

TÍCH HỢP MODULE TRÍ TUỆ NHÂN TẠO DALL-E ĐỂ TẠO ỨNG DỤNG WEB XỬ LÝ ẢNH KỸ THUẬT SỐ

Nguyễn Thị Như

Ngày nhận bài: 04/4/2024; Ngày phản biện thông qua: 26/7/2024; Ngày duyệt đăng: 27/7/2024

TÓM TẮT

Với sự phát triển không ngừng của Trí tuệ nhân tạo (AI), nhiều ứng dụng sáng tạo ảnh đã được phát triển dựa trên module DALL_E của OpenAI, các ứng dụng này cho phép người sử dụng tạo ảnh không cần kỹ năng tạo ảnh mà chỉ thông qua các thao tác cơ bản, đơn giản bằng cách mô tả bằng ngôn ngữ tự nhiên về bức ảnh mình mong muốn. Từ đó, chương trình sẽ tạo ra những bức ảnh cho người dùng như những gì người dùng mô tả và theo các phong cách hoạ sĩ cụ thể do người dùng chọn. Nghiên cứu này tập trung vào việc xây dựng các ứng dụng chỉnh sửa và tạo ảnh trên giao diện web trực quan và thân thiện với người dùng, tích hợp module trí tuệ nhân tạo DALL-E để người dùng có thể tạo những bức ảnh từ mô tả ngôn ngữ tự nhiên; cung cấp các chức năng chỉnh sửa ảnh đa dạng, tối ưu hóa hiệu suất và chất lượng, đồng thời đảm bảo tính bảo mật và quyền riêng tư của người dùng. Với mục tiêu này, nghiên cứu sẽ mang lại cho người dùng khả năng tạo ra những tác phẩm chỉnh sửa và sáng tạo ảnh độc đáo, đẩy mạnh sự phát triển và ứng dụng của công nghệ AI trong lĩnh vực xử lý ảnh trên nền tảng web, đồng thời đây cũng là mô hình mẫu cho việc giảng dạy lập trình, tích hợp AI vào website. Kết quả nghiên cứu đã thực hiện được việc tích hợp module DALL_E 2 vào website tự xây dựng, mô hình đã triển khai cho sinh viên học và nghiên cứu việc tích hợp trí tuệ nhân tạo vào website để tạo thành ứng dụng tạo hình ảnh mới, chỉnh sửa ảnh nghệ thuật.

Từ khóa: DALL-E, sáng tạo ảnh, chỉnh sửa ảnh, trí tuệ nhân tạo, OpenAI.

1. MỞ ĐẦU

Trí tuệ nhân tạo (AI) đang trải qua sự phát triển đáng kể và đã được tích hợp vào đa dạng lĩnh vực, mang lại nhiều tiện ích và sự linh hoạt cho cộng đồng người dùng. DALL-E 2 là một mô hình tạo ảnh từ văn bản được phát triển bởi OpenAI. Việc tích hợp AI mà cụ thể là module DALL-E 2 vào website để sáng tạo hình ảnh đã biến nền tảng này thành một trải nghiệm độc đáo, nhanh chóng cập nhật các tiện ích, và thực hiện xử lý thông minh, tạo ra một môi trường tốt nhất cho người sử dụng. DALL-E của OpenAI là phiên bản có 12 tỷ tham số của GPT-3 được đào tạo để tạo hình ảnh từ mô tả văn bản, sử dụng tập dữ liệu gồm các cặp văn bản-hình ảnh. Bao gồm tạo ra các phiên bản động vật và đồ vật được nhân cách hóa, kết hợp các khái niệm không liên quan theo những cách hợp lý, hiển thị văn bản và áp dụng các phép biến đổi cho hình ảnh hiện có (Victoria White 2023).

AI nói chung và DALL-E nói riêng, nhanh chóng được các hãng công nghệ quan tâm và tích hợp xây dựng ứng dụng sáng tạo, tùy chỉnh ảnh, nổi bật như DeepAI, Bing Image Creator, Canva, Dezgo,...

DALL-E, một trong những công nghệ AI tiên tiến, chức năng chính là chuyển văn bản thành hình ảnh theo các phong cách tùy chọn. Bằng cách này, các ứng dụng phát triển trên DALL_E cho

phép người dùng tạo và tương tác với hình ảnh một cách độc đáo.

Sự ra đời của DALL-E nhanh chóng thay đổi cách thức tạo và sử dụng ảnh trên mạng, giờ đây một người bất kỳ, không cần kiến thức hay kỹ năng hội hoạ, chỉ cần vài cú click chuột cũng có thể tạo ra những bức ảnh như mong muốn.

