



THÔNG TIN CHUNG CỦA NHÓM

- Link YouTube video của báo cáo: <https://youtu.be/IngGZ3gRdOI>
- Link slides:
<https://github.com/ngtuthanhan/CS519.N11/blob/main/Presentation.pdf>

<ul style="list-style-type: none">• Họ và Tên: Nguyễn Tư Thành Nhân• MSSV: 20520079 	<ul style="list-style-type: none">• Lớp: CS519.N11• Tự đánh giá (điểm tổng kết môn): 9,5/10• Số buổi vắng: 0• Số câu hỏi QT cá nhân: 10• Số câu hỏi QT của cả nhóm: 3• Link Github: https://github.com/ngtuthanhan/CS519.N11• Mô tả công việc và đóng góp của cá nhân cho kết quả của nhóm:<ul style="list-style-type: none">○ Đề xuất bài toán.○ Lên ý tưởng cho đề án.○ Làm Slide báo cáo.○ Chỉnh sửa báo cáo, slide.
<ul style="list-style-type: none">• Họ và Tên: Cao Văn Hùng• MSSV: 20520193 	<ul style="list-style-type: none">• Lớp: CS519.N11• Tự đánh giá (điểm tổng kết môn): 9/10• Số buổi vắng: 1• Số câu hỏi QT cá nhân: 10• Số câu hỏi QT của cả nhóm: 3• Link Github: https://github.com/hungcao0402/CS519.N11• Mô tả công việc và đóng góp của cá nhân cho kết quả của nhóm:

	<ul style="list-style-type: none"> ○ Khảo sát các nghiên cứu liên quan, viết giới thiệu và mục tiêu. ○ Đóng góp ý tưởng phần Nội dung và Phương pháp thực hiện. ○ Viết báo cáo đồ án
<ul style="list-style-type: none"> ● Họ và Tên: Lê Nhật Minh ● MSSV: 20520070 	<ul style="list-style-type: none"> ● Lớp: CS519.N11 ● Tự đánh giá (điểm tổng kết môn): 9/10 ● Số buổi vắng: 4 ● Số câu hỏi QT cá nhân: 9 ● Số câu hỏi QT của cả nhóm: 3 ● Link Github: https://github.com/minh1409/CS519.N11 ● Mô tả công việc và đóng góp của cá nhân cho kết quả của nhóm: <ul style="list-style-type: none"> ○ Làm Poster ○ Kiểm tra các lỗi văn bản, lỗi trình bày

ĐỀ CƯƠNG NGHIÊN CỨU

TÊN ĐỀ TÀI (IN HOA)

POLY-GAN CHO HỆ THỐNG THỬ TRANG PHỤC TRỰC TUYẾN

TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)

POLY-GAN FOR INTERACTIVE VIRTUAL TRY-ON SYSTEM

TÓM TẮT

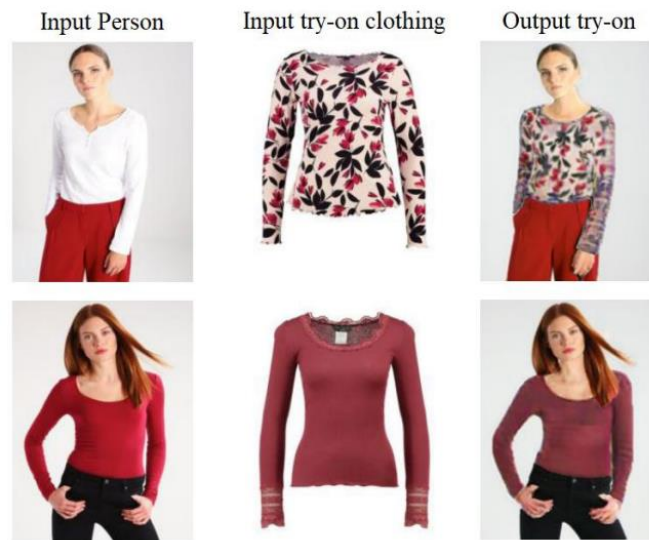
Trong thời đại của công nghệ hiện nay, nhiều ngành công nghiệp đang trải qua sự phát triển nhanh chóng, đặc biệt là ngành thời trang hướng theo mong muốn và nhu cầu của người tiêu dùng. Một trong những yêu cầu của khách hàng là có thể thử quần áo trực tuyến trước khi mua chúng. Để giải quyết vấn đề này, các giải pháp thường sử dụng các mô hình Generative Adversarial Network (GAN). Tuy nhiên, những giải pháp này gặp phải hạn chế khi các vùng cơ thể bị che phủ khiến cho trang phục sẽ bị mất hoặc mờ đi trong ảnh đầu ra. Để giải quyết vấn đề này, Poly-GAN đã được đề xuất, với cùng một kiến trúc GAN được sử dụng cho tất cả ba giai đoạn xử lý và một số thay đổi lớn trong hoạt động của từng giai đoạn. Trong đề tài này, nhóm tập trung nghiên cứu về sử dụng PolyGAN để đánh giá và chứng minh tính hiệu quả của phương pháp tiếp cận của chúng tôi đối với bài toán thử quần áo trực tuyến dựa trên hình ảnh người dùng.

