khung hình trung gian (intermediate frames) giữa các khung hình gốc, từ đó tái hiện chuyển động tự nhiên giữa các đối tượng trong ảnh.

Chất lượng đầu ra cao: Một ưu điểm nổi bật khác là độ phân giải đầu ra của KLint AI có thể đạt tới 4K (3840 x 2160 pixels), phù hợp cho các dự án chuyên nghiệp như phim tài liệu, quảng cáo hoặc video marketing.

Thời gian xử lý nhanh: KLint AI được thiết kế để tận dụng tối đa sức mạnh phần cứng, đặc biệt là GPU (Graphics Processing Unit). Nhờ vào việc tối ưu các mạng nơ-ron học sâu và bộ nhớ đệm, hệ thống có thể xử lý và tạo video trong vài phút, ngay cả trên các thiết bị có cấu hình trung bình.

Tính mã nguồn mở và khả năng mở rộng: KLint AI là một nền tảng mã nguồn mở (open-source) và có thể dễ dàng tích hợp vào các hệ thống web hoặc ứng dụng phần mềm khác.

Hạn chế:

Giới hạn đối với ảnh phức tạp: Mặc dù KLint AI hoạt động tốt với các ảnh chân dung hoặc ảnh đơn giản, nhưng đối với những ảnh có quá nhiều chi tiết phức tạp (như cảnh rừng rậm, đám đông người) hoặc các đối tượng chồng chéo hoặc có nhiều lớp chi tiết. Kết quả đầu ra có thể không chính xác hoặc làm mờ các chi tiết quan trọng.

Phụ thuộc vào dữ liệu đầu vào: Chất lượng video đầu ra phụ thuộc mạnh mẽ vào độ phân giải và độ nét của ảnh đầu vào. Điều này khiến KLint AI không thực sự phù hợp cho các nhiệm vụ phục hồi ảnh cực kỳ hư hỏng mà không có sự can thiệp từ các công cụ nâng cấp khác (như Real-ESRGAN).

Không tối ưu cho chuyển động phức tạp: Không phù hợp để tái tạo các chuyển động vật lý phức tạp như sóng nước hay gió thổi.

2.6. Tổng quan về Flask

2.6.1. Giới thiệu về Flask

Flask là một micro web framework mã nguồn mở được viết bằng ngôn ngữ lập trình Python. Được phát triển bởi Armin Ronacher vào năm 2010, Flask nổi bật nhờ tính nhẹ, linh hoạt và dễ sử dụng, phù hợp cho cả các dự án nhỏ và các ứng dụng web quy mô vừa.

Flask được thiết kế theo triết lý "micro", tức là cung cấp các thành phần cốt lõi để xây dựng ứng dụng web và cho phép lập trình viên tùy chỉnh, mở rộng dựa trên yêu cầu thực tế của dự án. Flask không tích hợp sẵn các tính năng như hệ thống quản lý cơ sở dữ liệu hay xác thực người dùng, nhưng hỗ trợ các thư viện bên thứ ba giúp mở rộng khả năng dễ dàng.

Các tính năng nổi bật của Flask:

- **Đơn giản và gọn nhẹ:** Cấu trúc code rõ ràng, dễ hiểu.
- **Tính mô-đun cao:** Dễ dàng mở rộng bằng các extensions.
- **Hỗ trợ WSGI (Web Server Gateway Interface):** Tích hợp tốt với các máy chủ web như Gunicorn, uWSGI.
- **Khả năng tương thích cao:** Tích hợp tốt với SQLAlchemy, Jinja2 và nhiều thư viện Python khác.

2.6.2. Cấu trúc ứng dụng Flask và nguyên lý hoạt động

Flask tuân theo mô hình Client-Server, trong đó:

- **Client (Người dùng):** Truy cập ứng dụng qua trình duyệt web.
- **Server (Máy chủ Flask):** Xử lý yêu cầu, truy vấn dữ liệu và trả kết quả về cho trình duyệt.

Cấu trúc cơ bản của một ứng dụng Flask:

Router (Định tuyến yêu cầu): Khi người dùng truy cập vào một đường link, Flask sẽ kiểm tra URL đó và gọi hàm xử lý tương ứng.

Controller (Xử lý logic nghiệp vụ): Thực hiện các thao tác như kiểm tra dữ liệu đầu vào, truy vấn CSDL hoặc gọi các mô hình AI.

Template Engine (Giao diện người dùng): Flask sử dụng Jinja2 để kết hợp dữ liệu và giao diện, tạo ra trang web động.

Database (Cơ sở dữ liệu): Flask không tích hợp sẵn quản lý CSDL nhưng thường kết hợp với SQLite hoặc SQLAlchemy để lưu trữ dữ liệu.

Quy trình xử lý cơ bản:

- Người dùng truy cập website.
- Flask xử lý yêu cầu và gọi hàm phù hợp.
- Nếu cần, Flask truy vấn CSDL và xử lý logic.
- Trả kết quả cho người dùng dưới dạng HTML hoặc JSON.

2.6.3. Ứng dụng Flask trong phát triển web

Flask không chỉ là một framework dùng để tạo website đơn giản, mà còn là giải pháp linh hoạt cho nhiều mục đích khác nhau:

Phát triển ứng dụng web động và quản lý nội dung:

- Xây dựng trang web cá nhân, blog và portfolio.
- Phát triển hệ thống quản lý nội dung (CMS) đơn giản cho các doanh nghiệp nhỏ.

Xây dựng API RESTful:

- Flask thường được sử dụng để tạo RESTful API, cho phép các ứng dụng như mobile apps hoặc frontend web giao tiếp với server.
- Các dự án AI như GFPGAN và Stable Diffusion có thể triển khai API để xử lý ảnh và trả kết quả qua JSON.

Ứng dụng trong trí tuệ nhân tạo (AI) và Machine Learning:

- Flask là lựa chọn lý tưởng để triển khai mô hình AI do tính nhẹ và dễ tích hợp.
- Được sử dụng phổ biến trong các dự án xử lý ảnh và video như phục chế ảnh bằng GFPGAN hoặc làm nét ảnh với Real-ESRGAN.

Xây dựng hệ thống quản lý người dùng:

- Tích hợp dễ dàng các tính năng như:
- Đăng ký, đăng nhập.
- Phân quyền người dùng.

- Quản lý tài khoản và bảo mật dữ liệu.

Do cấu trúc đơn giản, Flask phù hợp cho các dự án thử nghiệm hoặc xây dựng nguyên mẫu nhanh chóng. Được các startup và nhà phát triển lựa chọn để thử nghiệm ý tưởng phần mềm trước khi mở rộng quy mô.

2.7. Tổng quan về SQLAlchemy

2.7.1. Giới thiệu về SQLAlchemy

SQLAlchemy là một thư viện ORM (Object Relational Mapping) mạnh mẽ cho ngôn ngữ lập trình Python, được phát triển bởi Michael Bayer vào năm 2006. SQLAlchemy cho phép các nhà phát triển tương tác với cơ sở dữ liệu quan hệ (RDBMS) thông qua đối tượng Python thay vì phải viết trực tiếp các câu lệnh SQL phức tạp.

Khác với việc sử dụng truy vấn SQL thô, SQLAlchemy cung cấp một lớp trừu tượng cao cấp giúp quản lý và thao tác với dữ liệu theo cách trực quan hơn. Các đối tượng trong CSDL như bảng, cột, hàng dữ liệu được ánh xạ trực tiếp thành class và đối tượng Python, từ đó cho phép:

- Tạo, đọc, cập nhật, xóa (CRUD) dữ liệu một cách dễ dàng.
- Tự động ánh xạ (Mapping) giữa các bảng trong cơ sở dữ liệu và đối tượng trong code.
- Quản lý các quan hệ giữa bảng dữ liệu một cách rõ ràng và trực quan.

SQLAlchemy bao gồm hai thành phần chính:

- Core: Cung cấp quyền kiểm soát thấp hơn với SQL thuần.
- ORM (Object Relational Mapping): Cung cấp quyền kiểm soát cấp cao, cho phép làm việc với đối tượng Python thay vì SQL thô.

