

Chuyển đổi phong cách hình ảnh

Sinh viên Phạm Thành Trung & Lê Văn Trung & Nguyễn Tuấn Trung & Lưu Hoàng Ngọc Trinh
Giảng viên dạy PGS. TS Huỳnh Trung Hiếu & Lưu Giang Nam
Hội đồng phản biện PGS. TS Huỳnh Trung Hiếu & TS. Nguyễn Chí Kiên & TS. Bùi Thanh Hùng & Lưu Giang Nam

Tóm tắt nội dung

Bài toán dịch Image to image (I2I) là một trong những bài toán có khó trong chủ đề Thị giác máy tính. Với bài toán này, nhóm sẽ dựa trên một bài báo tên là StyleFlow For Content-Fixed Image to Image Translation của Fan et al công bố năm 2022, để hiện thực lại và so sánh (benchmark). Sử dụng mô hình StyleFlow theo kiến trúc normalizing flows và mô đun Style-Aware Normalization vào hai vấn đề Normally và Strongly constrained translation. Cụ thể, với vấn đề normally constrained là bài toán style transfer là chuyển đổi phong cách nghệ thuật của hình ảnh, còn bài strongly constrained là bài toán photo-realism là hiện thực hóa (làm cho nó giống thật nhất) ảnh tạo từ trò chơi. Các bộ dữ liệu sẽ dùng MS COCO, WikiArt, GTA5, Cityscapes.

Bố cục bài báo cáo

Nhóm sẽ dựa trên một paper là StyleFlow For Content-Fixed Image to Image Translation của Fan et al để hiện thực lại kết quả. Gồm các phần chính sau:

1. Giới thiệu bài toán: Mô tả bài toán dịch Image to Image (I2I), các ứng dụng của bài toán I2I, những model cho tới hiện tại đã đạt được gì, chưa đạt được gì, đề xuất cải tiến.
2. Phương pháp thực hiện: Sử dụng kiến trúc StyleFlow dựa trên model GLOW và đề xuất một module mới là SAN, mô tả về kiến trúc mô hình mới, cách hoạt động của SAN, công thức loss như content loss aligned style loss, smooth loss.
3. Thực nghiệm Data sử dụng: MS COCO, WikiArt, GTA5, Cityscapes Với bài toán style transfer là Coco to WikiArt, bài photo-realism sử dụng GTA5 to Cityscapes Vì bài báo gốc không để link dataset mà chỉ nêu tên nên nhóm em tìm kiếm và tạo bộ dataset khác dựa theo tên.
4. Đánh giá mô hình: Đánh giá định lượng sử dụng Structural Similarity Index Metrics (SSIM) và style loss so sánh với các kết quả của mô hình đã thực hiện trước đây. Đánh giá thực hiện độc lập trên 2 bài toán style transfers và photo-realism.
5. Kết luận và hướng phát triển.

Tài liệu

- [1] W. Fan, J. Chen, J. Ma, J. Hou, S. Yi (2022) StyleFlow For Content-Fixed Image to Image Translation *arXiv:2207.01909*, doi: 10.48550/arXiv.2207.01909.
- [2] D. P. Kingma, P. Dhariwal (2018) Glow: Generative Flow with Invertible 1x1 Convolutions *arXiv:1807.03039*, doi: 10.48550/arXiv.1807.03039.