Việc tích hợp DALL-E vào website không chỉ giúp người dùng tạo và chỉnh sửa hình ảnh một cách dễ dàng mà còn góp phần thúc đẩy sự phát triển của công nghệ AI trong lĩnh vực xử lý hình ảnh và cung cấp một nền tảng giáo dục mạnh mẽ cho việc học lập trình và AI.

2. VẬT LIỆU VÀ PHƯƠNG PHÁP

Tổng hợp các tài liệu của thế giới và Việt Nam về sáng tạo hình ảnh trên nền tảng số và ứng dụng trí tuệ nhân tạo trong sáng tạo nội dung ảnh số. Nghiên cứu tài liệu về cách sử dụng API của OpenAI để nhúng DALL-E2 vào trang web.

Triển khai ứng dụng: sử dụng PHP và framework Symfony để xây dựng website. Tích hợp gói phần mềm có các chức năng xử lý hình ảnh, nhúng và kết nối với API của DALL - E2. Kiểm tra và kiểm thử ứng dụng để đảm bảo tính kết nối, ổn định của các chức năng.

Đánh giá hiệu suất và hiệu quả của ứng dụng web trong việc chỉnh sửa và sáng tạo ảnh. So sánh

¹Khoa Khoa học Tự nhiên & Công nghệ, Trường Đại học Tây Nguyên;

Tác giả liên hệ: Nguyễn Thị Như; ĐT: 0906200625; Email: ntnhu@ttn.edu.vn.

kết quả nghiên cứu với các ứng dụng chỉnh sửa ảnh trên thị trường để đánh giá khả năng tích hợp AI trong việc sáng tạo ra ảnh kỹ thuật số từ mô tả văn bản.

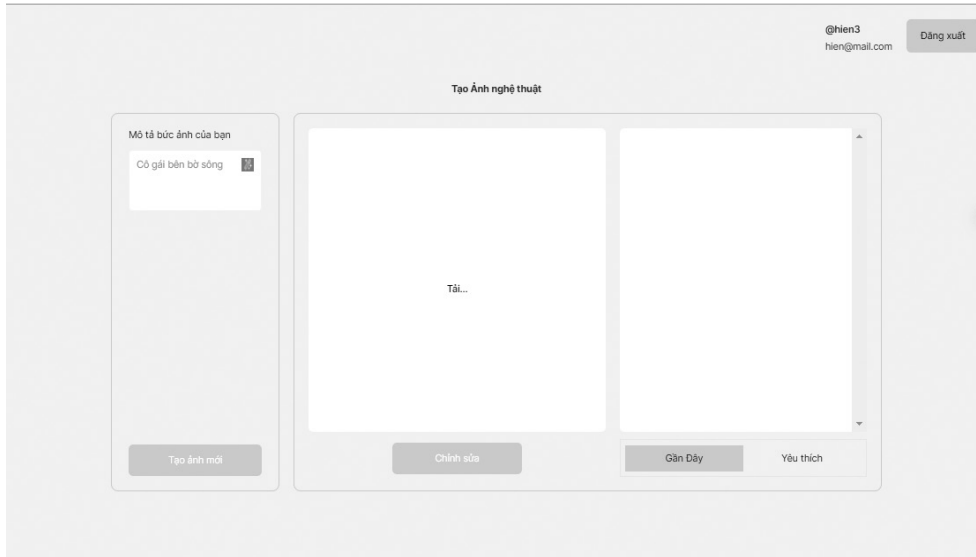
3. KẾT QUẢ VÀ THẢO LUẬN

3.1. Xây dựng website

Website được xây dựng đơn giản với phần giao diện chỉ hiển thị hình ảnh và các liên kết chức năng

(Jon Duckett 2022).

Website gồm hai phần; phần front-end hiển thị dành cho người dùng cuối: cho phép người dùng đăng kí tài khoản, đăng nhập, đăng xuất, thay đổi mật khẩu vào hệ thống; phần back-end dành cho người dùng quản lí bao gồm quản lí tài khoản người dùng, kiểm soát các phần mềm, module được tích hợp vào website.



Hình 1. Giao diện chính của website

Trong giao diện đã xây dựng, hình ảnh được lấy từ mạng Internet để minh họa, các nút lệnh được xây dựng và gắn vào cuối giao diện để người dùng kích hoạt và chuyển tiếp các chức năng khác, chẳng hạn: tải ảnh lên để chỉnh sửa, lưu ảnh về máy người dùng, sáng tạo ảnh.

