

ỨNG DỤNG SEGMENT ANYTHING MODEL CHO BÀI TOÁN ĐIỀN MÀU ẢNH XÁM 3 GIAI ĐOẠN CÓ NHẬN THỨC VỀ VÙNG

Lý Văn Nhật Tiến - 21521525

Tóm tắt

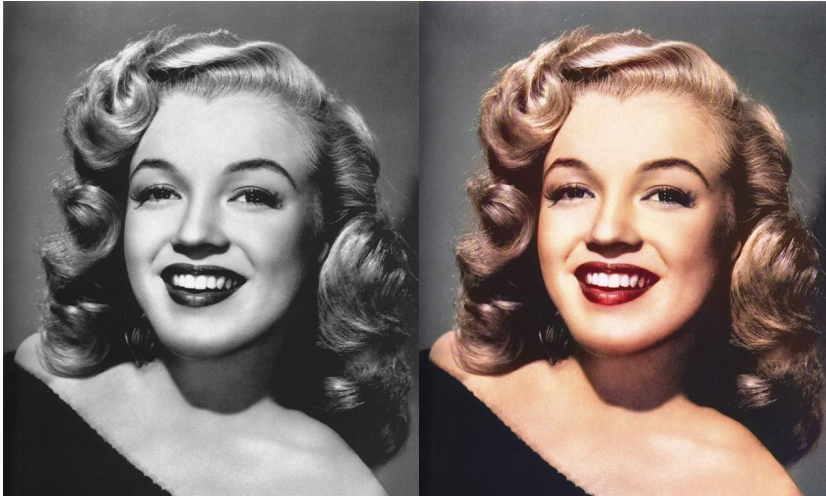
- Lớp: CS519.P11
- Link Github của nhóm: <https://github.com/ntien-281/cs519.p11>
- Link YouTube video: <https://youtu.be/x4w1PmM9kwY>



Lý Văn Nhật Tiến
MSSV: 21521525

Giới thiệu

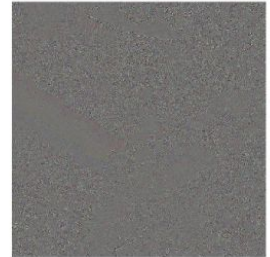
- Điền màu ảnh xám (*grayscale image colorization*) là tác vụ “tô màu” cho các bức ảnh này.
- Bài toán trên mang tính ứng dụng cực kỳ lớn.



Before compression



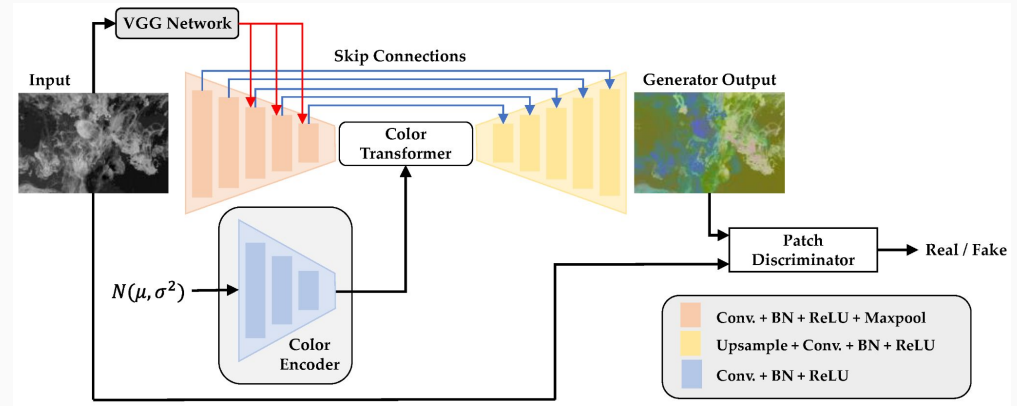
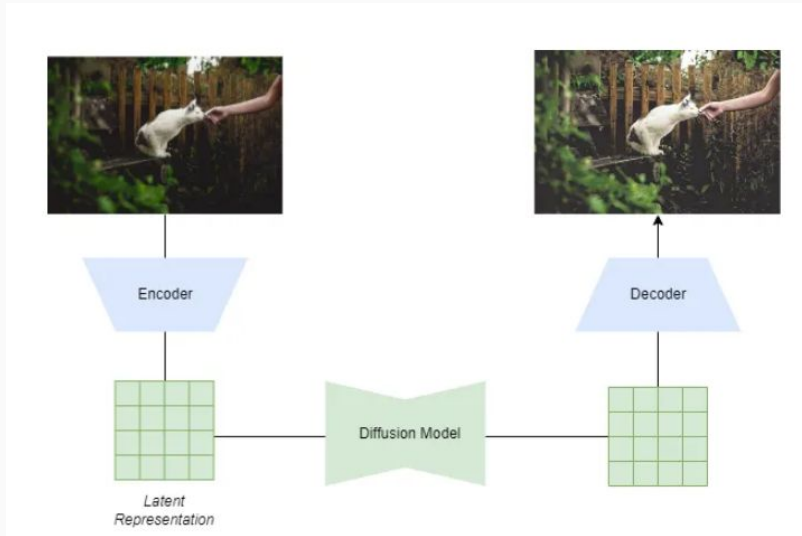
After JPEG decomposition
compression ratio: 2.82



