



## Sinh ảnh theo phong cách nghệ thuật

Đồ án nhập môn AI

Thành viên :Nguyễn Văn Mạnh

**Giáo viên hướng dẫn: Phan Mạnh Thường**

Trường Đại học Lạc Hồng, số 10 Huỳnh Văn Nghệ, Bửu Long, Biên Hòa, Đồng Nai, Việt Nam

### ARTICLE INFO

Đề tài này được lựa chọn vì Neural Style Transfer (NST) là kỹ thuật biến đổi phong cách hình ảnh, cho phép kết hợp nội dung của một ảnh với phong cách nghệ thuật của ảnh khác. NST có ứng dụng trong nhiều lĩnh vực như thiết kế đồ họa, nghệ thuật số, xử lý ảnh sáng tạo và ứng dụng mạng xã hội.

### KEYWORDS

Neural Style Transfer

Content Image

Style Image

Feature Extraction

VGG19

Gram Matrix

Image Postprocessing

Deep Learning

PyTorch

Phân tích hình ảnh sử dụng các mô hình CNN, đặc biệt là VGG19, để trích xuất các đặc trưng nội dung và phong cách. VGG19 vượt trội trong việc trích xuất các đặc trưng không gian cục bộ, trong khi việc tính Gram matrix giúp nắm bắt các mối quan hệ textural và stylistic trong ảnh. Các phương pháp tối ưu hóa như LBFGS được sử dụng để giảm thiểu content loss và style loss, từ đó tạo ra ảnh kết hợp hài hòa giữa nội dung và phong cách.

Mục tiêu cuối cùng là tạo ra các bức ảnh được “tô điểm” theo phong cách nghệ thuật mong muốn, đồng thời giữ nguyên đặc trưng nội dung, nhằm nâng cao trải nghiệm người dùng trong các ứng dụng sáng tạo và nghệ thuật số.

## 1. Giới thiệu

Neural Style Transfer (NST) là một lĩnh vực nghiên cứu đang phát triển trong học sâu, với khả năng biến đổi phong cách hình ảnh mà vẫn giữ nguyên nội dung gốc. Kỹ thuật này có tiềm năng ứng dụng rộng rãi trong nhiều lĩnh vực sáng tạo, thiết kế đồ họa, nghệ thuật số và mạng xã hội. Dưới đây là một số lợi ích nổi bật khi áp dụng NST.

### 1.1. Tạo ra các tác phẩm nghệ thuật cá nhân hóa

Ảnh gốc có thể được biến đổi theo phong cách của các họa sĩ nổi tiếng hoặc các phong cách nghệ thuật khác nhau.

Ví dụ: Một bức ảnh bình thường có thể được chuyển thành phong cách tranh Van Gogh, Monet hay trừu tượng.

Giúp người dùng tạo ra các sản phẩm sáng tạo độc đáo, tăng trải nghiệm trực quan.

### 1.2. Ứng dụng trong thiết kế đồ họa và marketing

Cho phép thiết kế nhanh các poster, banner, ảnh quảng cáo theo nhiều phong cách khác nhau.

Ví dụ: Thử nghiệm nhiều phong cách hình ảnh mà không cần thiết kế lại từ đầu.

Tăng tính linh hoạt và hiệu quả trong các chiến dịch truyền thông.

### 1.3. Tích hợp trong các ứng dụng di động và mạng xã hội

Ứng dụng chỉnh sửa ảnh trực tiếp cho người dùng cuối, tạo ảnh nghệ thuật từ ảnh selfie hay ảnh phong cảnh.

Ví dụ: Khi người dùng tải ảnh lên, ứng dụng có thể gợi ý nhiều phong cách khác nhau để lựa chọn.

Tăng sự tương tác và trải nghiệm người dùng.

### 1.4. Hỗ trợ trong giáo dục (EdTech)

Giúp học sinh, sinh viên hiểu được sự khác biệt giữa các phong cách nghệ thuật.