2.7.2. Ứng dụng SQLAlchemy trong các hệ thống nhỏ gọn

SQLAlchemy là một giải pháp lý tưởng cho các hệ thống nhỏ gọn và vừa, đặc biệt trong các dự án Python, nhờ khả năng tích hợp linh hoạt và dễ sử dụng. Một số ứng dụng tiêu biểu bao gồm:

Quản lý CSDL cho ứng dụng web (Flask/Django):

- SQLAlchemy thường được kết hợp với Flask để quản lý CSDL nhỏ như SQLite.
- Được sử dụng để quản lý người dùng, sản phẩm, giao dịch trong các ứng dụng web cỡ nhỏ.

Ứng dụng quản lý dữ liệu cá nhân:

- Quản lý danh sách công việc, ghi chú cá nhân, quản lý liên hệ.
- Tối ưu cho các ứng dụng desktop và hệ thống offline.

Phát triển hệ thống thử nghiệm và nguyên mẫu:

- SQLAlchemy phù hợp cho giai đoạn thử nghiệm và tạo nguyên mẫu do tính đơn giản trong cấu hình.
- Thường sử dụng với SQLite cho các hệ thống cần ít tài nguyên, dễ triển khai.

Hệ thống học máy và xử lý dữ liệu:

- Tích hợp với các thư viện Pandas, NumPy để xử lý và lưu trữ dữ liệu.
- Được dùng trong các dự án AI và Machine Learning để quản lý dữ liệu huấn luyện quy mô nhỏ.

Quản lý giao dịch và phiên làm việc: SQLAlchemy còn hỗ trợ quản lý phiên làm việc (Session Management) và các giao dịch dữ liệu phức tạp trong các hệ thống quản lý tài chính, thương mại điện tử.

2.7.3. Ưu điểm và hạn chế của SQLAlchemy

Ưu điểm:

Mã nguồn mở và cộng đồng lớn: SQLAlchemy là một thư viện mã nguồn mở, được cộng đồng Python hỗ trợ mạnh mẽ, có tài liệu đầy đủ và dễ hiểu, dễ dàng cho người mới bắt đầu.

Tính trừu tượng và dễ sử dụng: Cung cấp giao diện đối tượng (ORM), cho phép làm việc với dữ liệu dưới dạng đối tượng Python, loại bỏ nhu cầu viết SQL thủ công trong hầu hết các trường hợp.

Tương thích đa CSDL: Hỗ trợ nhiều hệ quản trị CSDL phổ biến như SQLite, PostgreSQL, MySQL, Oracle, Microsoft SQL Server.

Tích hợp tốt với Flask: Được hỗ trợ mặc định trong Flask-SQLAlchemy, đồng thời phù hợp cho cả các hệ thống API RESTful và quản lý dữ liệu web.

Hạn chế:

Hiệu suất thấp hơn SQL thuần: Vì hoạt động trên lớp trừu tượng, SQLAlchemy có thể chậm hơn so với câu lệnh SQL thuần trong các truy vấn phức tạp.

Không phù hợp cho hệ thống lớn và phức tạp: Khi quản lý dữ liệu hàng triệu bản ghi, các ORM-based framework như SQLAlchemy có thể không tối ưu.

Độ phức tạp trong các hệ thống lớn: Khi làm việc với nhiều bảng và mối quan hệ phức tạp, SQLAlchemy có thể trở nên khó bảo trì.

Phụ thuộc vào Python: Chỉ hoạt động với ngôn ngữ Python, không phù hợp cho các hệ thống sử dụng ngôn ngữ khác.

SQLAlchemy là một thư viện ORM mạnh mẽ dành cho Python, đặc biệt phù hợp với các hệ thống vừa và nhỏ nhờ khả năng quản lý dữ liệu linh hoạt và tích hợp dễ dàng với các framework như Flask. Tuy nhiên, đối với các hệ thống lớn và phức tạp, cần cân nhắc việc tối ưu hoặc sử dụng SQL thuần để đảm bảo hiệu suất.

2.8. Tổng quan về PayPal API

2.8.1. Giới thiệu PayPal API

PayPal API là một nền tảng giao tiếp lập trình ứng dụng do PayPal phát triển, cho phép các nhà phát triển tích hợp chức năng thanh toán trực tuyến vào ứng dụng hoặc website của mình. Được ra mắt vào năm 1998, PayPal hiện là một trong những hệ thống thanh toán trực tuyến lớn nhất thế giới với hơn 300 triệu tài khoản đang hoạt động.

PayPal API đóng vai trò như một cầu nối giữa người mua, người bán và hệ thống ngân hàng, giúp thực hiện các giao dịch thanh toán điện tử một cách bảo mật và tiện lợi. Thay vì yêu cầu người dùng nhập thông tin thẻ trực tiếp trên website, hệ thống PayPal cung cấp giao diện bảo mật để xử lý thanh toán bên ngoài, giúp giảm thiểu rủi ro gian lận và truy cập trái phép.

Các tính năng chính của PayPal API:

- **Thanh toán một lần:** Hỗ trợ thanh toán trực tiếp cho sản phẩm hoặc dịch vụ đơn lẻ.
- **Thanh toán định kỳ:** Dành cho các dịch vụ có mô hình subscription.
- **Quản lý giao dịch:** Hỗ trợ kiểm tra trạng thái giao dịch, hủy giao dịch hoặc hoàn tiền.
- **Xử lý nhiều loại tiền tệ:** PayPal hỗ trợ hơn 25 loại tiền tệ trên toàn cầu.
- **Bảo mật cao:** Tích hợp công nghệ mã hóa SSL và chuẩn bảo mật PCI DSS.

2.8.2. Ứng dụng PayPal API trong các hệ thống thanh toán

PayPal API đã trở thành một trong những nền tảng thanh toán trực tuyến phổ biến nhất trên toàn cầu nhờ khả năng bảo mật cao, dễ dàng tích hợp và hỗ trợ đa loại giao dịch. Dưới đây là những ứng dụng tiêu biểu của PayPal API trong các hệ thống thanh toán hiện nay:

Thanh toán cho Thương mại điện tử (E-Commerce)

- Tích hợp để xử lý thanh toán một lần cho sản phẩm hoặc dịch vụ.
- Tương thích nhiều loại tiền tệ, giúp doanh nghiệp mở rộng quy mô ra thị trường quốc tế.

Thanh toán cho các dịch vụ kỹ thuật số và phần mềm SaaS

- **Thanh toán định kỳ:** PayPal API cho phép thiết lập các gói subscription với chu kỳ thanh toán tự động theo tháng hoặc năm.
- **Quản lý hủy gia hạn:** Người dùng có thể hủy thanh toán định kỳ trực tiếp qua giao diện PayPal Dashboard.

Nạp tiền và hệ thống tín dụng (Credit-based Systems)

PayPal API có thể ứng dụng trong các hệ thống tín dụng nội bộ, nơi người dùng:

- **Nạp tiền (Credit top-up):** Người dùng thực hiện giao dịch và nhận được số điểm tín dụng trong hệ thống.
- **Thanh toán bằng điểm:** Khi thực hiện một hành động như phục chế ảnh hoặc xóa đối tượng trong ảnh, hệ thống trừ điểm từ tài khoản thay vì thực hiện giao dịch tiền thật.

Quản lý giao dịch và hoàn tiền (Refund Management)

- **Hoàn tiền toàn phần:** Khi khách hàng không hài lòng, PayPal API cho phép hoàn toàn bộ số tiền đã thanh toán.
- **Hoàn tiền một phần:** Áp dụng khi một phần dịch vụ hoặc sản phẩm không đáp ứng đúng thỏa thuận.

Thanh toán Quốc tế và Đa loại Tiền tệ

- Hỗ trợ hơn 25 loại tiền tệ toàn cầu.
- Tự động chuyển đổi ngoại tệ theo tỷ giá PayPal.
- Phù hợp cho các nền tảng như Freelancer, Upwork, Fiverr.

Tích hợp trong ứng dụng di động và hệ thống nhúng

- PayPal API hỗ trợ SDK di động cho iOS và Android, cho phép thanh toán trực tiếp từ ứng dụng.
- Các hệ thống ứng dụng IoT và POS cũng có thể sử dụng PayPal như một cổng thanh toán an toàn.