GIỚI THIỆU

Ngày nay, người tiêu dùng ngày càng quan tâm đến việc mua sắm trực tuyến. Tuy nhiên, việc mua sắm online đã nảy sinh ra một trở ngại đối với những người muốn mua sắm trang phục là: Người mua không có cơ hội thử trước những trang phục mà mình muốn mua để biết được trang phục đó có thật sự hợp với cơ thể, vóc dáng của mình hay không giống như việc mua sắm trực tiếp. Vì vậy, để giúp cho khách hàng có được một đánh giá khách quan về sự phù hợp của trang phục với bản thân khi mua sắm online, bài toán thử quần áo ảo dựa trên hình ảnh (Image-based Virtual try-on)[2] đã ra đời.

Với bài toán này, dữ liệu đầu vào sẽ bao gồm bức ảnh cá nhân người dùng - source model (bức ảnh này phải bao gồm phần cơ thể tương ứng với trang phục cần mặc thử) và bức ảnh

trang phục cần mặc thử - reference garment. Đầu ra sẽ là bức ảnh mà người sử dụng đang mặc trang phục đó.



Hình 1. *Input/Output của bài toán Virtual try-on*

Có nhiều giải pháp đã được đề xuất để giải quyết bài toán này và hầu hết các giải pháp đó đều sử dụng các mô hình Generative Adversarial Network (GAN) [4] [5] [6] và gồm ba giai đoạn. Giai đoạn 1 sử dụng phép biến đổi affine để căn chỉnh trang phục cần mặc thử cho khớp với tư thế của người trong source model. Giai đoạn 2 dán trang phục lên phần cơ thể tương ứng của người trong source model. Giai đoạn 3 là giai đoạn tinh chỉnh/hậu xử lý để giảm thiểu các tạo tác (artifacts) phát sinh sau giai đoạn 2.

Tuy nhiên, những giải pháp sử dụng các mô hình GAN có pipeline như trên gặp phải vấn đề là: Khi có vùng cơ thể (ví dụ như tóc, cánh tay, bàn tay) che đi phần trang phục trong source model thì vùng cơ thể đó sẽ bị mất hoặc bị mờ đi ở bức ảnh đầu ra.

Đối với hạn chế này, một mô hình GAN mới đã được đề xuất để giải quyết nó. Mô hình mới này, với tên gọi Poly-GAN [1], chỉ sử dụng một kiến trúc GAN chung cho cả 3 giai đoạn xử lý (1, 2 và 3) thay vì sử dụng 3 networks khác nhau cho 3 giai đoạn như các mô hình trước đó, đi cùng với đó là một số thay đổi lớn trong hoạt động của từng giai đoạn. Nhờ những thay đổi này mà Poly-GAN đã đạt được những kết quả đáng kể trong việc khắc phục vấn đề trên của

các giải pháp trước.