Ví dụ: Học sinh có thể thực hành biến đổi một bức ảnh gốc sang các phong cách khác nhau để học về nghệ thuật.

Tạo môi trường học tập trực quan và sinh động.

## 1.5. Hỗ trợ nghiên cứu AI và xử lý ảnh

NST là cơ sở để nghiên cứu các mô hình trích xuất đặc trưng hình ảnh, tối ưu hóa loss function (content loss, style loss).

Phát triển các kỹ thuật học sâu, như CNN, VGG19, để nâng cao khả năng trích xuất nội dung và phong cách.

## 1.6. Ứng dụng trong nghệ thuật và giải trí

Tạo video, animation, hoặc game với phong cách hình ảnh nghệ thuật độc đáo.

Ví dụ: Biến thể giới 3D trong game thành các phong cách hội họa khác nhau, tăng trải nghiệm người chơi.

## 1.7. Thúc đẩy nghiên cứu và phát triển các kỹ thuật học sâu

Việc áp dụng NST giúp nâng cao khả năng nghiên cứu các kiến trúc mạng CNN, tính toán Gram matrix và tối ưu hóa loss function.

Là cơ sở để phát triển các ứng dụng sáng tạo, nghệ thuật số và xử lý hình ảnh nâng cao.

## 2. Thách thức về Neural Style Transfer

### 2.1. Hạn chế về độ phân giải và chất lượng ảnh

Các ảnh content và style thường có độ phân giải khác nhau hoặc chất lượng thấp, khiến nhiều chi tiết về texture hoặc màu sắc không được tái hiện chính xác. Điều này làm giảm hiệu quả của các mô hình NST, vốn phụ thuộc nhiều vào đặc trưng hình ảnh để chuyển phong cách. Một số nghiên cứu đã áp dụng kỹ thuật tăng độ phân giải hoặc tiền xử lý ảnh, nhưng kết quả vẫn chưa đạt tối ưu.

### 2.2. Sự tương đồng giữa các biểu cảm cảm xúc

Một số phong cách có đặc trưng quá nổi bật hoặc phức tạp, dẫn đến việc nội dung ảnh gốc bị mất đi khi áp dụng NST. Dù có các giải pháp như cân bằng trọng số content/style hoặc multi-style blending, vẫn khó đảm bảo

*Sinh ảnh theo phong cách nghệ thuật*  
rằng cả nội dung và phong cách đều được thể hiện hài hòa.

### 2.3 Vấn đề tính toán và hiệu suất

NST thường yêu cầu tối ưu hóa lặp nhiều bước (iterative optimization) trên ảnh, dẫn đến tốn nhiều thời gian và tài nguyên GPU. Khi xử lý ảnh lớn hoặc nhiều ảnh đồng thời, hiệu suất tổng thể bị hạn chế, đặc biệt trên các thiết bị có giới hạn phần cứng.

### 2.4. Mất cân bằng trong dữ liệu phong cách

Khi sử dụng nhiều style image, nếu một số phong cách chiếm ưu thế hoặc không được chuẩn hóa, kết quả tạo ảnh có thể thiên lệch về phong cách nào đó. Các kỹ thuật blending hay weighting cần được điều chỉnh cẩn thận để tạo kết quả hài hòa.

### 2.5. Khó khăn trong giữ chi tiết nội dung

Một thách thức lớn của NST là giữ nguyên chi tiết nội dung trong khi chuyển phong cách. Các texture hoặc pattern phức tạp có thể làm mờ các chi tiết quan trọng, khiến ảnh kết quả không trung thực với ảnh gốc.

## 3. Phương pháp nghiên cứu Sinh ảnh nghệ thuật

### 3.1 Bộ Dữ Liệu

#### 3.1.1 Mô tả bộ dữ liệu công khai

Sử dụng ảnh Content (nội dung) và ảnh Style (phong cách) từ các nguồn công khai hoặc do người dùng upload.