RGB compression error
mean RMSE: 7.03
mean SNR: 24.6 dB

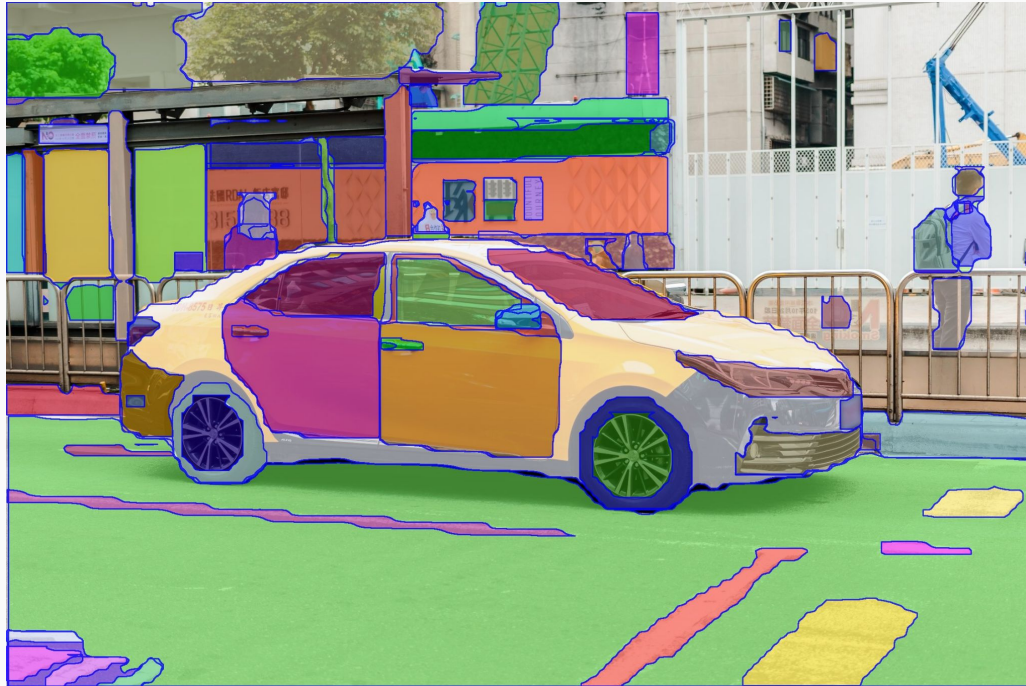
Giới thiệu

- Các phương pháp hiện có



Giới thiệu

- Đề xuất lời giải mới cho bài toán này. Sử dụng *Segment Anything Model*.