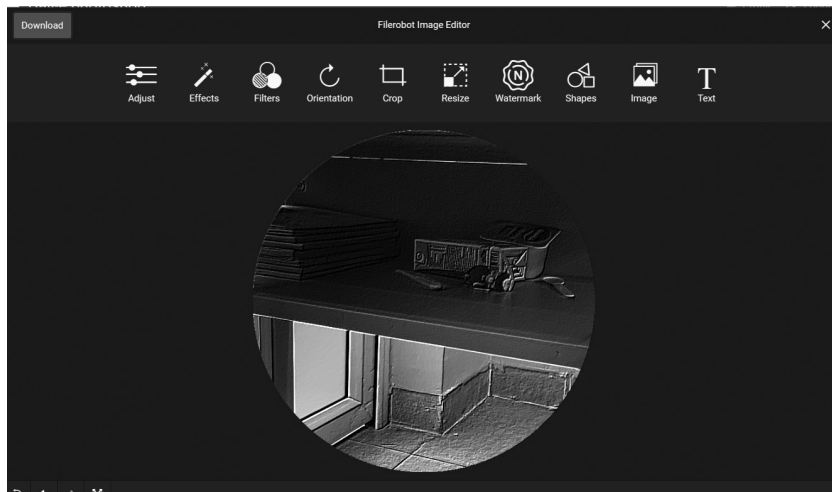
3.2. Tích hợp phần mềm chỉnh sửa ảnh

Sau khi hoàn thiện trang giao diện chính, một phần mềm chỉnh sửa ảnh cần được tích hợp để giúp người dùng có thể chỉnh sửa ảnh đã tạo hoặc ảnh lấy từ nguồn khác.

Phần mềm *Filerobot Image Editor* được chọn

lựa vì đây là phần mềm miễn phí, có đầy đủ các chức năng cơ bản cho việc chỉnh sửa ảnh, gồm: cắt ảnh, thay đổi kích thước ảnh, ảnh chế độ nền (watermark), xoay ảnh, tạo hiệu ứng, thêm họa tiết, thêm văn bản, điều chỉnh các thuộc tính...

Github cung cấp mã nguồn, liên kết cho phép người dùng lấy và đưa vào ứng dụng web. Sau khi nhúng được mã nguồn phần mềm chỉnh sửa ảnh vào web, trang web được trang bị thêm các chức năng tải ảnh lên từ máy tính và tải ảnh xuống sau khi đã chỉnh sửa (*Filerobot Image Editor Demo 2024*).



Hình 2. Giao diện chỉnh sửa ảnh

Các chức năng của website và các chức năng của phần mềm tích hợp đã hoàn thành quy trình chỉnh sửa một ảnh có sẵn.

3.3. Tích hợp DALL-E 2 vào website

DALL-E 2 là một module trí tuệ nhân tạo, có thể tạo ra hình ảnh nghệ thuật chân thực từ mô tả bằng ngôn ngữ tự nhiên. Hệ thống Dall-E 2 sử dụng mô hình GPT-3 giúp nó hiểu văn bản và tạo ra hình ảnh minh họa phức tạp hơn và có thể được sử dụng trong nhiều lĩnh vực khác nhau từ thiết kế đến nghệ thuật sáng tạo. Dall-E 2 sở hữu nhiều tính năng nổi bật có thể kể đến như (DALL – E 2 prompt book 2022):

Tạo hình ảnh từ mô tả văn bản: Dall-E 2 có khả năng tạo ra hình ảnh, hoặc minh họa dựa trên mô tả văn bản. Điều này có nghĩa rằng người dùng có thể mô tả hình ảnh một cách chi tiết, và Dall-E 2 sẽ tạo ra hình ảnh tương ứng. Tuy nhiên, module này yêu cầu đóng phí nên mỗi khi tạo hình ảnh sẽ mất 1 credit.

Tích hợp trí tuệ nhân tạo: Dall-E 2 sử dụng mô hình GPT-3 mạnh mẽ để hiểu văn bản và tạo ra hình ảnh chất lượng cao dựa trên văn bản cung cấp.

Khả năng tạo ra ảnh độc đáo: Dall-E 2 có khả năng tạo ra hình ảnh độc đáo, sáng tạo và thú vị. Nó có thể tạo ra hình ảnh không giới hạn dựa trên mô tả văn bản và không bị giới hạn bởi kiểu dáng hoặc chủ đề.