Trong khuôn khổ đề tài này, nhóm chúng em xin tập trung nghiên cứu về sử dụng PolyGAN trong việc giải quyết bài toán Image-based Virtual try-on.

MỤC TIÊU

- Tìm hiểu bài toán Image-based Virtual Try-on và cách xây dựng bộ dữ liệu huấn luyện, kiểm thử.
- Tìm hiểu và phân tích mô hình Poly-GAN. Từ đó áp dụng mô hình Poly-GAN để huấn luyện trên bộ dữ liệu đã được xây dựng và đánh giá kết quả đạt được.
- Xây dựng ứng dụng minh họa dựa trên nền tảng web.

NỘI DUNG VÀ PHƯƠNG PHÁP

Nội dung 1. Tìm hiểu bài toán Image-based Virtual Try-on và cách xây dựng bộ dữ liệu huấn luyện, kiểm thử.

Phương pháp thực hiện:

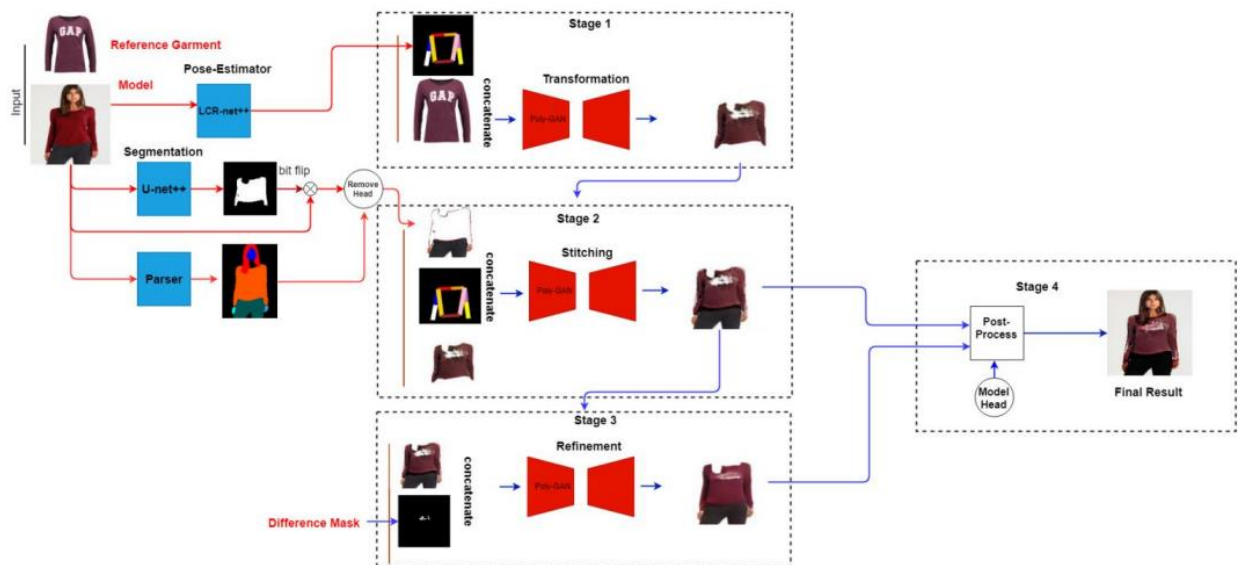
- Tìm hiểu về bài toán Image-based Virtual try-on.
- Tìm hiểu các độ đo dùng để đánh giá cho bài toán Image-based Virtual try-on: Structure Similarity Index Measure (SSIM) và Inception Score (IS).
- Tìm hiểu cách xây dựng bộ dữ liệu huấn luyện và kiểm thử cho mô hình từ bộ dữ liệu DeepFashion.

Nội dung 2. Tìm hiểu và phân tích mô hình Poly-GAN. Từ đó áp dụng mô hình Poly-GAN để huấn luyện trên bộ dữ liệu đã được xây dựng và đánh giá kết quả đạt được.