2.8.3. Ưu điểm và hạn chế của PayPal API

Ưu điểm:

Tính bảo mật cao và đáng tin cậy: Một trong những ưu điểm nổi bật của PayPal API là mức độ bảo mật và đáng tin cậy cao. PayPal áp dụng chuẩn mã hóa SSL (Secure Sockets Layer) tiên tiến để đảm bảo rằng tất cả dữ liệu nhạy cảm, như thông tin thẻ tín dụng hoặc tài khoản ngân hàng, đều được mã hóa trước khi truyền qua mạng. Ngoài ra, hệ thống còn tuân thủ chuẩn bảo mật PCI DSS (Payment Card Industry Data Security Standard) – một tiêu chuẩn bắt buộc đối với các tổ chức xử lý thanh toán trực tuyến.

Phổ biến và dễ nhận diện trên toàn cầu: PayPal là một trong những công thanh toán trực tuyến được sử dụng phổ biến nhất trên toàn cầu, với hơn 300 triệu tài khoản đang hoạt động. Điều này tạo nên sự tin tưởng cao đối với người dùng khi thanh toán trên các trang web tích hợp PayPal API.

Dễ dàng tích hợp và tài liệu đầy đủ: PayPal API được thiết kế với RESTful API, cho phép các nhà phát triển dễ dàng tích hợp vào hệ thống web và ứng dụng. Bộ tài liệu chính thức của PayPal rất chi tiết, cung cấp các ví dụ và mã mẫu cho nhiều ngôn ngữ lập trình như Python, JavaScript, PHP.

Nhược điểm:

Phí giao dịch cao: Một trong những nhược điểm lớn nhất của PayPal API là chi phí giao dịch cao hơn so với một số công thanh toán khác. PayPal hiện áp dụng mức phí $2.9\% + \$0.30$ USD cho mỗi giao dịch quốc tế và tỷ lệ này có thể cao hơn đối với một số quốc gia. Điều này làm tăng đáng kể chi phí vận hành, đặc biệt đối với các doanh nghiệp nhỏ hoặc giao dịch có giá trị thấp.

Không phù hợp cho giao dịch vi mô (Microtransaction): Do áp dụng mức phí cố định trên mỗi giao dịch, PayPal không phù hợp cho các giao dịch vi mô (dưới \$5 USD). Ví dụ, một giao dịch có giá trị \$2 USD có thể bị trừ tới 20% tổng giá trị chỉ tính riêng chi phí xử lý. Điều này khiến PayPal trở nên kém hấp dẫn đối với các nền tảng như bán sản phẩm kỹ thuật số giá trị thấp, chẳng hạn như ebook, ảnh kỹ thuật số, khóa học ngắn hạn hoặc các dịch vụ sáng tạo nội dung nhỏ.

Hạn chế tùy chỉnh giao diện thanh toán: Dù PayPal cung cấp giao diện thanh toán chuẩn mực, nhưng khả năng tùy chỉnh lại hạn chế. Điều này gây khó khăn cho các doanh nghiệp muốn đồng nhất giao diện giữa trang web chính và trang thanh toán, từ đó ảnh hưởng đến nhận diện thương hiệu.

CHƯƠNG 3: HIỆN THỰC HÓA NGHIÊN CỨU

3.1. Đặt tả ứng dụng

Trong bối cảnh công nghệ trí tuệ nhân tạo (AI) không ngừng phát triển và khẳng định vai trò quan trọng trong nhiều lĩnh vực, nhu cầu phục chế ảnh cũ và nâng cao chất lượng hình ảnh đang trở nên cấp thiết hơn bao giờ hết. Những bức ảnh cũ, mờ nhòe theo thời gian, đặc biệt là ảnh tư liệu lịch sử, di sản văn hóa hoặc ảnh kỷ niệm gia đình, đều mang giá trị tinh thần và tư liệu to lớn. Tuy nhiên, việc phục chế những hình ảnh này thường đòi hỏi kỹ thuật chỉnh sửa phức tạp, yêu cầu chuyên môn cao và các phần mềm chuyên dụng. Điều này đặt ra nhu cầu về một giải pháp tự động, thân thiện với người dùng nhưng vẫn đảm bảo chất lượng đầu ra cao, phục vụ cả mục đích lưu trữ, bảo tồn và sáng tạo nội dung.

Đề tài "Nghiên cứu GFPGAN và xây dựng trang web phục chế ảnh bằng AI" được phát triển với mục tiêu giải quyết bài toán trên thông qua việc ứng dụng các công nghệ trí tuệ nhân tạo tiên tiến. Hệ thống cho phép người dùng phục chế ảnh chân dung cũ, làm nét và phóng to ảnh, xóa đối tượng không mong muốn, tạo ảnh từ văn bản và chuyển đổi ảnh thành video một cách hoàn toàn tự động. Để hỗ trợ việc vận hành và quản lý dịch vụ, hệ thống tích hợp PayPal API cho phép người dùng nạp tín dụng (credit) và sử dụng các tính năng nâng cao. Bên cạnh đó, đề tài còn xây dựng một giao diện quản trị viên nhằm quản lý dữ liệu người dùng, theo dõi lịch sử giao dịch và kiểm soát quá trình xử lý ảnh. Với kiến trúc web thân thiện, đề tài không chỉ hướng đến việc phục chế ảnh mang tính cá nhân mà còn mở ra tiềm năng ứng dụng trong lưu trữ lịch sử, bảo tồn văn hóa và sản xuất nội dung sáng tạo trong thời đại số hóa.

3.2. Thiết kế kiến trúc hệ thống

3.2.1. Thiết kế kiến trúc tổng quan của hệ thống

Hệ thống được thiết kế theo mô hình Client-Server kết hợp với nguyên tắc modular design nhằm đảm bảo phân tách rõ ràng giữa giao diện người dùng (Frontend) và phần xử lý logic (Backend). Kiến trúc này không chỉ giúp tổ chức mã nguồn khoa học, mà còn tối ưu hóa khả năng mở rộng và bảo trì hệ thống trong tương lai.

Phần Frontend tập trung vào việc hiển thị giao diện và xử lý tương tác của người dùng, trong khi phần Backend sử dụng mô hình Flask Blueprint để tổ chức các chức năng theo từng module độc lập, từ quản lý người dùng, phục chế ảnh, cho đến xử lý thanh toán.

Dưới đây là các thành phần chính của hệ thống:

1. Frontend (Giao diện người dùng)

- Được phát triển bằng HTML, CSS và JavaScript.
- Thư mục: frontend/css, frontend/js.
- Giao tiếp với backend thông qua các HTTP request.

2. Backend (Flask Server)

- Ngôn ngữ chính: Python.
- Framework: Flask.
- Được thiết kế theo mô hình **Blueprint** để chia nhỏ từng tính năng xử lý:
 - **routes_auth.py**: Xử lý xác thực và quản lý tài khoản người dùng.
 - **routes_gfpgan.py**: Phục chế ảnh chân dung bằng GFPGAN.
 - **routes_realesrgan.py**: Làm nét ảnh bằng Real-ESRGAN.
 - **routes_lama.py**: Xóa đối tượng bằng LaMa.
 - **routes_stable.py**: Tạo ảnh từ văn bản bằng Stable Diffusion.
 - **routes_admin.py**: Quản lý người dùng và dữ liệu (Admin).
 - **routes_home.py**: Quản lý trang chủ.
 - **routes_frontend.py**: Quản lý tệp tĩnh của frontend.
 - **routes_footer_assets.py** và **routes_header_assets.py**: Quản lý tài nguyên giao diện.

3. Cơ sở dữ liệu (SQLite)

Hệ thống sử dụng SQLite để lưu trữ và quản lý dữ liệu, với các bảng dữ liệu trọng yếu bao gồm:

- **NguoiDung**: Lưu trữ thông tin tài khoản và vai trò của người dùng.
- **LichSuSuDung**: Ghi nhận lịch sử sử dụng các tính năng xử lý ảnh trên hệ thống.
- **AnhXuLy**: Quản lý thông tin các tệp ảnh đã qua xử lý.
- **ThanhToan**: Lưu trữ thông tin các giao dịch thanh toán từ người dùng.
- **GoiThanhToan**: Chứa thông tin về các gói tín dụng có sẵn.
- **LienKet**: Liên kết giữa các giao dịch và gói tín dụng tương ứng.

