

Image to Image

Thái Thị Hiền* và Trịnh Ngọc Đức*

*Trường đại học Công Nghiệp thành phố Hồ Chí Minh

Môn học: Thị giác máy tính
Giáo viên hướng dẫn: Huỳnh Trung Hiếu, Lưu Giang Nam



Nội dung

- 1 Giới thiệu đề tài
- 2 Data
- 3 Mô hình
- 4 Đánh giá
- 5 Kết luận
- 6 Hướng phát triển trong tương lai

Nội dung

- 1 Giới thiệu đề tài
- 2 Data
- 3 Mô hình
- 4 Đánh giá
- 5 Kết luận
- 6 Hướng phát triển trong tương lai

Giới thiệu

Mục tiêu bài toán: Sử dụng CGAN để dịch hình ảnh qua hình ảnh



Semantically Labeled Image



Generated Image

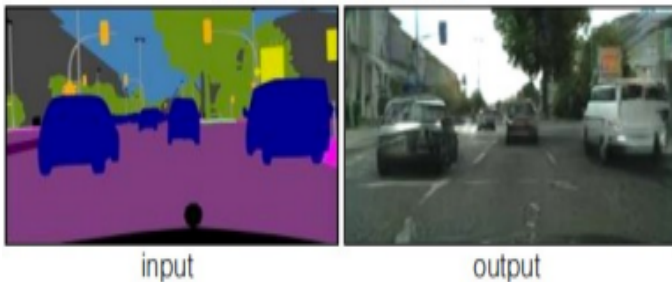
Hình 1: Tổng quan bài toán

Nội dung

- 1 Giới thiệu đề tài
- 2 Data**
- 3 Mô hình
- 4 Đánh giá
- 5 Kết luận
- 6 Hướng phát triển trong tương lai

Data

- Đây là bộ dữ liệu cho mô hình pix2pix nhằm mục đích hoạt động như một giải pháp đa năng cho các vấn đề dịch từ hình ảnh sang hình ảnh.
- Trong bộ dữ liệu này nhóm em đã sử dụng 3 bộ dữ liệu nhỏ gồm: Cityscapes, Edges2shoes, Maps để đào tạo mô hình



Hình 2: Dữ liệu phân vùng ảnh

Data

- Edget to Photo là bài toán chuyển các hình ảnh vẽ tay sang hình ảnh thật, từ các bộ phận cạnh của dữ liệu đã học(mèo, túi xách, giày ,...).



Hình 3: Dữ liệu về ảnh phác thảo

Data

- Aerial to Map là bài toán ứng dụng trong google map.



Hình 4: Dữ liệu về ảnh bản đồ

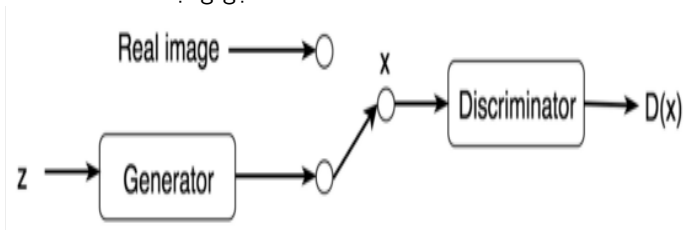
Nội dung

- 1 Giới thiệu đề tài
- 2 Data
- 3 Mô hình**
- 4 Đánh giá
- 5 Kết luận
- 6 Hướng phát triển trong tương lai

GAN

- GAN là viết tắt “generative adversarial network”, hướng tới việc sinh ra dữ liệu mới sau quá trình học.

GAN được cấu thành từ 2 mạng gọi là Generator và Discriminator.



Hình 5: Kiến trúc mạng Gan

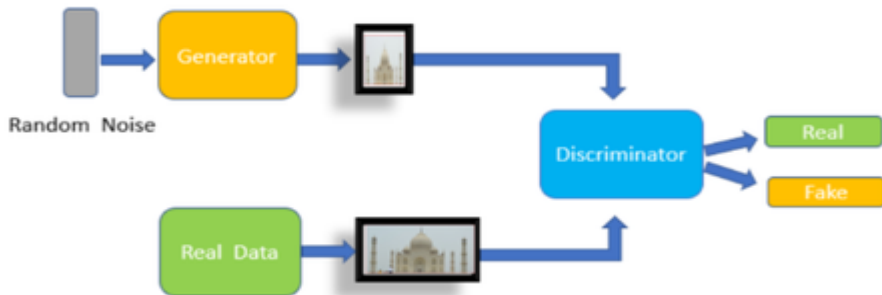
CGAN

- CGAN: là một loại GAN liên quan đến việc tạo hình ảnh có điều kiện bằng mô hình trình tạo.
- Hàm mục tiêu của GAN có điều kiện có thể được biểu thị bằng:

$$L_{cGAN}(G, D) = E_x[\log D(x)] + E_z[\log(1 - D(G(z)))]$$

Generator và Discriminator

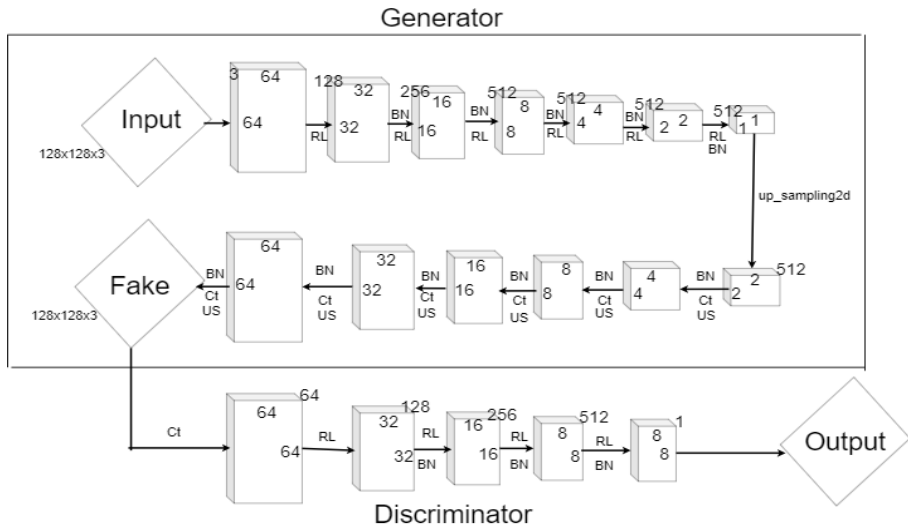
- Generator là mạng sinh ra dữ liệu
- Discriminator là mạng để phân biệt xem dữ liệu là thật (dữ liệu từ dataset) hay giả (dữ liệu sinh ra từ Generator)



Kiến trúc mạng trong Generator

- Explicit model: Mô hình hiện, một dạng của Generative model sử dụng các hàm phân phối xác suất lý thuyết để sinh mẫu dữ liệu.
- Implicit model: Mô hình ẩn, một dạng khác của Generative model, không sử dụng phân phối xác suất lý thuyết mà thay vào đó, mẫu sinh được sinh ra từ mô hình. GAN là một lớp mô hình như vậy.

CGan



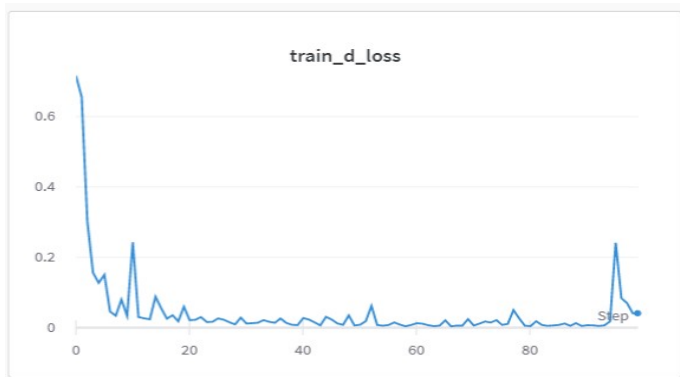
Hình 6: Kiến trúc mạng CGan

Nội dung

- 1 Giới thiệu đề tài
- 2 Data
- 3 Mô hình
- 4 Đánh giá**
- 5 Kết luận
- 6 Hướng phát triển trong tương lai