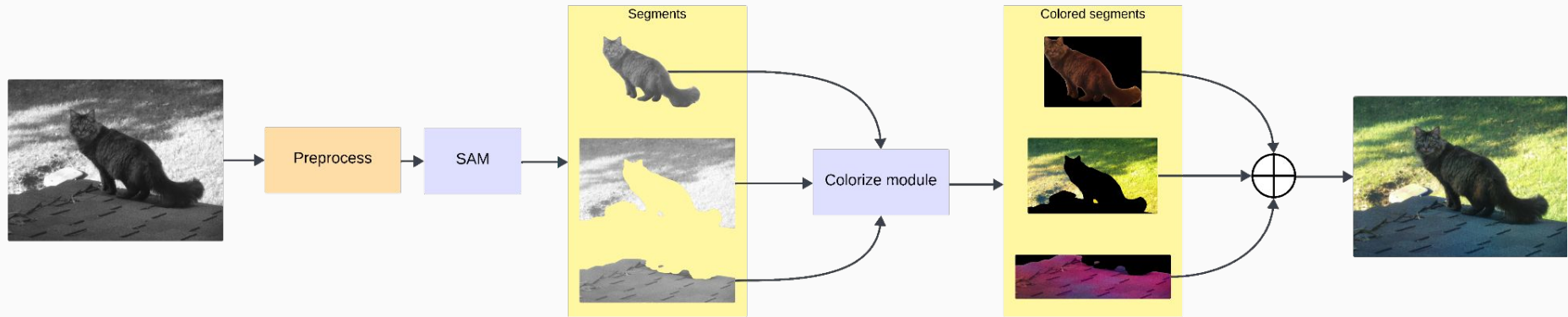
Mục tiêu

- Xây dựng hệ thống điền màu cho ảnh gồm 3 thành phần: module phân chia vùng, module điền màu và module hợp nhất.
- Đánh giá được hệ thống đã xây dựng và so sánh với các nghiên cứu khác dựa trên các benchmark, bộ dữ liệu và metrics khác nhau.
- Đánh giá được sự ảnh hưởng của cấu hình các module lên hiệu năng và độ chính xác của hệ thống (*Ablation study*). Đồng thời phát hiện các yếu điểm và đề xuất được hướng phát triển trong tương lai.

Nội dung và Phương pháp

- Tìm hiểu, lựa chọn mô hình và bộ dữ liệu cho module điền màu.
- Xây dựng module hợp nhất các segment đã được điền màu.
- Finetune mô hình *SAM* trên tập dữ liệu đã được chọn.
- Xây dựng, huấn luyện và đánh giá hệ thống.
- Đánh giá sự phụ thuộc của hệ thống vào các module và đề xuất phát triển.

Nội dung và Phương pháp



Kết quả dự kiến

- Xây dựng, huấn luyện và đánh giá được một hệ thống ổn định.
- Hệ thống đạt kết quả **ít nhất** là ngang ngửa với các phương pháp hiện có trên độ đo FDI và SSIM, điều này thể hiện tiềm năng phát triển của các nghiên cứu theo sau. Bên cạnh đó, thời gian suy luận phải thấp, hướng đến ứng dụng trong thực tiễn.
- Đánh giá được sự phụ thuộc vào cấu hình module, đồng thời đóng gói mã nguồn phục vụ mục đích tham khảo trong nghiên cứu ứng dụng.

Tài liệu tham khảo

- [1]A. Kirillov et al., “Segment Anything,” in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, IEEE, Oct. 2023. Accessed: Dec. 29, 2024. [Online]. Available: <https://doi.org/10.1109/iccv51070.2023.00371>
- [2]H. Shafiq, T. Nguyen, and B. Lee, “Colorformer: A Novel Colorization Method Based on a Transformer,” Elsevier BV, 2024. Accessed: Jan. 01, 2025. [Online]. Available: <https://doi.org/10.2139/ssrn.4937072>
- [3]C. Saharia et al., “Palette: Image-to-Image Diffusion Models,” in *Special Interest Group on Computer Graphics and Interactive Techniques Conference Proceedings*, New York, NY, USA: ACM, Aug. 2022, pp. 1–10. Accessed: Jan. 01, 2025. [Online]. Available: <https://doi.org/10.1145/3528233.3530757>
- [4]J.-W. Su, H.-K. Chu, and J.-B. Huang, “Instance-Aware Image Colorization,” in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2020, pp. 7965–7974. Accessed: Jan. 01, 2025. [Online]. Available: <https://doi.org/10.1109/cvpr42600.2020.00799>
- [5]R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, “High-Resolution Image Synthesis with Latent Diffusion Models,” in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jun. 2022, pp. 10674–10685. Accessed: Jan. 12, 2025. [Online]. Available: <https://doi.org/10.1109/cvpr52688.2022.01042>