Cấu trúc cơ sở dữ liệu được thiết kế chặt chẽ nhằm đảm bảo tính toàn vẹn dữ liệu, hỗ trợ tốt cho việc truy vấn và mở rộng hệ thống trong tương lai.

4. Công nghệ xử lý ảnh AI

- **GFPGAN:** Phục chế ảnh chân dung.
- **Real-ESRGAN:** Làm nét và phóng to ảnh.
- **LaMa:** Xóa đối tượng không mong muốn trong ảnh.
- **Stable Diffusion:** Tạo ảnh từ văn bản.
- **KLing AI:** Chuyển ảnh thành video.

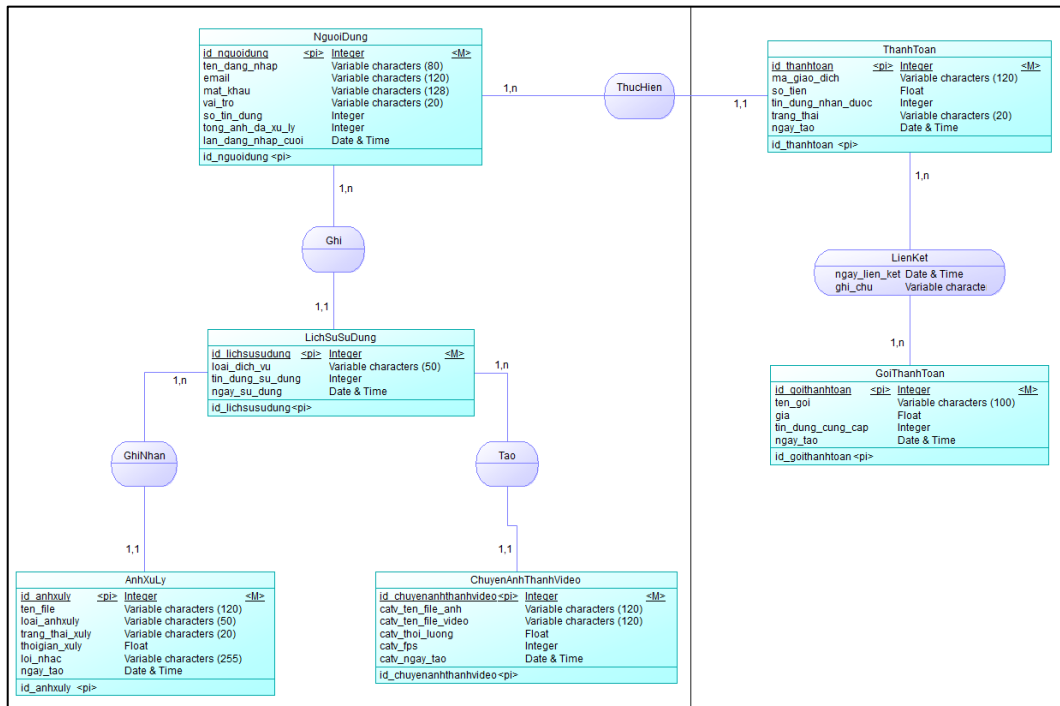
Tất cả các công nghệ này được gọi thông qua các route riêng biệt, đảm bảo tách biệt logic xử lý.

5. Hệ thống thanh toán

Tích hợp PayPal API để xử lý thanh toán tín dụng, cho phép người dùng nạp tiền và sử dụng các tính năng cao cấp.

3.2.2. Thiết kế cơ sở dữ liệu

Cơ sở dữ liệu của hệ thống được thiết kế theo mô hình quan hệ (Relational Database) sử dụng SQLite nhằm đảm bảo tính toàn vẹn dữ liệu, dễ quản lý và phù hợp với quy mô ứng dụng. Cấu trúc dữ liệu được chuẩn hóa theo dạng 3NF (Third Normal Form) nhằm tránh dư thừa dữ liệu và tối ưu hóa hiệu suất truy vấn.



Hình 3. 1. Minh họa mô hình ERD

Các bảng dữ liệu chính trong hệ thống:

Bảng NgoiDung:

Bảng 3. 1. Bảng cơ sở dữ liệu thông tin người dùng

Tên cột	Kiểu dữ liệu	Ràng buộc	Mô tả
id_nguoidung	INTEGER	PK	Khóa chính định danh người dùng.
ten_dang_nhap	VARCHAR(50)	NOT NULL, UNIQUE	Tên đăng nhập người dùng.
email	VARCHAR(100)	NOT NULL, UNIQUE	Địa chỉ email của người dùng.
mat_khau	VARCHAR(255)	NOT NULL	Mật khẩu đã mã hóa.
vai_tro	VARCHAR(20)	NOT NULL	Vai trò (user, admin).
so_tin_dung	INTEGER	DEFAULT 0	Số tin dụng hiện tại.
tong_anh_da_xu_ly	INTEGER	DEFAULT 0	Tổng số ảnh đã xử lý.
lan_dang_nhap_cuoi	DATETIME	NULLABLE	Lần cuối cùng người dùng đăng nhập.

Bảng LichSuSuDung:

Bảng 3. 2. Bảng cơ sở dữ liệu thông tin lịch sử sử dụng

Tên cột	Kiểu dữ liệu	Ràng buộc	Mô tả
id_lichsusudung	INTEGER	PK	Khóa chính.
id_nguoidung	INTEGER	FK → NguoiDung	Khóa ngoại liên kết người dùng.
loai_dich_vu	VARCHAR(50)	NOT NULL	Loại dịch vụ sử dụng.
tin_dung_su_dung	INTEGER	NOT NULL	Số tin dụng đã sử dụng.
ngay_su_dung	DATETIME	NOT NULL	Thời gian thực hiện thao tác.

Bảng AnhXuLy:

Bảng 3. 3. Bảng cơ sở dữ liệu ảnh xử lý

Tên cột	Kiểu dữ liệu	Ràng buộc	Mô tả
id_anhxuly	INTEGER	PK	Khóa chính.
id_lichsusudung	INTEGER	FK → LichSuSuDung	Khóa ngoại liên kết lịch sử sử dụng.
ten_file	VARCHAR(120)	NOT NULL	Tên tệp ảnh đã xử lý.
loai_anhxuly	VARCHAR(50)	NOT NULL	Loại xử lý ảnh
trang_thai_xuly	VARCHAR(20)	NOT NULL	Trạng thái xử lý.
thoi_gian_xuly	TIME	NULLABLE	Thời gian xử lý.
loi_nhac	VARCHAR(255)	NULLABLE	Gợi ý nội dung cho ảnh (nếu có).
ngay_tao	DATETIME	NOT NULL	Ngày tạo bản ghi.

Bảng ChuyenAnhThanhVideo:

Bảng 3. 4. Bảng cơ sở dữ liệu chuyển ảnh thành video

Tên cột	Kiểu dữ liệu	Ràng buộc	Mô tả
id_chuyenanhthanhvideo	INTEGER	PK	Khóa chính.
catv_ten_file_anh	VARCHAR(120)	NOT NULL	Tên tệp ảnh đầu vào.
catv_ten_file_video	VARCHAR(120)	NOT NULL	Tên tệp video đầu ra.
catv_thoi_luong	INTEGER	NOT NULL	Thời lượng video (giây).
catv_fps	INTEGER	NOT NULL	Số khung hình trên giây.
catv_ngay_tao	DATETIME	NOT NULL	Ngày tạo bản ghi.

Tên cột	Kiểu dữ liệu	Ràng buộc	Mô tả
---------	--------------	-----------	-------

Bảng ThanhToan:

Bảng 3. 5. Bảng cơ sở dữ liệu thanh toán

Tên cột	Kiểu dữ liệu	Ràng buộc	Mô tả
id_thanhtoan	INTEGER	PK, Auto-increment	Khóa chính.
id_nguoidung	INTEGER	FK → NguoiDung	Khóa ngoại liên kết người dùng.
ma_giao_dich	VARCHAR(100)	NOT NULL	Mã giao dịch từ PayPal.
so_tien	DECIMAL(10,2)	NOT NULL	Số tiền thanh toán.
tin_dung_nhan_duoc	INTEGER	NOT NULL	Tín dụng nhận được sau giao dịch.