Đánh giá

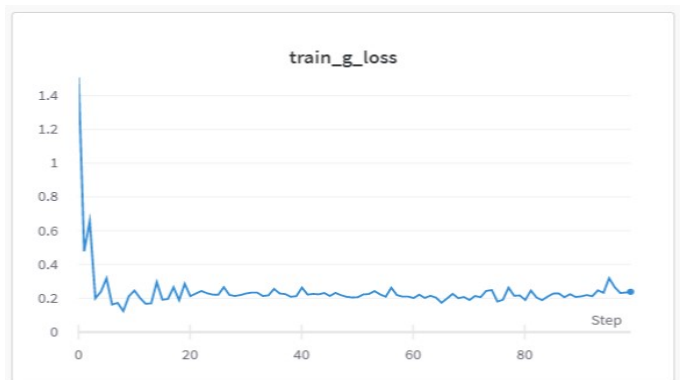
Cityscapes



Giá trị loss trên tập train discriminator Cityscapes

Đánh giá

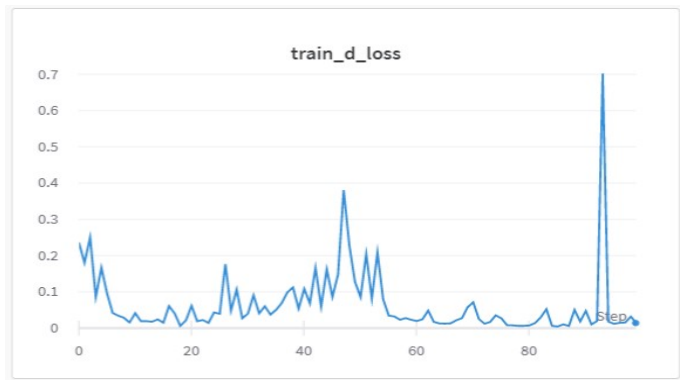
Cityscapes



Giá trị loss trên tập train generator Cityscapes

Đánh giá

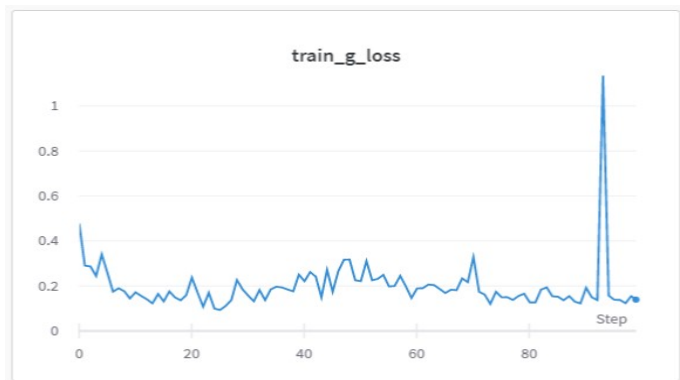
Edges2shoes



Giá trị loss trên tập train discriminator Edges2shoes

Đánh giá

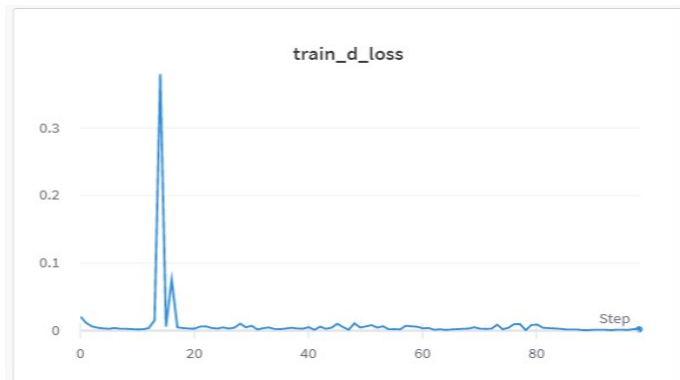
Edges2shoes



Giá trị loss trên tập train generator Edges2shoes

Đánh giá

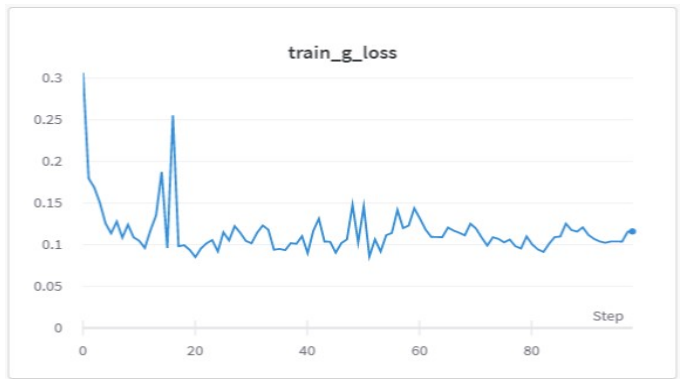
Maps



Giá trị loss trên tập train discriminator Map

Đánh giá

Maps



Giá trị loss trên tập train generator Map

Đánh giá

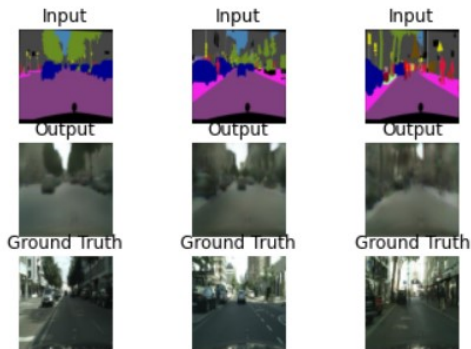
- Cuối cùng mô hình cho ra kết quả như sau:
- + Đầu tiên là về điểm dự đoán được

