

# ỨNG DỤNG MÔ HÌNH DIFFUSION VÀ PIXEL-WISE SEMANTIC SEGMENTATION TRONG BÀI TOÁN TÔ MÀU ẢNH

Nguyễn Hồng Thái - 21520445

# Tóm tắt

- Lớp: CS519.011
- Link YouTube video:  
[https://www.youtube.com/watch?v=9xT\\_TyvFYP8](https://www.youtube.com/watch?v=9xT_TyvFYP8)
- Ảnh + Họ và tên thành viên:



Nguyễn Hồng Thái – 21520445

# Giới thiệu

- Image colorization
- Diffusion model
- Pixel-wise semantic segmentation

# Mục tiêu

- Hiện thực hóa việc kết hợp mô hình Diffusion và pixel-wise semantic segmentation để giải quyết bài toán tô màu ảnh.
- Nghiên cứu tác động của việc kết hợp mô hình Diffusion và pixel-wise semantic segmentation đối với chất lượng và tốc độ tô màu ảnh.
- Xây dựng chương trình minh họa và so sánh đánh giá với các phương pháp khác đã giải quyết được bài toán tô màu ảnh.

# Nội dung

- Các lý thuyết liên quan đến các dạng cấu trúc mô hình Diffusion, pixel-wise semantic segmentation, và bài toán tô màu ảnh sẽ được nghiên cứu để có thể hiểu rõ về cách thức hoạt động của các mô hình và phương pháp tô màu ảnh hiện có.
- Nghiên cứu quá trình tích hợp thông tin pixel-wise semantic segmentation vào mô hình Diffusion trước khi huấn luyện mô hình.
- Mô hình Diffusion sẽ được huấn luyện trên tập ImageNet.
- Xây dựng chương trình minh họa sử dụng mô hình Diffusion đã được huấn luyện để tô màu ảnh.

# Phương pháp

- Thay đổi và tinh chỉnh kiến trúc của mô hình Diffusion để có thể tích hợp thông tin pixel-wise semantic segmentation vào mô hình trước khi huấn luyện.
- Phương pháp đánh giá: Đánh giá chất lượng hình ảnh của mô hình bằng PSNR (Peak Signal-to-Noise Ratio) và SSIM (Structural Similarity Index Measure).

# Kết quả dự kiến

- Mô hình kết hợp sẽ cải thiện chất lượng tô màu ảnh lên 5 ~ 10% so với các mô hình không kết hợp và phương pháp hiện có.**[6]**
- Mô hình kết hợp sẽ giải quyết được các vấn đề lem màu cho ảnh gốc có độ phân giải thấp, hay khó tô màu chính xác cho đối tượng phức tạp.

# Tài liệu tham khảo

- [1] Naoki Matsunaga, Masato Ishii, Akio Hayakawa, Kenji Suzuki, Takuya Narihira: Fine-grained Image Editing by Pixel-wise Guidance Using Diffusion Models. CoRR abs/2212.02024 (2022)
- [2] Yuri Viazovetskyi, Vladimir Ivashkin, Evgeny Kashin: StyleGAN2 Distillation for Feed-forward Image Manipulation. CoRR abs/2003.03581 (2020)
- [3] Richard Zhang, Phillip Isola, Alexei A. Efros: Colorful Image Colorization. CoRR abs/1603.08511 (2016)
- [4] Emiel Hoogeboom, Alexey A. Gritsenko, Jasmijn Bastings, Ben Poole, Rianne van den Berg, Tim Salimans: Autoregressive Diffusion Models. CoRR abs/2110.02037 (2021)
- [5] Aäron van den Oord, Nal Kalchbrenner, Oriol Vinyals, Lasse Espeholt, Alex Graves, Koray Kavukcuoglu: Conditional Image Generation with PixelCNN Decoders. CoRR abs/1606.05328 (2016)
- [6] Saeed Anwar, Muhammad Tahir, Chongyi Li, Ajmal Mian, Fahad Shahbaz Khan, Abdul Wahab Muzaffar: Image Colorization: A Survey and Dataset. CoRR abs/2008.10774 (2020)