Bảng GoiThanhToan:

Bảng 3. 6. Bảng cơ sở dữ liệu gói thanh toán

Tên cột	Kiểu dữ liệu	Ràng buộc	Mô tả
id_goithanhtoan	INTEGER	PK	Khóa chính.
ten_goi	VARCHAR(100)	NOT NULL	Tên gói tín dụng.
gia	DECIMAL(10,2)	NOT NULL	Giá gói tín dụng.
tin_dung_cung_cap	INTEGER	NOT NULL	Số tín dụng cung cấp.
ngay_tao	DATETIME	NOT NULL	Ngày tạo gói tín dụng.

Bảng LienKet:

Bảng 3. 7. Bảng cơ sở dữ liệu liên kết

Tên cột	Kiểu dữ liệu	Ràng buộc	Mô tả
---------	--------------	-----------	-------

id_thanhtoan	INTEGER	FK → ThanhToan	Khóa ngoại liên kết giao dịch.
id_goithanhtoan	INTEGER	FK → GoiThanhToan	Khóa ngoại liên kết gói tín dụng.
ngay_lien_ket	DATETIME	NOT NULL	Ngày liên kết.
ghi_chu	VARCHAR(255)	NULLABLE	Ghi chú thêm (nếu có).

3.3. Xây dựng các Route theo kiến trúc Blueprint

3.3.1. Route khởi tạo ứng dụng

Route khởi tạo ứng dụng là thành phần trung tâm đảm bảo hệ thống web hoạt động đúng cấu trúc và sẵn sàng phục vụ yêu cầu từ người dùng. Cấu trúc này được thiết kế theo hướng modular, giúp quản lý và mở rộng dễ dàng.

Quy trình khởi tạo ứng dụng:

Khởi tạo ứng dụng và cấu hình:

- Tạo một đối tượng ứng dụng web Flask.
- Nạp các biến môi trường (.env) như khóa bảo mật, đường dẫn cơ sở dữ liệu và thư mục lưu trữ.
- Thiết lập kết nối với cơ sở dữ liệu SQLite.

Cấu hình CSDL và Thư mục:

- CSDL SQLite được thiết lập mặc định.
- Tạo các thư mục lưu trữ ảnh gốc và ảnh đã qua xử lý.

Đăng ký Route (Blueprint):

- Các route xử lý chính như: phục chế ảnh, làm nét ảnh, xóa đối tượng, tạo ảnh từ văn bản, chuyển đổi video, và quản lý tài khoản được tổ chức thành từng module riêng biệt (Blueprint).
- Điều này đảm bảo mã nguồn tách biệt theo từng chức năng.

Chạy Ứng dụng:

- Ứng dụng khởi chạy ở chế độ debug để hỗ trợ phát hiện lỗi trong quá trình phát triển.

Route khởi tạo ứng dụng đóng vai trò cốt lõi trong việc thiết lập toàn bộ hệ thống, đảm bảo môi trường làm việc ổn định, an toàn và sẵn sàng mở rộng. Cấu trúc này không chỉ giúp tổ chức mã nguồn rõ ràng mà còn hỗ trợ khả năng mở rộng, bảo trì hệ thống trong tương lai. Các route khác

3.3.2. Route quản lý người dùng

Route quản lý người dùng là thành phần quan trọng trong hệ thống, chịu trách nhiệm xử lý các tác vụ liên quan đến tài khoản người dùng như đăng ký, đăng nhập, quản lý vai trò và bảo mật dữ liệu. Hệ thống được thiết kế đảm bảo tính bảo mật, dễ dàng kiểm soát và quản lý thông tin người dùng.

Quy trình quản lý người dùng:

Đăng ký tài khoản:

- Người dùng cung cấp thông tin cá nhân như tên đăng nhập, email và mật khẩu.
- Mật khẩu được mã hóa trước khi lưu trữ vào cơ sở dữ liệu.
- Kiểm tra trùng lặp tài khoản trước khi xác nhận đăng ký.

Xác thực và đăng nhập:

- Thông tin đăng nhập được đối chiếu với dữ liệu trong cơ sở dữ liệu.
- Nếu xác thực thành công, hệ thống khởi tạo phiên làm việc (session) để theo dõi trạng thái đăng nhập.
- Ngược lại, người dùng nhận thông báo lỗi và được yêu cầu thử lại.

Quản lý vai trò và quyền hạn:

- Phân quyền theo vai trò: Người dùng thông thường và quản trị viên.
- Quản trị viên có quyền thêm, xóa và chỉnh sửa thông tin người dùng.

Đăng xuất và hủy phiên:

- Khi người dùng thực hiện đăng xuất, phiên làm việc sẽ được xóa bỏ để đảm bảo tính bảo mật.

Bảo mật và kiểm tra:

- Các thao tác như mã hóa mật khẩu và kiểm tra mật khẩu bị rò rỉ đều được thực hiện để bảo vệ dữ liệu người dùng.
- Thiết lập cơ chế giới hạn phiên hoạt động và ngăn chặn truy cập trái phép được áp dụng.

Route quản lý người dùng không chỉ là thành phần quan trọng trong việc đảm bảo trải nghiệm người dùng liền mạch, mà còn đóng vai trò bảo vệ hệ thống khỏi các

rủi ro liên quan đến bảo mật dữ liệu. Việc thiết kế tách biệt và tuân theo mô hình phân quyền giúp đảm bảo hệ thống an toàn, minh bạch và dễ quản lý.

3.3.3. Route quản trị viên

Route quản trị viên là thành phần quan trọng đảm bảo việc giám sát và điều hành toàn bộ hoạt động trên hệ thống web. Đây là nơi tập trung các chức năng quản lý tài khoản người dùng, giao dịch và giám sát lịch sử sử dụng dịch vụ, đảm bảo mọi hoạt động diễn ra minh bạch và an toàn.

Quy trình hoạt động:

- **Truy cập bảng điều khiển:** Chỉ những tài khoản có vai trò "Quản trị viên" (Admin) mới được phép truy cập trang quản trị. Người quản trị sau khi đăng nhập sẽ được chuyển hướng tới giao diện quản lý riêng biệt.
- **Quản lý người dùng:** Tại giao diện quản trị, quản trị viên có thể thực hiện các thao tác như thêm mới, chỉnh sửa thông tin, gán quyền và xóa tài khoản người dùng. Các hành động này được ghi nhận trong CSDL nhằm đảm bảo theo dõi lịch sử thay đổi.
- **Kiểm soát thanh toán:** Quản trị viên có thể xem chi tiết lịch sử thanh toán, trạng thái của các giao dịch, xử lý hoàn tiền nếu cần và đảm bảo rằng các giao dịch thông qua PayPal API được xử lý đúng quy trình.
- **Theo dõi lịch sử sử dụng:** Toàn bộ lịch sử sử dụng các dịch vụ (phục chế ảnh, làm nét ảnh, xóa đối tượng,...) đều có thể được kiểm tra trong giao diện quản trị để phục vụ cho mục đích phân tích hoặc xử lý các yêu cầu hỗ trợ người dùng.

Route quản trị viên đóng vai trò trung tâm trong việc đảm bảo hệ thống vận hành an toàn, bảo mật và minh bạch.

3.3.4. Route phục chế ảnh chân dung

Route này triển khai tính năng phục chế ảnh chân dung sắc nét sử dụng mô hình GFPGAN, giúp cải thiện độ nét và khôi phục chi tiết gương mặt trong các bức ảnh cũ hoặc bị mờ.

Quy trình hoạt động:

- **Tải lên ảnh:** Người dùng chọn tệp ảnh cần phục chế từ thiết bị cá nhân và tải lên hệ thống thông qua giao diện web.

- **Kiểm tra và xử lý:** Hệ thống kiểm tra định dạng và dung lượng tệp ảnh trước khi chuyển dữ liệu vào mô hình GFPGAN.
- **Xử lý ảnh:** GFPGAN sẽ phân tích khuôn mặt trong ảnh, phục chế các chi tiết gương mặt bằng các thuật toán AI và trả về kết quả đã cải thiện.
- **Lưu trữ ảnh phục chế:** Sau khi xử lý, kết quả ảnh sẽ được lưu trữ tại thư mục static/restored và thông tin ảnh được ghi nhận trong bảng AnhXuLy trong cơ sở dữ liệu.
- **Hiển thị kết quả:** Người dùng sẽ thấy ảnh phục chế và có thể tải xuống tệp kết quả từ giao diện.

