**BỘ CÔNG THƯƠNG**

**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THÀNH PHỐ HỒ CHÍ MINH**

**Khoa Công Nghệ Thông Tin**

Logo

Description automatically generated

**BÁO CÁO**

**MÔN HỌC : THỊ GIÁC MÁY TÍNH**

**Virtual Interior Designer**

**Giảng viên hướng dẫn: PGS. TS. Huỳnh Trung Hiếu**

**Lớp: DHKHDL17A**

**Tác giả**

**Nguyễn Thanh Hoài - 21023751**

**Phạm Chí Nguyên - 21104551**

**Trần Quang Vinh – 21071061**

**LỜI CẢM ƠN**

Chúng em xin gửi lời chân thành cảm ơn đến **Trường Đại học Công Nghiệp thành phố Hồ Chí Minh** đã tạo điều kiện thuận lợi cho chúng em học tập và hoàn thành đề tài nghiên cứu này. Đặc biệt, chúng em xin bày tỏ lòng biết ơn sâu sắc đến thầy **PGS. TS. Huỳnh Trung Hiếu** đã tận tâm chỉ bảo hướng dẫn chúng em trong quá trình học tập.

Cả nhóm đã cố gắng vận dụng những kiến thức đã học được trong thời gian qua để hoàn thành bài tiểu luận. Nhưng do kiến thức hạn chế và không có nhiều kinh nghiệm thực tiễn nên khó tránh khỏi những thiếu sót trong quá trình nghiên cứu và làm bài. Rất kính mong sự góp ý của quý thầy cô để bài tiểu luận của chúng em được hoàn thiện hơn.

Một lần nữa, chúng em xin chân thành cảm ơn, sự quan tâm giúp đỡ của các thầy cô dành cho chúng em trong quá trình thực hiện bài tiểu luận này.

Xin chân thành cảm ơn!

**ĐỒ ÁN ĐƯỢC HOÀN THÀNH**

**TẠI TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP TP HỒ CHÍ MINH**

Tôi xin cam đoan đây là sản phẩm đồ án của chúng tôi và được sự hướng dẫn của thầy **PGS. TS. Huỳnh Trung Hiếu**. Các nội dung nghiên cứu, kết quả trong đề tài này là trung thực và chưa công bố dưới bất kỳ hình thức nào trước đây. Những số liệu trong các bảng biểu phục vụ cho việc phân tích, nhận xét, đánh giá được chính tác giả thu thập từ các nguồn khác nhau có ghi rõ trong phần tài liệu tham khảo.

Ngoài ra, trong đồ án còn sử dụng một số nhận xét, đánh giá cũng như số liệu của các tác giả khác, cơ quan tổ chức khác đều có trích dẫn và chú thích nguồn gốc.

**Nếu phát hiện có bất kỳ sự gian lận nào tôi xin hoàn toàn chịu trách nhiệm về nội dung đồ án của mình.** Trường đại học Công nghiệp TP Hồ Chí Minh không liên quan đến những vi phạm tác quyền, bản quyền do tôi gây ra trong quá trình thực hiện (nếu có).

*TP. Hồ Chí Minh, ngày tháng năm*

*Tác giả*

*(ký tên và ghi rõ họ tên)*

*Nguyễn Thanh Hoài*

*Phạm Chí Nguyên*

*Trần Quang Vinh*

**PHẦN ĐÁNH GIÁ CỦA GIẢNG VIÊN**

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Tp. Hồ Chí Minh, ngày tháng năm

(kí và ghi họ tên)

**Tóm Tắt**

Dự án Virtual Interior Designer là một ứng dụng web thông minh sử dụng công nghệ AI tiên tiến để hỗ trợ thiết kế và chỉnh sửa nội thất. Hệ thống tích hợp bốn công nghệ AI chính: SAM (Segment Anything Model) cho phân đoạn vật thể chính xác, LaMa (Large Mask Inpainting) cho phục hồi nền tự nhiên, Stable Diffusion cho việc tạo nội thất mới, và AI Vision cho việc phân tích và tạo prompt từ ảnh tham khảo.

Ứng dụng được xây dựng trên nền tảng Django 4.2 với giao diện người dùng hiện đại sử dụng Glass Morphism Design và GSAP Animations. Người dùng có thể dễ dàng tải lên ảnh phòng, chọn và xóa vật thể không mong muốn, sau đó thay thế bằng nội thất mới được tạo bởi AI. Hệ thống hỗ trợ hơn 20 loại nội thất khác nhau bao gồm sofa, giường, bàn, đèn, cây cảnh và các vật dụng trang trí.