Data	D_loss	G_loss	Accuracy
Cityscapes	0.04109	0.23862	98% - 99%
Edges2shoes	0.01404	0.13925	98% - 99%
Maps	0.001749	0.111907	98% - 99%

Điểm số của các mô hình của 3 tập dữ liệu khác nhau

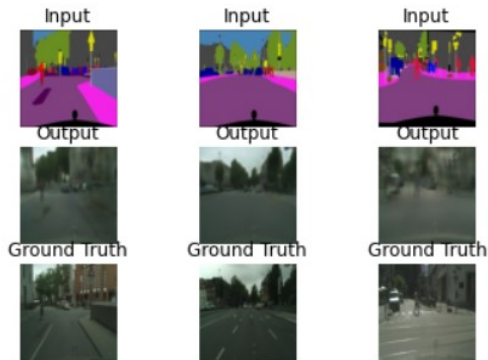
Đánh giá

+ Thứ hai là kết quả sau khi chạy mô hình của tập dữ liệu Cityscapes



Kết quả hình ảnh

Đánh giá



Kết quả hình ảnh

Đánh giá

+ Kết quả sau khi chạy mô hình của tập dữ liệu Edges2shoes



Kết quả hình ảnh

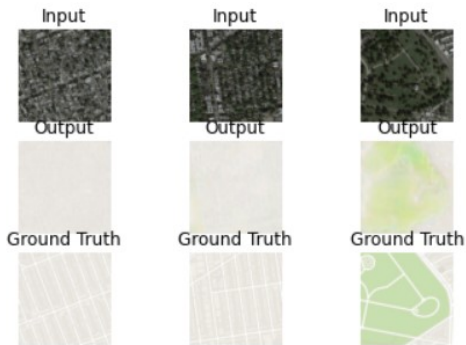
Đánh giá



Kết quả hình ảnh

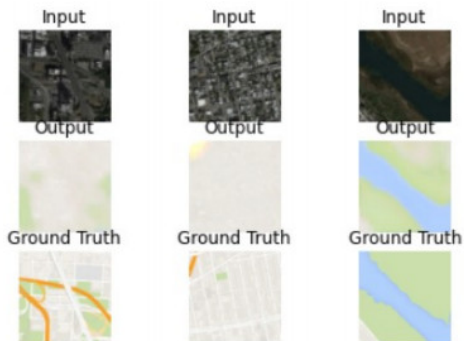
Đánh giá

+ Kết quả sau khi chạy mô hình của tập dữ liệu Maps



Kết quả hình ảnh

Đánh giá



Kết quả hình ảnh

Nội dung

- 1 Giới thiệu đề tài
- 2 Data
- 3 Mô hình
- 4 Đánh giá
- 5 Kết luận**
- 6 Hướng phát triển trong tương lai

Kết luận

1. Dựa vào bài đánh giá có thể thấy được qua 3 tập dữ liệu độ chính xác của mô hình từ 0.98 - 0.99
 2. Sai số của mô hình discriminator tương đối thấp từ 0.0017 - 0.04
 3. Sai số của mô hình generator từ 0.11 - 0.23
- ==> Qua 3 tập dữ liệu Cityscapes, Edges2shoes, thì 2 mô hình discriminator và generator có chênh lệch nhưng không đáng kể. Ngoài ra độ chính xác của bài toán cũng rất cao.

Nội dung

- 1 Giới thiệu đề tài
- 2 Data
- 3 Mô hình
- 4 Đánh giá
- 5 Kết luận
- 6 Hướng phát triển² trong tương lai

Hướng phát triển

Với đề tài này nhóm chúng em đã tiến hành tìm hiểu, triển khai mô hình và tìm những biện pháp cải tiến. Bên cạnh đó kết quả đem lại tương đối cao nhưng nhóm chúng em có định hướng sẽ tiếp tục tìm hiểu rộng và sâu hơn về đề tài từ đó có thể đưa ra các mô hình đào tạo khác có kết quả tốt hơn mô hình hiện tại.

Vì thế trong thời gian tới bài báo này sẽ được áp dụng các mô hình khác và đánh giá sự tiến bộ giữa các mô hình.

Thank you!

CẢM ƠN THẦY VÀ CÁC BẠN ĐÃ LẮNG NGHE BÀI THUYẾT TRÌNH CỦA
NHÓM 2