Route này cung cấp giải pháp phục chế ảnh chân dung một cách tự động và hiệu quả, phù hợp cho việc bảo tồn ảnh lịch sử hoặc cải thiện chất lượng ảnh cũ.

3.3.5. Route làm nét và phóng to ảnh

Route này sử dụng mô hình Real-ESRGAN để làm nét và phóng to ảnh với độ chi tiết cao, đặc biệt phù hợp cho các hình ảnh có độ phân giải thấp hoặc cần cải thiện chất lượng.

Quy trình hoạt động:

- **Tải lên ảnh:** Người dùng tải lên ảnh cần làm nét và có thể chọn mức độ phóng to ảnh (2x, 4x).
- **Xác thực dữ liệu:** Hệ thống sẽ kiểm tra dung lượng ảnh và định dạng lại tệp trước khi xử lý.
- **Tiền xử lý:** Ảnh được chuyển đổi thành dạng base64 để truyền qua API để xử lý.
- **Xử lý ảnh:** Real-ESRGAN thực hiện phân tích và cải thiện chi tiết ảnh bằng thuật toán AI chuyên biệt cho việc phóng to mà không làm mờ ảnh.
- **Lưu trữ và hiển thị:** Ảnh sau khi làm nét được lưu trữ tại thư mục static/restored và hiển thị trên giao diện web.
- **Ghi nhận lịch sử:** Thông tin ảnh và kết quả làm nét được ghi nhận vào bảng AnhXuLy.

Route làm nét và phóng to ảnh giúp cải thiện chất lượng hình ảnh, đặc biệt trong các trường hợp ảnh cũ hoặc có độ phân giải thấp.

3.3.6. Route xóa đối tượng trong ảnh

Route này triển khai tính năng xóa đối tượng không mong muốn trong ảnh bằng cách sử dụng mô hình LaMa (Large Masked Autoencoder).

- **Quy trình hoạt động:**
- **Tải ảnh lên và tạo mask:** Người dùng tải ảnh cần chỉnh sửa và vẽ vùng mask trực tiếp trên giao diện web để chỉ định đối tượng cần xóa.
- **Kiểm tra dữ liệu:** Hệ thống kiểm tra dữ liệu mask và kiểm tra xem ảnh có hợp lệ trước khi xử lý.
- **Xử lý ảnh:** Mô hình LaMa nhận ảnh và mask, loại bỏ đối tượng không mong muốn bằng cách tái tạo lại vùng ảnh bị xóa dựa trên các đặc trưng xung quanh.
- **Lưu trữ và hiển thị:** Kết quả ảnh đã xử lý được lưu trữ và hiển thị cho người dùng xem trước.
- **Cập nhật lịch sử sử dụng:** Mọi thao tác chỉnh sửa sẽ được ghi nhận vào bảng AnhXuLy trong CSDL.

Route này cung cấp khả năng xóa đối tượng không mong muốn tự động, đặc biệt phù hợp cho việc chỉnh sửa ảnh chuyên nghiệp.

3.3.7. Route tạo ảnh từ văn bản

Route này ứng dụng mô hình Stable Diffusion thông qua API từ Huggingface để tạo ảnh từ văn bản đầu vào của người dùng.

Quy trình hoạt động:

- **Nhập văn bản:** Người dùng nhập một đoạn mô tả (prompt) vào giao diện.
- **Xác thực đầu vào:** Hệ thống kiểm tra nội dung mô tả và độ dài hợp lệ.
- **Xử lý tạo ảnh:** Prompt được gửi tới mô hình Stable Diffusion, thực hiện quá trình tạo ảnh dựa trên nội dung đã mô tả.
- **Hiển thị ảnh:** Ảnh sinh ra được hiển thị trực tiếp trên giao diện web.
- **Lưu trữ dữ liệu:** Thông tin ảnh và mô tả được ghi nhận vào bảng AnhXuLy.

Route này hỗ trợ người dùng sáng tạo nội dung hình ảnh nhanh chóng và chuyên nghiệp thông qua lệnh prompt đơn giản.

3.3.8. Route chuyển đổi ảnh thành video

Route này sử dụng mô hình Kling AI để chuyển đổi các bức ảnh tĩnh thành video chuyển động mượt mà.

Quy trình hoạt động:

- **Tải lên bộ ảnh:** Người dùng tải lên các ảnh cần chuyển đổi thành video.
- **Tiền xử lý ảnh:** Hệ thống xác thực số lượng ảnh và dung lượng trước khi xử lý truyền đến Kling AI.
- **Chuyển đổi ảnh:** Kling AI thực hiện tạo hiệu ứng chuyển động và kết hợp các ảnh thành video hoàn chỉnh.
- **Lưu trữ và hiển thị:** Video được hiển thị trực tiếp trên web và lưu trữ trong thư mục static/restored.
- **Cập nhật lịch sử:** Thông tin video và quá trình xử lý được ghi nhận vào bảng ChuyenAnhThanhVideo.

Route này mang lại khả năng sáng tạo video từ ảnh tĩnh, phù hợp cho các dự án truyền thông và sáng tạo nội dung.

3.3.9. Route giao diện chính và xử lý trang chủ

Route này đảm nhận vai trò hiển thị giao diện chính của hệ thống, giúp người dùng dễ dàng truy cập và điều hướng giữa các tính năng.

Quy trình hoạt động:

- **Trang chủ:** Hiển thị giao diện chính với các liên kết dẫn đến các tính năng xử lý ảnh và video.
- **Điều hướng:** Cung cấp các liên kết trực tiếp đến các route phục chế ảnh, làm nét ảnh, xóa đối tượng, tạo video,...
- **Quản lý phiên làm việc:** Kiểm tra trạng thái đăng nhập và hiển thị tùy chọn tài khoản phù hợp (đăng nhập, đăng ký, đăng xuất).
- **Tích hợp nội dung động:** Kết hợp các route xử lý và hiển thị dữ liệu theo thời gian thực.

Route giao diện chính là cổng vào trung tâm, đảm bảo tính trực quan, dễ sử dụng cho người dùng và đồng nhất trong trải nghiệm hệ thống.

3.3.10. Route thanh toán

Route thanh toán đảm nhận vai trò quản lý toàn bộ quy trình giao dịch và xử lý thanh toán trên hệ thống, đảm bảo tính bảo mật, minh bạch và tiện lợi cho người dùng khi thực hiện giao dịch thông qua PayPal API. Tính năng thanh toán được thiết kế để người dùng có thể nạp tín dụng (credit) nhằm sử dụng các dịch vụ nâng cao của hệ thống.

Quy trình hoạt động:

Khởi tạo giao dịch:

- Người dùng chọn một gói tín dụng phù hợp từ danh sách các gói đã được thiết lập trước trong bảng GoiThanhToan.
- Sau khi chọn gói, hệ thống hiển thị chi tiết giá trị và số lượng credit tương ứng.
- Người dùng xác nhận và hệ thống khởi tạo giao dịch qua PayPal API.

Chuyển hướng tới cổng thanh toán:

- Hệ thống thực hiện lệnh gọi API tới PayPal để khởi tạo yêu cầu thanh toán, bao gồm thông tin người dùng, số tiền cần thanh toán và đơn vị tiền tệ.
- Người dùng được chuyển hướng tới cổng PayPal chính thức để hoàn tất thanh toán.

Xác nhận và xử lý phản hồi từ PayPal:

- Sau khi thanh toán thành công, PayPal gửi phản hồi (callback) về hệ thống thông qua một webhook.
- Hệ thống xác thực giao dịch bằng cách kiểm tra thông tin phản hồi như: mã giao dịch, trạng thái thành công, và số tiền đã thanh toán.

Cập nhật thông tin thanh toán và tín dụng:

- Nếu giao dịch thành công, hệ thống cập nhật số tín dụng (credit) vào tài khoản người dùng trong bảng NguoiDung.
- Thông tin giao dịch được ghi nhận vào bảng ThanhToan, bao gồm mã giao dịch, thời gian thanh toán, và số tín dụng nạp thành công.