Kết quả thử nghiệm cho thấy hệ thống có khả năng phân tích ảnh chính xác, tạo ra các prompt chi tiết cho việc sinh nội thất mới, và tạo ra kết quả với chất lượng chuyên nghiệp. Ứng dụng không chỉ giúp người dùng tiết kiệm thời gian trong việc thiết kế nội thất mà còn cung cấp các giải pháp sáng tạo và linh hoạt cho việc tái thiết kế không gian sống.

**Giới Thiệu**

**1. Bối cảnh và động lực**

Trong thời đại công nghệ số phát triển mạnh mẽ, việc ứng dụng trí tuệ nhân tạo (AI) vào lĩnh vực thiết kế nội thất đang trở thành xu hướng tất yếu. Thiết kế nội thất truyền thống thường tốn nhiều thời gian, chi phí và đòi hỏi chuyên môn cao. Điều này tạo ra rào cản lớn cho những người muốn tự thiết kế không gian sống của mình. Trong bối cảnh đó, việc phát triển một giải pháp thông minh hỗ trợ thiết kế nội thất bằng AI trở nên cấp thiết.

**2. Mục tiêu nghiên cứu**

Dự án Virtual Interior Designer được phát triển với các mục tiêu chính sau:

1. Xây dựng một hệ thống thông minh cho phép người dùng dễ dàng chỉnh sửa và thiết kế lại không gian nội thất thông qua giao diện web trực quan.

2. Tích hợp và tối ưu hóa các mô hình AI tiên tiến để:

    - Phân đoạn chính xác các vật thể trong không gian

  - Phục hồi nền tự nhiên sau khi xóa vật thể

    - Tạo ra nội thất mới phù hợp với phong cách và không gian hiện có

3. Phát triển một hệ thống phân tích ảnh thông minh có khả năng:

   - Nhận diện phong cách nội thất

   - Phân tích màu sắc và chất liệu

    - Tạo ra các prompt chi tiết cho việc sinh nội thất mới

1.3. Phạm vi nghiên cứu

Dự án tập trung vào các lĩnh vực chính sau:

1. Phân tích và xử lý ảnh:

   - Phân đoạn vật thể sử dụng SAM (Segment Anything Model)

   - Phục hồi nền với LaMa (Large Mask Inpainting)

   - Xử lý và tối ưu hóa hình ảnh

2. Sinh nội thất thông minh:

   - Tích hợp Stable Diffusion cho việc tạo nội thất mới

   - Phát triển hệ thống prompt engineering

   - Hỗ trợ đa dạng loại nội thất (hơn 20 loại)

3. Giao diện người dùng:

    - Thiết kế UI/UX hiện đại với Glass Morphism

    - Tối ưu hóa trải nghiệm người dùng

1.4. Ý nghĩa thực tiễn

Dự án mang lại những giá trị thực tiễn sau:

1. Cho người dùng:

    - Tiết kiệm thời gian và chi phí trong thiết kế nội thất

    - Dễ dàng thử nghiệm các ý tưởng thiết kế khác nhau

  - Có thể tự thiết kế không gian sống mà không cần chuyên môn sâu

2. Cho ngành thiết kế nội thất:

    - Tự động hóa một số công đoạn thiết kế

    - Cung cấp công cụ hỗ trợ cho các nhà thiết kế chuyên nghiệp

    - Tạo ra các giải pháp thiết kế sáng tạo và độc đáo

3. Cho sự phát triển công nghệ:

  - Đóng góp vào việc ứng dụng AI trong lĩnh vực thiết kế

    - Tạo tiền đề cho các nghiên cứu và phát triển tiếp theo

    - Thúc đẩy sự phát triển của công nghệ thực tế ảo trong thiết kế nội thất

**Phương Pháp**

**1. Kiến trúc hệ thống**

**2.1 Kiến trúc tổng thể**

Hệ thống Virtual Interior Designer được thiết kế theo mô hình 3-tầng (3-tier architecture) bao gồm :

1. **Presentation Layer (Tầng giao diện)**:

- Frontend web application sử dụng Django Templates

- Glass Morphism Design cho UI hiện đại

- GSAP (GreenSock Animation Platform) cho hiệu ứng động

- Bootstrap 5.3 framework cho responsive design

1. **Application Layer (Tầng ứng dụng)**:

- Django 4.2 web framework

- REST API endpoints cho giao tiếp với frontend

- Business logic xử lý yêu cầu người dùng

- Session management và authentication

3. **Data Layer (Tầng dữ liệu)**:

   - SQLite database cho development

- Media storage cho ảnh tải lên

    - Model AI được tích hợp như microservices

**2.2 Luồng xử lí chính**

User → Upload Image → SAM → LaMa → AI Vision → Stable Diffusion → Output

**2. Phương pháp**

**2.1 Phương pháp phân đoạn và xóa vật thể**

**2.1.1 SAM (Segment Anything Model)**

SAM (Segment Anything Model) được sử dụng để thực hiện phân đoạn vật thể chính xác trong ảnh phòng do người dùng tải lên. Mô hình này, phát triển bởi Meta AI, có khả năng nhận diện và phân tách các vật thể trong hình ảnh mà không cần huấn luyện trước cho từng loại vật thể cụ thể. Quy trình thực hiện bao gồm:

1. Nhập liệu : Người dùng chọn vật thể cần phân đoạn thông qua giao diện web
2. Xử lý: SAM sử dụng kiến trúc Transformer để phân tích hình ảnh, tạo ra các mask phân đoạn chi tiết với độ chính xác cao.
3. Kết quả: Các vật thể được phân đoạn thành các vùng độc lập, cho phép người dùng chọn vật thể cần xóa hoặc chỉnh sửa

**2.1.2 Phương pháp lama-inpainting**

Sau khi vật thể được phân đoạn và xóa bởi SAM, LaMa (Large Mask Inpainting) được sử dụng để phục hồi nền một cách tự nhiên. Phương pháp này hoạt động như sau :

1. Nhập liệu: Mask của vật thể đã xóa từ SAM được chuyển đến LaMa.

2. Xử lý: LaMa sử dụng mạng nơ-ron sâu với kiến trúc dựa trên học sâu để tái tạo vùng bị xóa, dựa trên bối cảnh xung quanh (màu sắc, kết cấu, ánh sáng).

3. Kết quả: Vùng nền được phục hồi trông tự nhiên, không để lại dấu vết của vật thể đã xóa.

**2.1.3 Thuật toán GrabCut**

Dùng để phân đoạn vùng, người dùng có thể phân đoạn 1 vùng nào đó cụ thể.

**2.2 Phương pháp phân tích ảnh và sinh nội thất**

**2.2.1 AI Vision cho Image Analysis**

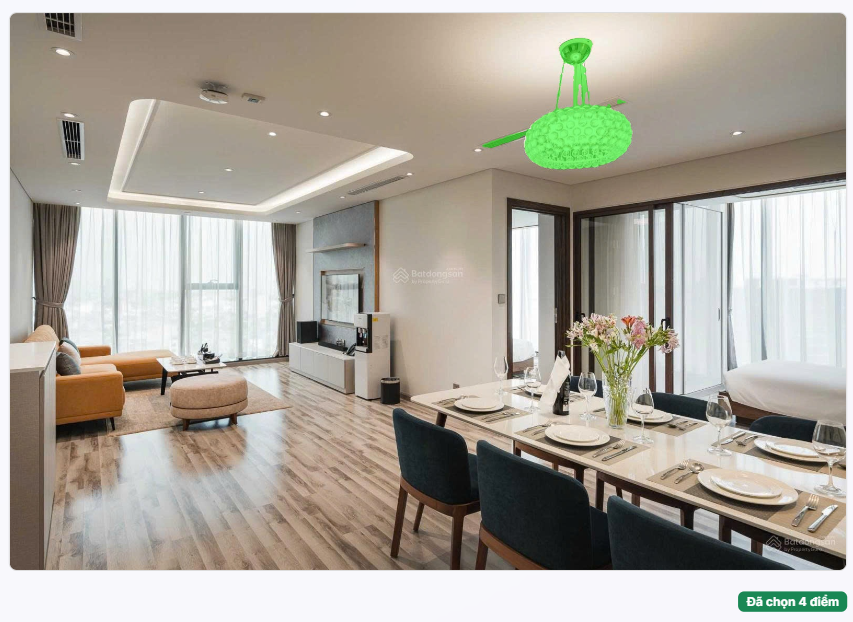
AI Vision được tích hợp để phân tích ảnh phòng nhằm tạo ra các prompt chi tiết cho việc sinh nội thất mới

**2.2.2 Stable-diffusion-inpainting**

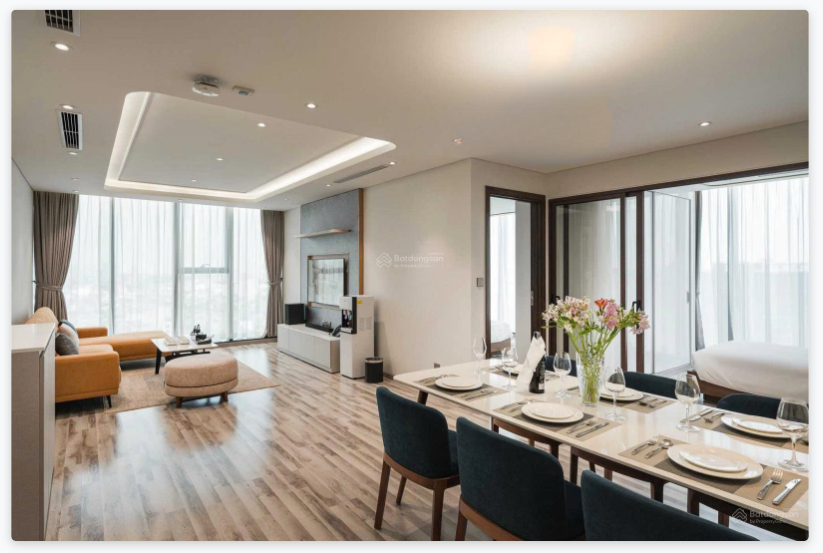
Stable-diffusion-inpainting được sử dụng để tạo nội thất mới dựa trên prompt từ AI Vision hoặc người dùng. Quy trình thực hiện như sau:

1. Nhập liệu: Prompt chi tiết từ AI Vision hoặc người dùng, cùng với ảnh vùng đã phục hồi từ LaMa và mask .
2. Xử lý: Stable Diffusion sử dụng mô hình học sâu dựa trên diffusion để tạo ra hình ảnh nội thất mới, phù hợp với phong cách và không gian của ảnh gốc.
3. Tích hợp không gian: Nội thất mới được chèn vào vùng đã phục hồi, đảm bảo tỷ lệ, ánh sáng, và góc nhìn phù hợp với không gian phòng.

**3. Kết Quả**



Ảnh 1. Phân đoạn vật thể



Ảnh 2. Xóa và phục hồi nền



Ảnh 3. Sinh vật thể dựa trên câu prompt

1. **Thảo luận**
2. **Ưu điểm nổi bật**

+ Tính tự động và linh hoạt : Hệ thống cho phép người dùng thực hiện các thao tác xóa vật thể và phục hồi nền một cách dễ dàng mà không cần chuyên môn nhờ vào mô hình SAM và LAMA

+ Chât lượng hình ảnh : ảnh được giữ nguyên với chất lượng ban đầu

+ Khả năng tạo prompt : Hệ thống AI Vision giúp người dùng nhận diện được prompt chi tiết giúp sinh nội thất

1. **Hạn chế và thách thức**

+ Tốc độ xử lí chưa tối ưu : một số model đòi hỏi thời gian chạy khá lâu. Và yêu cầu tài nguyên GPU lớn

+ Độ chính xác của câu prompt : Trong một số trường hợp, ảnh đầu vào bị mờ, thiếu sáng hoặc bố cục phức tạp có thể khiến hệ thống phân tích sai hoặc tạo ra prompt không phù hợp.

+ Chưa thể sinh ra ảnh đúng với hoàn toàn ảnh đầu vào của người dùng.

**5. Kết Luận**

Dự án **Virtual Interior Designer** đã xây dựng thành công hệ thống thiết kế nội thất ảo ứng dụng AI, giúp người dùng chỉnh sửa và tái tạo không gian sống dễ dàng từ hình ảnh. Hệ thống kết hợp các mô hình SAM, LaMa và Stable Diffusion để xử lý ảnh và sinh nội thất theo phong cách mong muốn. Kết quả cho thấy tiềm năng ứng dụng cao trong thiết kế, dù vẫn còn một số hạn chế cần cải thiện như tốc độ xử lý và độ chính xác khi ảnh đầu vào chưa rõ nét

1. **Tương lai**

Dự án Virtual Interior Designer sẽ được phát triển thêm như sau:

* Tối ưu hóa: Cải thiện tốc độ xử lý, tối ưu tài nguyên GPU.
* Nâng cao chất lượng: Tăng độ chính xác phân tích ảnh, cải thiện prompt engineering, hỗ trợ tùy chỉnh chi tiết nội thất.
* Mở rộng tính năng: Tích hợp VR/AR, kết nối nhà cung cấp nội thất, hỗ trợ thiết kế không gian thương mại.

**7. Tài Liệu Tham Khảo**

[1] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... & Girshick, R. (2023). Segment Anything. *arXiv preprint arXiv:2304.02643*.

[2] Suvorov, R., Logacheva, E., Mashikhin, A., Remizova, A., Ashukha, A., Silvestrov, A., ... & Lempitsky, V. (2022). Resolution-robust Large Mask Inpainting with Fourier Convolutions. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2149-2159.

[3] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., & Ommer, B. (2022). High-Resolution Image Synthesis with Latent Diffusion Models. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 10684-10695.

[4] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... & Amodei, D. (2020). Language Models are Few-Shot Learners. *Advances in Neural Information Processing Systems (NeurIPS)*, 33, 1877-1901..

[5] Rother, C., Kolmogorov, V., & Blake, A. (2004). "GrabCut": Interactive Foreground Extraction using Iterated Graph Cuts. *ACM Transactions on Graphics (TOG)*, 23(3), 309-314.