Ảnh Content chứa nội dung chính (ví dụ: phong cảnh, chân dung), ảnh Style chứa phong cách nghệ thuật (tranh sơn dầu, trừu tượng...).

#### 3.1.2 Chuẩn hóa ảnh đầu vào

Thay đổi kích thước ảnh về max\_size (ví dụ 512x512) để đồng nhất.

Chuyển ảnh về dạng RGB và chuẩn hóa theo mean/std giống ImageNet.

#### 3.1.3. Mô tả bộ dữ liệu công khai

Sử dụng mô hình VGG19 pretrained để trích xuất content features và style features.

Tính Gram matrix cho các lớp style để nắm bắt texture và pattern.

#### 3.1.4. Chia tập dữ liệu

Khi thử nghiệm với nhiều ảnh content và style, có thể chia tập dữ liệu thành **tập kiểm tra / tập huấn luyện nội bộ** để đánh giá kết quả NST.

#### 3.1.5. Đối tượng và mẫu nghiên cứu

Các ảnh đại diện cho nhiều phong cách khác nhau để đảm bảo độ đa dạng.

Sử dụng từ 1–5 style images cho mỗi ảnh content để thử nghiệm multi-style blending.

### 3.2. Phương pháp nghiên cứu

#### 3.2.1. Mô hình và tối ưu hóa

Mô hình NST dựa trên CNN (VGG19), tối ưu ảnh đầu ra target image sao cho content loss và style loss đạt tối ưu.

Content Loss:

$$\mathcal{L}_{\text{content}} = \frac{1}{2} \sum_{i,j} \left( F_{ij}^{\text{target}} - F_{ij}^{\text{content}} \right)^2$$

Style Loss:

$$\mathcal{L}_{\text{style}} = \sum_l w_l \frac{1}{4N_l^2 M_l^2} \sum_{i,j} \left( G_{ij}^{l,\text{target}} - G_{ij}^{l,\text{style}} \right)^2$$

#### 3.2.2. Phương pháp tối ưu hóa

Sử dụng LBFGS và Adam để cập nhật target image.

Lặp nhiều bước (step) để giảm tổng loss:

$$\mathcal{L}_{\text{total}} = \alpha \mathcal{L}_{\text{content}} + \beta \mathcal{L}_{\text{style}}$$

#### 3.2.3. Kết quả đầu ra

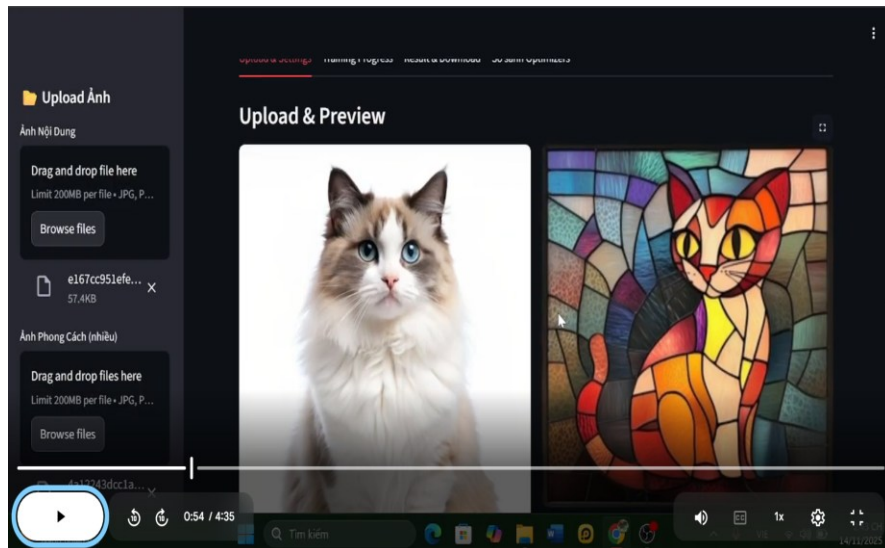
Ảnh output giữ nội dung của ảnh gốc, nhưng mang phong cách nghệ thuật mong muốn.