Phương pháp thực hiện:

- Tìm hiểu kiến trúc Generative Adversarial Network cho bài toán Image-based Virtual try-on
- Tài liệu chi tiết mô hình Poly-GAN, các kỹ thuật liên quan và các pretrained model được sử dụng (LCR-net++[7], U-Net++[8]).
- Cài đặt, huấn luyện, mô hình trên bộ dữ liệu đã được xây dựng từ DeepFashion và đánh giá kết quả nhận được dựa trên các độ đo SSIM và IS.

- Đánh giá, so sánh kết quả của mô hình Poly-GAN với các mô hình đã có (VITON[5], CP-VITON[6]) trên cùng bộ dataset đã được xây dựng từ DeepFashion.



Hình 2. Poly-GAN pipeline. Giai đoạn 1: Chuyển đổi trang phục sử dụng Poly-GAN với đầu vào là khung RGB của mô hình và áo quần tham chiếu. Giai đoạn 2: Dán áo quần sử dụng Poly-GAN với đầu vào là mô hình phân đoạn, khung RGB của mô hình và trang phục đã biến đổi. Giai đoạn 3: Tinh chỉnh để lấp đầy lỗ sử dụng Poly-GAN với đầu vào là ảnh đã được dán áo quần và mask cho biết vùng bị thiếu. Giai đoạn 4: Hậu xử lý để kết hợp output của giai đoạn 2 và giai đoạn 3 với đầu của source model).

Nội dung 3. Xây dựng ứng dụng minh họa dựa trên nền tảng web.

Phương pháp thực hiện:

- Tìm hiểu và xây dựng hệ thống minh họa trên nền tảng web cho phép người dùng tải lên hình ảnh về trang phục và cơ thể người. Với kết quả trả về sẽ là cơ thể người trong bộ trang phục đó.
- Cho phép người dùng kiểm soát các khía cạnh khác nhau của quy trình tạo hình ảnh, chẳng hạn như màu sắc, hoa văn và phong cách, để tạo hình ảnh thời trang tùy chỉnh.

KẾT QUẢ MONG ĐỢI

- Tài liệu về bài toán Image based Virtual Try-on, tài liệu về độ đo SSIM, IS. Tài liệu về dataset DeepFashion.

- Source code cho mô hình Poly-GAN đã được tìm hiểu và chú thích. Bảng kết quả đánh giá mô hình Poly-GAN trên bộ dataset DeepFashion.
- Ứng dụng minh họa trên nền tảng web cho phép người dùng mặc thử áo quần trực tuyến.

TÀI LIỆU THAM KHẢO

- [1] N. Pandey and A. Savakis, “Poly-gan: Multi-conditioned gan for fashion synthesis,” 2019. [Online]. Available: <https://arxiv.org/abs/1909.02165>
- [2] H. Ghodhbani, A. Alimi, and M. Neji, “Image-based virtual try-on system: A survey of deep learning-based methods,” 02 2021
- [3] F. Kınılı, B. Ozcan, and F. Kırac, “A benchmark for inpainting of ” clothing images with irregular holes,” in *Computer Vision – ECCV 2020 Workshops*, A. Bartoli and A. Fusiello, Eds. Cham: Springer International Publishing, 2020, pp. 182–199.
- [4] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, “Generative adversarial nets,” in *Advances in Neural Information Processing Systems*, Z. Ghahramani, M. Welling, C. Cortes, N. Lawrence, and K. Weinberger, Eds., vol. 27. Curran Associates, Inc., 2014. [Online]. Paper.pdf
- [5] X. Han, Z. Wu, Z. Wu, R. Yu, and L. S. Davis, “Viton: An image-based virtual try-on network,” in *CVPR*, 2018
- [6] B. Wang, H. Zheng, X. Liang, Y. Chen, and L. Lin, “Toward characteristic-preserving image-based virtual try-on network,” in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 589–604
- [7] G. Rogez, P. Weinzaepfel, and C. Schmid, “LCR-Net: Localization Classification Regression for Human Pose,” in *CVPR*, 2017
- [8] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, “Unet++: A nested u-net architecture for medical image segmentation,” in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Springer, 2018, pp. 3–11.