Chỉnh sửa ảnh trực tiếp: Dall-E 2 cũng cho phép người dùng chỉnh sửa hình ảnh từ máy tính cá nhân từ chức năng “Upload an image to edit”.

Tạo biến thể hình ảnh: Dall-E 2 còn cung



Hình 3. Hình ảnh và các biến thể hình ảnh được tạo ra từ DALL – E 2

Ảnh được tạo từ DALL – E có thể trở thành nguồn ảnh mới, người dùng có thể sử dụng một số công cụ có sẵn trên công cụ cơ bản như di chuyển, kéo thả ảnh, xóa chi tiết, thêm frame và upload ảnh mới, hoặc sử dụng các công cụ từ phần mềm tích hợp khác để thay đổi và biến hình ảnh trở thành ảnh mới cho riêng mình.

4. KẾT LUẬN

Trong giai đoạn bùng nổ thông tin, trí tuệ nhân tạo được ứng dụng rộng rãi trong hầu hết các lĩnh

cấp một tính năng vô cùng ấn tượng khi cho phép người dùng tạo biến thể hình ảnh mới dựa vào tính năng “Variations”. Tức là Dall-E 2 cung cấp nhiều phiên bản hình ảnh mới dựa trên ảnh gốc, người dùng sẽ có thêm nhiều sự lựa chọn thú vị hơn.

Ứng dụng đa dạng: Dall-E 2 đã có thể được sử dụng trong nhiều lĩnh vực khác nhau, bao gồm thiết kế, truyền thông, nghệ thuật sáng tạo, giáo dục và nhiều lĩnh vực khác.

Tích hợp công nghệ tiên tiến: Dall-E 2 được xây dựng trên các công nghệ và mô hình trí tuệ nhân tạo hàng đầu, giúp nó có khả năng tạo ra hình ảnh chất lượng cao và phức tạp.

Bộ lọc nội dung thông minh: OpenAI cũng đưa ra những quy định nghiêm ngặt về nội dung sáng tạo hình ảnh trên OpenAI. Những nội dung yêu cầu sẽ được giám sát thủ công và tự động để ngăn chặn những nội dung không lành mạnh. Từ đó, những nội dung sáng tạo hình ảnh bạo lực, phản cảm sẽ bị cấm ngay lập tức.

Để tích hợp DALL – E 2 vào website, người dùng cần tạo tài khoản thuộc hệ sinh thái AI của OpenAI, đăng kí trả phí để khai thác module DALL – E 2.

Sau khi đã lập trình giao diện và tích hợp công cụ chỉnh sửa, sáng tạo ảnh thì tiến hành lập trình nhúng API vào trang web vừa tạo bằng cách code các JSON để nhúng API của DALL – E vào website.

Đăng nhập vào open AI và nhận mã API của open AI để trang web có thể liên kết với DALL – E 2 cho phép người dùng sử dụng AI để sáng tạo và chỉnh sửa ảnh bằng văn bản hoặc thao tác.

vực của đời sống, sáng tạo nội dung ảnh số theo mô tả hoặc theo các nguồn do người dùng cung cấp đang trở nên phổ biến.

So sánh Dall-E2 với Adobe Firefly: Nhìn chung, có thể thấy cả DALL-E 2 và Adobe Firefly có nhiều đặc điểm giống nhau, đều là những phần mềm, công cụ tạo ảnh AI. Người dùng chỉ cần nhập câu lệnh, phần mềm sẽ sử dụng trí tuệ nhân tạo để xử lý và sáng tạo hình ảnh dựa trên mô tả được cung cấp. Cả hai phần mềm đều cho ra những

hình ảnh rất sáng tạo và cuốn hút. DALL-E 2 tuy có kém hơn một chút về những tính năng đi kèm nhưng công cụ này lại được tích hợp khá nhiều trong đa số những công cụ AI trên thị trường hiện nay như Tome AI để người dùng nhanh chóng tạo hình ảnh ngay trên những tác vụ khác mà không phải mất thêm thời gian chuyển tác vụ. Một công cụ tiện ích có thể tích hợp và áp dụng nhiều trong mọi trường hợp (Victoria White 2023).