Có thể áp dụng postprocessing: sharpen, saturation,

Sinh ảnh theo phong cách nghệ thuật  
contrast để cải thiện trực quan.

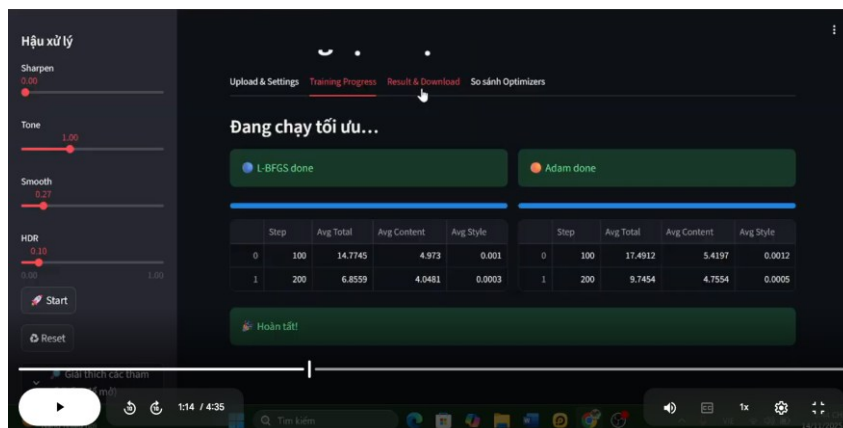
## 4.Demo ảnh và kết quả

### 4.1 đầu vào



Hình 1 . Ảnh đầu vào bao gồm ảnh cần phong cách và ảnh phong cách

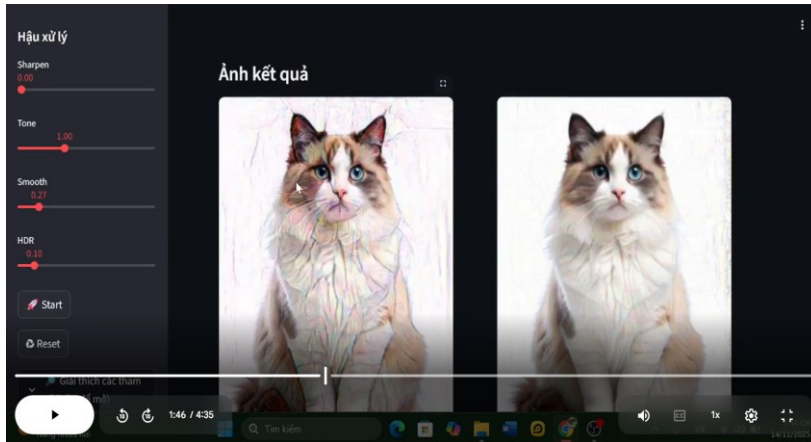
### 4.2 .Quá trình học khi người dùng bắt đầu



Hình 2 . Ảnh Quá trình chạy 2 model L-BSGS và  
dam

### 4.3 . kết quả

Hình 3 . Ảnh kết quả chạy 2 model L-BSGS và



đam

### 4.4 Bảng thống kê

Hậu xử lý

Sharpen

0.0

Tone

1.00

Smooth

0.01

HDR

0.10

Start

Reset

Giải thích các tham số (bấm để mở)

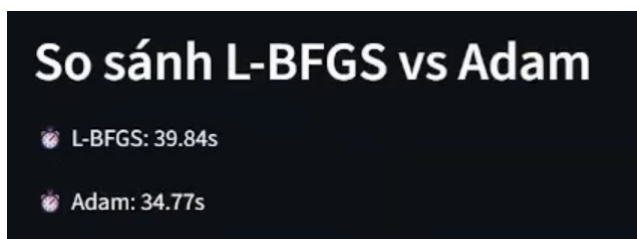
Bảng tóm tắt mỗi 100 step (hoặc mốc cuối)