Xử lý lỗi và hoàn tiền (nếu cần):

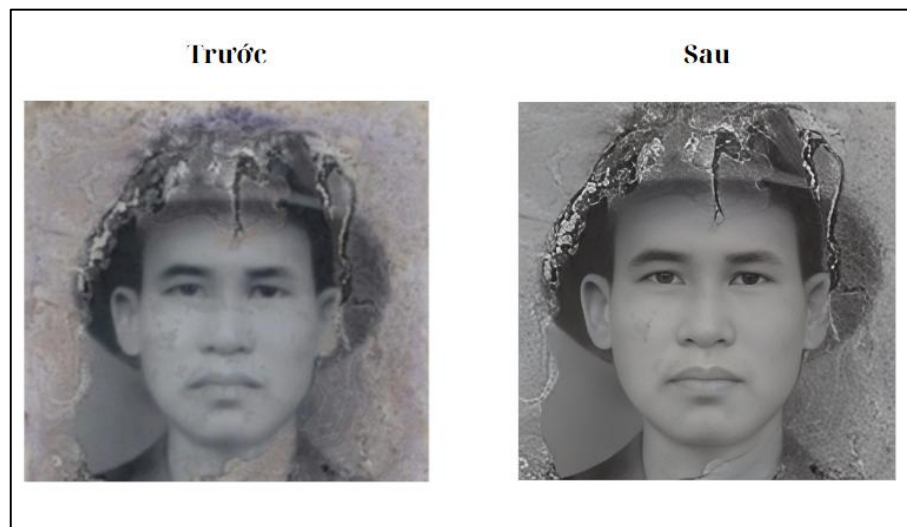
- Trong trường hợp giao dịch thất bại hoặc bị hủy, hệ thống sẽ thông báo lỗi đến người dùng.
- Nếu có yêu cầu hoàn tiền, quản trị viên có thể thực hiện hoàn tiền thông qua giao diện quản trị hoặc trực tiếp qua PayPal API.

Route thanh toán không chỉ đảm bảo quy trình giao dịch diễn ra tự động và minh bạch mà còn đóng vai trò quan trọng trong việc quản lý tài nguyên và doanh thu của hệ thống. Việc tích hợp PayPal API giúp đảm bảo an toàn cho người dùng và tạo ra trải nghiệm thanh toán trực tuyến chuyên nghiệp.

CHƯƠNG 4: KẾT QUẢ NGHIÊN CỨU

4.1. Dữ liệu thử nghiệm

4.1. Dữ liệu thử nghiệm với chức năng phục chế ảnh chân dung



Hình 4. 1. Minh họa kết quả phục chế ảnh chân dung bằng GFPGAN

4.2. Dữ liệu thử nghiệm với chức năng xóa đối tượng trong ảnh



Hình 4. 2. Minh họa kết quả chức năng xóa đối tượng trong ảnh

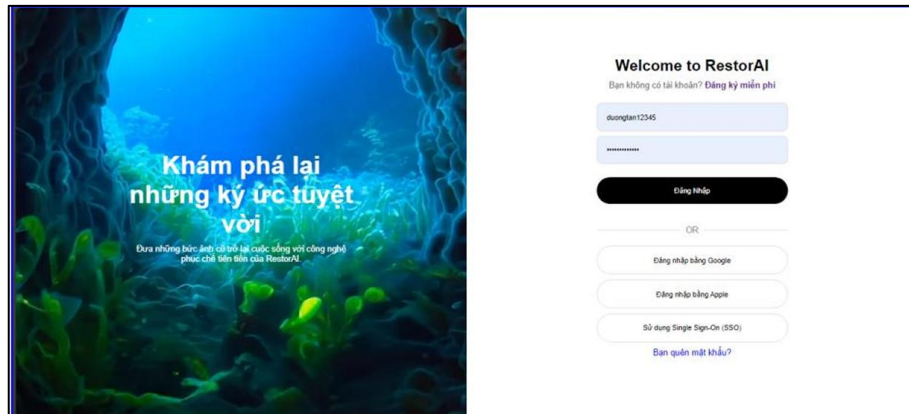
4.3. Dữ liệu thử nghiệm với chức năng tạo ảnh từ văn bản



Hình 4. 3. Minh họa kết quả tính năng tạo ảnh từ văn bản

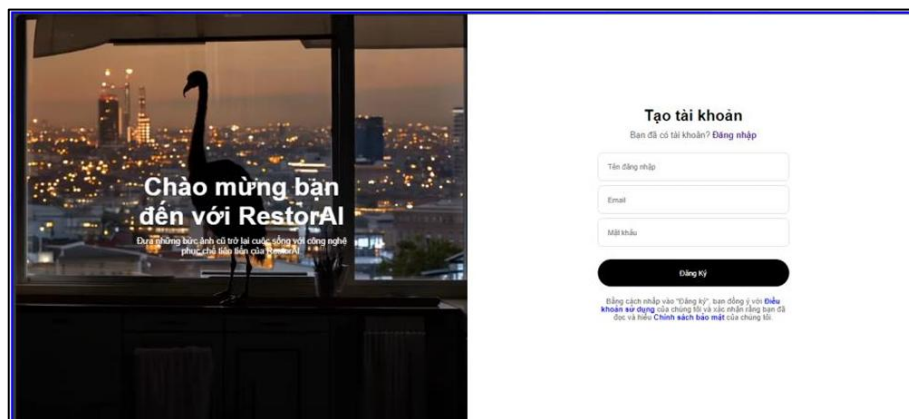
4.2. Kết quả thực nghiệm hệ thống

4.2.1. Chức năng đăng nhập



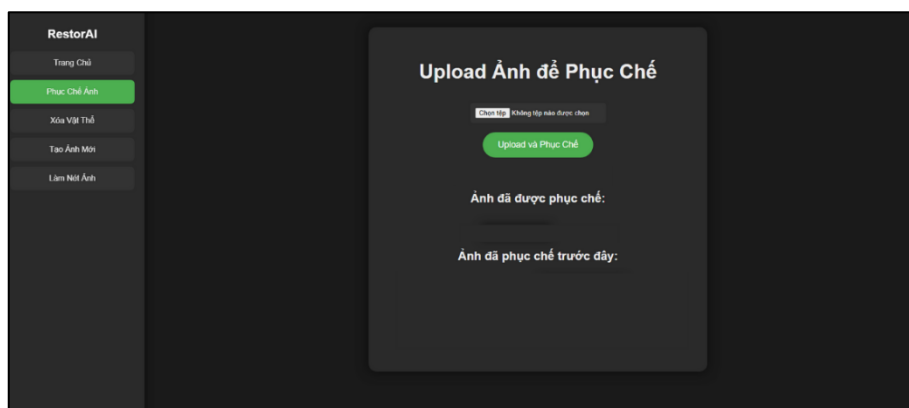
Hình 4. 4. Minh họa giao diện đăng nhập

4.2.2. Chức năng đăng ký



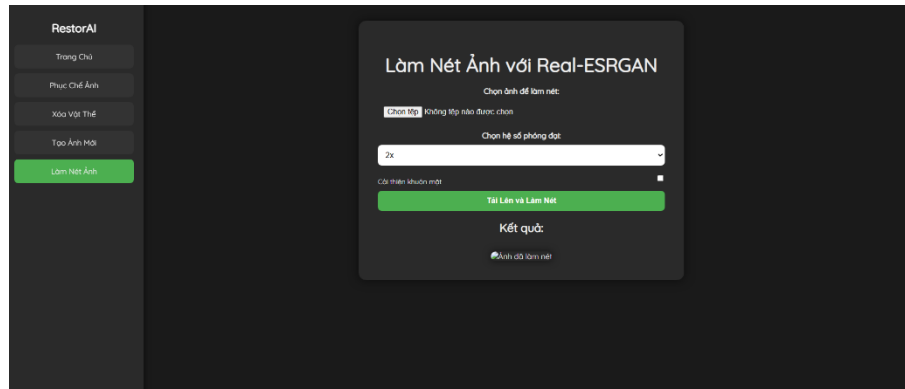
Hình 4. 5. Minh họa giao diện đăng ký

4.2.3. Chức năng chức năng phục chế ảnh chân dung



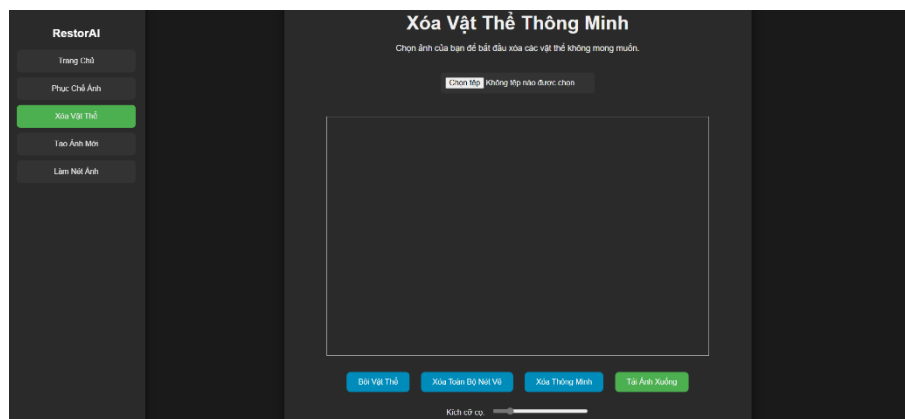
Hình 4. 6. Minh họa giao diện chức năng phục chế ảnh

