

TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI

ĐỒ ÁN TỐT NGHIỆP

Ứng dụng mô hình GANs giải bài toán dịch ảnh đa miền

NGUYỄN HỮU MINH

minh.nh152464@sis.hust.edu.vn

Ngành: Toán Tin

Chuyên ngành: Toán Tin

Giảng viên hướng dẫn: ThS. Nguyễn Tuấn Dũng

Bộ môn: Toán Tin

Viện: Toán ứng dụng và Tin học

Chữ ký của GVHD

Hà Nội, 06/2020

NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN

1. Mục đích và nội dung của đề án:

2. Kết quả đạt được:

3. Ý thức làm việc của sinh viên:

Hà Nội, ngày tháng năm
Giảng viên hướng dẫn

Lời cảm ơn

Thành công nào cũng đều gắn liền với sự hỗ trợ của những người xung quanh, dù cho là ít hay nhiều, trực tiếp hay gián tiếp. Trong suốt thời gian làm Đồ án, em đã nhận được sự quan tâm, giúp đỡ của thầy cô, gia đình, anh chị và bạn bè xung quanh.

Với tấm lòng biết ơn vô cùng sâu sắc, em xin gửi lời cảm ơn chân thành nhất từ đáy lòng đến quý Thầy Cô của Viện Toán ứng dụng và Tin học, trường Đại học Bách Khoa Hà Nội đã tạo điều kiện và cùng dùng những tri thức, tâm huyết của mình để có thể truyền đạt cho em trong vốn kiến thức quý báu.

Đặc biệt, em xin chân thành cảm ơn ThS. Nguyễn Tuấn Dũng đã tận tâm chỉ bảo hướng dẫn em trong suốt thời gian làm Đồ án qua. Nhờ có những lời hướng dẫn, dạy bảo của thầy mà Đồ án của em đã hoàn thành một cách tốt nhất. Một lần nữa, em xin gửi lời cảm ơn chân thành đến thầy.

Đồ án của em không tránh khỏi những thiếu sót, em rất mong nhận được ý kiến đóng góp của quý Thầy Cô và các bạn học để Đồ án của em được hoàn thiện hơn.

Em xin chân thành cảm ơn!

Tóm tắt nội dung đồ án

Cách mạng công nghiệp 4.0 đã mang đến cho con người một kỷ nguyên khai phá dữ liệu. Điều này đặt ra bài toán không chỉ làm sao để khai phá được dữ liệu một cách hiệu quả mà còn làm sao để sinh ra được thêm nhiều dữ liệu một cách tự động và với số lượng lớn. Do đó, trong khuôn khổ của đồ án tốt nghiệp, em sẽ nghiên cứu về ý tưởng chung, kiến trúc, hàm loss, phương pháp đánh giá, những vấn đề tồn đọng và cách giải quyết các vấn đề đó của mô hình Generative Adversarial Networks (GANs) tổng quát nhằm giải quyết bài toán sinh dữ liệu ảnh nói và mô hình Multimodal Unsupervised Image-to-Image Translation (MUNIT) giúp giải quyết bài toán sinh dữ liệu ảnh mới từ ảnh đã có sẵn (Image-to-Image Translation). Bên cạnh những kết quả đã được công bố trong bài báo, em đã thử nghiệm mô hình với bộ dữ liệu khác để chứng minh tính đúng đắn của mô hình và lấy đó làm tiền đề cho việc phát triển những mô hình phức tạp hơn trong tương lai.

Hà Nội, ngày tháng năm
Sinh viên thực hiện

Mục lục

Lời cảm ơn	1
Tóm tắt nội dung đề án	2
Danh sách hình vẽ	4
Danh sách bảng	5
Kết luận và phương hướng phát triển	7
Phụ lục	8
Tài liệu tham khảo	9

Danh sách hình vẽ

Danh sách bảng

hello content [1]

Kết luận và phương hướng phát triển

Các mô hình dựa trên ý tưởng về Generative Adversarial Networks đã tạo ra những kết quả vô cùng ấn tượng trong việc sinh ảnh hay cải thiện chất lượng ảnh. Những kết quả này mở ra một tương lai trong việc tạo ra thật nhiều dữ liệu nhằm phục vụ cho các mô hình Deep Learning khác. Hơn nữa, với những thành công trong việc sinh ảnh, các mô hình GANs hoàn toàn có tiềm năng trong việc sinh ra các loại dữ liệu khác như video, âm thanh ... trong tương lai.

Mô hình Multimodal Unsupervised Image-to-Image Translation (MUNIT) đã để lại những kết quả rất ấn tượng trong việc giải quyết bài toán Image-to-Image Translation và cụ thể là Style Transfer. Kiến trúc quen thuộc theo dạng auto-encoder của MUNIT được điều chỉnh sao cho việc biến đổi các đặc trưng của ảnh hiệu quả nhất. Bên cạnh đó, hệ thống các hàm loss của MUNIT với các mục đích khác nhau đã định hướng cho mô hình, giúp giải quyết các vấn đề về unsupervised learning, multimodal. Những kết quả từ trong bài báo của tác giả đã minh chứng cho tính khả thi của MUNIT đối với nhiều bộ dữ liệu khác nhau. Ngoài ra, thử nghiệm với bộ dữ liệu ảnh chân dung như CelebA-HQ mang đến một kỳ vọng trong việc xử lý những bộ dữ liệu khó (liên quan đến con người), xử lý những style mang tính trừu tượng hơn.

Bên cạnh những kết quả có lợi đó, MUNIT cũng còn một số điểm yếu cần phải khắc phục như xử lý với ảnh có kích thước lớn hơn, kích thước mô hình vẫn rất lớn đòi hỏi yêu cầu phần cứng cao, số lượng các phép tính toán nhiều khiến việc luyện mô hình mất khá nhiều thời gian.

Phụ lục

Tài liệu tham khảo

[1] test name. test title. *test journal*, test year.