L-BFGS

Adam

Step	Avg Total	Avg Content	Avg Style	Step	Avg Total	Avg Content	Avg Style		
0	100	14.7745	4.973	0.001	0	100	17.4912	5.4197	0.0012
1	200	6.8559	4.5481	0.0003	1	200	9.7454	4.7554	0.0005
2	300	24.3967	6.6278	0.0009	2	300	29.5371	6.9485	0.0011
3	200	9.3028	5.5026	0.0002	3	200	14.6302	6.4507	0.0004
4	300	8.0546	5.0023	0.0002	4	300	12.3884	6.0223	0.0003
5	400	7.522	4.8384	0.0001	5	400	11.145	5.7747	0.0003
6	500	7.2279	4.7142	0.0001	6	500	10.4146	5.6003	0.0002

Hình 4 . Ảnh kết quả sau 100 step của 2 thuật toán



So sánh L-BFGS vs Adam	
L-BFGS:	39.84s
Adam:	34.77s

Hình 5 . Ảnh Thời gian chạy của 2 loại thuật toán sau 500 step

## Mục lục

ARTICLE INFO .....	1
1.GIỚI THIỆU .....	2
1.1.TẠO RA CÁC TÁC PHẨM NGHỆ THUẬT CÁ NHÂN HÓA.....	2
1.2.ỨNG DỤNG TRONG THIẾT KẾ ĐỒ HỌA VÀ MARKETING .....	2
1.3.TÍCH HỢP TRONG CÁC ỨNG DỤNG ĐỘNG VÀ MẠNG XÃ HỘI.....	2
1.4.HỖ TRỢ TRONG GIÁO DỤC (EdTECH).....	2
1.5.HỖ TRỢ NGHIÊN CỨU AI VÀ XỬ LÝ ẢNH.....	2
1.6.ỨNG DỤNG TRONG NGHỆ THUẬT VÀ GIẢI TRÍ .....	2
1.7.THÚC ĐẨY NGHIÊN CỨU VÀ PHÁT TRIỂN CÁC KỸ THUẬT HỌC SÂU.....	2
2.THÁCH THỨC VỀ NEURAL STYLE TRANSFER.....	2
2.1.HẠN CHẾ VỀ ĐỘ PHÂN GIẢI VÀ CHẤT LƯỢNG ẢNH .....	2
2.2.SỰ TƯƠNG ĐỒNG GIỮA CÁC BIỂU CẢM CẢM XÚC.....	2
2.3VẤN ĐỀ TÍNH TOÁN VÀ HIỆU SUẤT .....	3
2.4.MẤT CÂN BẰNG TRONG DỮ LIỆU PHONG CÁCH.....	3
2.5.KHÓ KHĂN TRONG GIỮ CHI TIẾT NỘI DUNG.....	3
3.PHƯƠNG PHÁP NGHIÊN CỨU SINH ẢNH NGHỆ THUẬT.....	3
3.1BỘ DỮ LIỆU.....	3
3.1.1Mô tả bộ dữ liệu công khai.....	3
3.1.2 Chuẩn hóa ảnh đầu vào .....	3
3.1.3. Mô tả bộ dữ liệu công khai .....	3
3.1.4. Chia tập dữ liệu .....	3
3.1.5. Đối tượng và mẫu nghiên cứu .....	3
3.2. PHƯƠNG PHÁP NGHIÊN CỨU .....	3
3.2.1. Mô hình và tối ưu hóa .....	3
3.2.2. Phương pháp tối ưu hóa .....	3
3.2.3. Kết quả đầu ra .....	3
4.DEMO ẢNH VÀ KẾT QUẢ .....	4
4.1 ĐẦU VÀO .....	4
4.2 .QUÁ TRÌNH HỌC KHI NGƯỜI DÙNG BẮT ĐẦU .....	4
4.3 . KẾT QUẢ.....	5
4.4 BẢNG THỐNG KÊ .....	5