4.2.4. Chức năng làm nét và phóng to ảnh



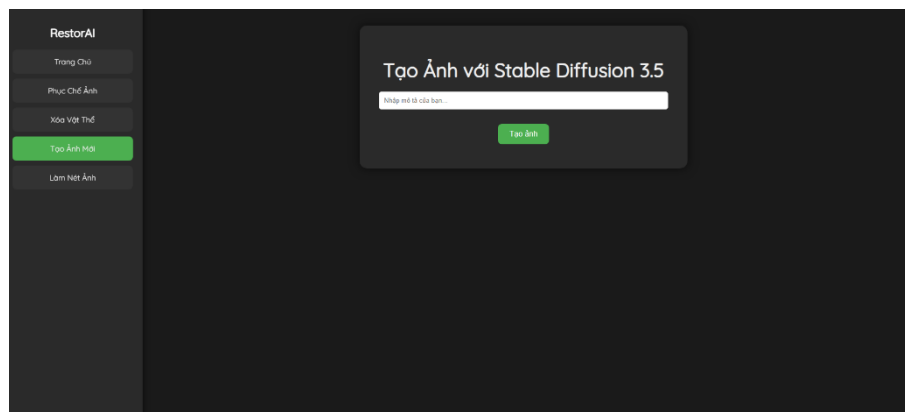
Hình 4. 7. Minh họa giao diện chức năng làm nét ảnh

4.2.5. Chức năng xóa đối tượng trong ảnh



Hình 4. 8. Minh họa giao diện chức năng xóa đối tượng trong ảnh

4.2.6. Chức năng tạo ảnh từ văn bản



Hình 4. 9. Minh họa giao diện chức năng tạo ảnh từ văn bản

CHƯƠNG 5: KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

5.1. Kết luận

Trong quá trình thực hiện đề tài “Nghiên cứu GFPGAN và xây dựng trang web phục chế ảnh bằng AI”, hệ thống đã được thiết kế và xây dựng thành công với các chức năng cốt lõi như phục chế ảnh chân dung, làm nét và phóng to ảnh, xóa đối tượng không mong muốn, tạo ảnh từ văn bản và chuyển đổi ảnh thành video. Các thuật toán và mô hình trí tuệ nhân tạo như GFPGAN, Real-ESRGAN, LaMa, Stable Diffusion và Kling AI đã được tích hợp và thử nghiệm với kết quả khả quan. Ngoài ra, hệ thống còn bao gồm cơ chế quản lý tài khoản người dùng, quản trị viên và tích hợp thanh toán qua PayPal API, đảm bảo tính ứng dụng thực tế cao.

Hệ thống được thiết kế theo mô hình Client-Server với kiến trúc modular, giúp tách biệt rõ ràng các thành phần xử lý, đảm bảo việc quản lý và mở rộng trong tương lai. Cơ sở dữ liệu SQLite được xây dựng theo nguyên tắc chuẩn hóa, lưu trữ đầy đủ thông tin người dùng, lịch sử sử dụng và giao dịch. Kết quả thực nghiệm cho thấy hệ thống hoạt động ổn định, các tính năng xử lý ảnh và thanh toán vận hành đúng yêu cầu, giao diện thân thiện với người dùng.

5.2. Hạn chế

Dù đã đạt được những thành công nhất định, hệ thống vẫn tồn tại một số hạn chế cần khắc phục:

Hiệu suất xử lý ảnh lớn: Khi xử lý ảnh có độ phân giải cao hoặc dữ liệu phức tạp, tốc độ xử lý có thể chậm.

Khả năng mở rộng cơ sở dữ liệu: Hệ quản trị dữ liệu SQLite phù hợp với quy mô nhỏ, tuy nhiên, để hỗ trợ lượng người dùng lớn hơn, cần cân nhắc chuyển đổi sang các hệ quản trị như PostgreSQL hoặc MySQL.

Tính phụ thuộc vào bên thứ ba: Một số tính năng như làm nét ảnh, tạo video từ ảnh và tạo ảnh từ văn bản đang phụ thuộc vào các API bên thứ ba, gây hạn chế khi cần sử dụng cục bộ.

Trải nghiệm người dùng: Giao diện hiện tại chưa thực sự trực quan và cần cải thiện để phù hợp với người dùng phổ thông hơn.

Thanh toán hạn chế: Hệ thống mới chỉ tích hợp PayPal API, chưa hỗ trợ các cổng thanh toán nội địa như VNPay hay MoMo.

5.3. Phương hướng phát triển

Trong tương lai, hệ thống có thể được cải tiến và mở rộng theo các hướng sau:

- **Nâng cấp hạ tầng xử lý và CSDL:** Chuyển từ SQLite sang PostgreSQL để đảm bảo khả năng mở rộng và tốc độ xử lý dữ liệu lớn hơn.
- **Tối ưu hóa hiệu suất xử lý ảnh:** Cải tiến thuật toán và cân nhắc triển khai trên nền tảng điện toán đám mây như Azure hoặc AWS để tăng tốc xử lý ảnh dung lượng lớn.
- **Tích hợp thêm công nghệ AI:** Nghiên cứu và áp dụng thêm các mô hình tiên tiến như CodeFormer hoặc SwinIR để cải thiện khả năng xử lý và phục chế ảnh.
- **Cải tiến giao diện người dùng (UI/UX):** Tái thiết kế giao diện để thân thiện hơn với người dùng, tăng cường trải nghiệm thao tác kéo thả và xử lý trực quan.
- **Tích hợp đa dạng cổng thanh toán:** Bổ sung các phương thức thanh toán như VNPay, MoMo và Stripe để đáp ứng đa dạng nhu cầu người dùng.
- **Mở rộng chức năng và dịch vụ:** Thêm các chức năng như ghép ảnh, chỉnh màu sắc và phát triển tính năng tạo video chuyên sâu từ AI.

Những cải tiến này sẽ giúp hệ thống trở nên toàn diện hơn, đáp ứng tốt hơn nhu cầu thực tế và tăng tính cạnh tranh trong lĩnh vực ứng dụng AI vào xử lý hình ảnh.

DANH MỤC TÀI LIỆU THAM KHẢO

- [1] X. Wang, Y. Chen, H. Zhang, and L. Zhang, “GFPGAN: Towards Real-World Blind Face Restoration with Generative Facial Prior,” arXiv preprint, arXiv:2101.04061, 2021.
- [2] X. Wang, Y. Chen, H. Zhang, and L. Zhang, “Real-ESRGAN: Training Real-World Blind Super-Resolution with Pure Synthetic Data,” arXiv preprint, arXiv:2107.07677, 2021.
- [3] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, and C. Dong, “ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks,” arXiv preprint, arXiv:1809.00219, 2018.
- [4] A. Matveev, A. Bochkovskiy, and K. Lempitsky, “LaMa: Resolution-robust Large Mask Inpainting with Fourier Convolutions,” arXiv preprint, arXiv:2109.07161, 2021.
- [5] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, “High-Resolution Image Synthesis with Latent Diffusion Models,” arXiv preprint, arXiv:2112.10752, 2022.
- [6] KLing AI Research Group, “KLing AI for Image-to-Video Transformation,” Research Paper, 2023.