Tuy nhiên, Adobe Firefly tốt hơn Dall-E 2 là do kho tàng hình ảnh khổng lồ cùng nhiều tính năng thú vị đi kèm công cụ. Người dùng có thể áp dụng trực tiếp những tính năng này để điều chỉnh hình ảnh trước khi tải về. Chúng ta dễ thấy rằng những tính năng này khá đầy đủ và có nhiều tùy chọn hơn để người dùng ứng dụng, như kích thước khung hình, style hình ảnh, tỷ lệ, phối màu, sắp xếp bố cục,...và còn có cả Generative AI để làm nhiều tác vụ khác như xóa đối tượng, xóa background, thêm đối tượng hay thay đổi phong nền,...

Nghiên cứu đã thực hiện và triển khai thử nghiệm mô hình sáng tạo nội dung ảnh số từ nguồn văn bản mô tả bằng ngôn ngữ tự nhiên do người dùng cung cấp. Kết quả đã thực hiện hoàn chỉnh

hệ thống cho phép người dùng tạo ảnh, chỉnh sửa ảnh và lưu thành ảnh cá nhân. Bằng phương thức này, người dùng được phép sản xuất hàng loạt hình ảnh trong nhiều lĩnh vực khác nhau. Sự sáng tạo trong từng hình ảnh còn phụ thuộc vào sự mô tả của người sử dụng hệ thống thông qua khai báo các tham số và mô tả yêu cầu của người biên tập bài viết.

Sau khi thử nghiệm, mô hình chạy ổn định và đảm bảo tính hợp lý như quy định của OpenAI, cùng với phần mềm chỉnh sửa ảnh được tích hợp, sản phẩm trở thành ứng dụng đáp ứng hầu hết các tính năng tạo ảnh, chỉnh sửa ảnh. Tuy nhiên, khi thực hiện chạy chức năng tạo ảnh mới nhiều lần trên cùng một nguồn mô tả thì nội dung của các hình ảnh mới không có nhiều khác biệt. Điều này thể hiện sự ổn định về chất lượng hình ảnh, tính hiệu đúng đắn mô tả của module DALL – E 2 nhưng hạn chế sự sáng tạo của sản phẩm được tạo ra. Để mô hình này tiếp tục phát triển theo hướng sáng tạo tốt hơn thì mô hình cần thời gian huấn luyện dài hơn, các tham số cần được cải tiến và người dùng cũng phải học cách mô tả yêu cầu tốt hơn cho mỗi lần tạo sử dụng chức năng tạo ảnh.

INTEGRATING DALL-E ARTIFICIAL INTELLIGENCE MODULE TO CREATE A WEB APPLICATION FOR DIGITAL IMAGE PROCESSING

Nguyen Thi Nhu¹

Received Date: 04/4/2024; Revised Date: 26/7/2024; Accepted for Publication: 27/7/2024

ABSTRACT

With the continuous development of Artificial Intelligence (AI), many creative image applications have been developed based on OpenAI's DALL-E module. Users without any image creation skills, just type some description the desired image in natural language and choose specific artist styles, these applications will generate expected images in a moment. This research focuses on building image editing and creation applications on an intuitive and user-friendly web interface, integrating the DALL-E AI module so that users can create images from natural language descriptions; providing diverse image editing functions, optimizing performance and quality, while ensuring user security and privacy. As a result, the research will offer users the ability to create unique edited and creative images, promoting the development and application of AI technology in the field of image processing on web platforms. In addition, it serves as a model for teaching programming and integrating AI into websites. The research

¹Faculty of Natural Science and Technology, Tay Nguyen University;

Corresponding author: Nguyen Thi Nhu; Tel: 0906200625; Email: ntnhu@ttn.edu.vn.

results have achieved the integration of the DALL-E 2 module into a website, and the model has been deployed for students to learn and research the integration of AI into websites to create new image creation and artistic editing applications.

Keywords: *DALL-E, photo creation, photo editing, artificial intelligent, OpenAI.*

TÀI LIỆU THAM KHẢO

Jon Duckett (2022). PHP & MySQL Server-side Web Development. *Wiley publisher*. ISBN: 978-1119149224.

Oliver Theobald (2022, updated 2023). Generative AI Art: A Beginner's Guide to 10x Your Output with Smart Text Prompts. ISBN: 979-8366143608

Victoria White (2023). Adobe Firefly: 100+ Prompts To Ignite Your Everyday Creativity. ISBN: 979-8864499252.

DALL-E 2 prompt book (2022). Xem 22/03/2024. <<https://dallery.gallery/wp-content/uploads/2022/07/The-DALL%E2%B7E-2-prompt-book-v1.02.pdf>>

Filerobot Image Editor Demo (2024). Xem 10/03/2024. <https://scaleflex.github.io/filerobot-image-editor>.