

**BỘ CÔNG THƯƠNG**  
**TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP HÀ NỘI**



**ĐỒ ÁN TỐT NGHIỆP**

**ĐỀ TÀI: ỨNG DỤNG HỌC SÂU TRONG**  
**BÀI TOÁN STYLE TRANSFER**

Sinh viên thực hiện  
**Trần Danh Quang**

Lớp – khóa  
**2019DHKHM01 – K14**

Giảng viên hướng dẫn: TS. Trần Hùng Cường

**Hà Nội – Tháng 11 / Năm 2024**

## LỜI CẢM ƠN

Để hoàn thành đề tài nghiên cứu khoa học này, nhóm thực tập vô cùng biết ơn thầy cô khoa Công nghệ thông tin trường Đại học Công nghiệp Hà Nội, trong suốt thời gian qua đã truyền đạt những kiến thức và bài học quý báu. Những tâm kiến thức, tâm huyết truyền đạt của thầy cô là hành trang, cơ sở khoa học để cho em có thể theo đuổi ước mơ, đam mê và hoàn thành đồ án tốt nghiệp này. Quá trình thực hiện đề tài cũng giúp cho em củng cố thêm về kiến thức, cũng như có thêm động lực, sự tự tin trên con đường phát triển sau này. Em xin xin gửi lời cảm ơn chân thành đến quý thầy cô, mong thầy cô luôn thành công trên sự nghiệp trồng người cao quý.

Đặc biệt, em xin gửi lời cảm ơn chân thành đến thầy, tiến sĩ Trần Hùng Cường, là người hướng dẫn đề tài của em. Trong thời gian thực hiện đề tài, thầy luôn quan tâm, hướng dẫn cho nhóm biết và thực hiện đề tài một cách khoa học nhất. Chúng em chúc thầy luôn có nhiều sức khỏe và niềm vui trong sự nghiệp giáo dục.

Nhóm thực hiện cũng xin cảm ơn những anh chị, bạn bè trong nhóm nghiên cứu đã giúp đỡ, cùng đồng hành với nhóm trong chặng đường vừa qua. Chúc cho các bạn luôn luôn thành công trong sự nghiệp học tập và có nhiều ước mơ, đam mê với sự nghiệp học tập cũng như khởi nghiệp sau này.

Hà Nội, tháng 11, năm 2024

# MỤC LỤC

LỜI CẢM ƠN.....	2
MỤC LỤC .....	3
DANH MỤC HÌNH VẼ .....	4
PHẦN MỞ ĐẦU .....	1
1. Lý do chọn đề tài .....	1
2. Mục tiêu đề tài .....	1
3. Ý nghĩa đề tài.....	2
4. Phương pháp nghiên cứu.....	2
5. Đối tượng và phạm vi nghiên cứu .....	2
6. Bố cục đề tài .....	3
PHẦN NỘI DUNG.....	4
Chương 1. Giới thiệu bài toán.....	4
1. Giới thiệu đề tài .....	4
2. Lịch sử phát triển.....	4
Chương 2. Cơ sở lý thuyết .....	6
1. Không sử dụng mạng nơ-ron.....	6
1.1. Stroke-Based Rendering .....	6
1.2. Region-Based Techniques.....	7
1.3. Example-Based Rendering.....	8
1.4. Tổng kết .....	8
2. Sử dụng mạng nơ-ron .....	9
2.1. Lấy nội dung của ảnh .....	10
2.2. Lấy phong cách của ảnh.....	12
2.3. Tái tạo hình ảnh.....	14

## DANH MỤC HÌNH VẼ

Hình 1: Mô tả thuật toán Stroke-Based Rendering	6
Hình 2: Mô tả thuật toán Region-Based Techniques	7
Hình 3: Mô tả kết quả của kỹ thuật example-based rendering	8
Hình 4: Vẽ lại ảnh theo phong cách	9
Hình 5: Ma trận của bộ lọc sobel	10
Hình 6: Phép toán convolution trên ảnh	11
Hình 7: Mô tả công thức Gram-matrix	13
Hình 8: Quá trình tổng hợp điểm ảnh bằng cách tìm lân cận	14
Hình 9: Kiến trúc một mô hình thuật toán Style transfer	14
Hình 10: Kiến trúc mạng style transfer	17
Hình 11: Kiến trúc mạng VGG19	18
Hình 12: Trích xuất đặc trưng content	33
Hình 13: Trích xuất đặc trưng style	34
Hình 14: Mô phỏng quá trình tái tạo ảnh	35
Hình 15: Mô tả thuật toán Multi-style Transfer	36
Hình 16: Style transfer với 1 content và 2 layer	38
Hình 17: Ảnh style được xoay theo các góc trước khi training	39
Hình 18: ảnh style và content thực nghiệm	44
Hình 19: kết quả thực nghiệm	44

## PHẦN MỞ ĐẦU

### 1. Lý do chọn đề tài

Trong thời đại công nghệ phát triển như hiện nay, việc ứng dụng những công nghệ tiên tiến của khoa học vào đời sống đã giúp chúng ta tăng hiệu suất công việc, đồng thời tạo ra được những kết quả vô cùng kinh ngạc. Trong đó, lĩnh vực giải trí, với sự giúp đỡ của những ứng dụng xử lý hình ảnh và thị giác máy tính đã mang đến những tiến bộ vượt bậc. Một số đóng góp của thị giác máy tính có thể thấy như: những ứng dụng làm nét hình ảnh, ứng dụng tạo ra hình ảnh mới với các bộ lọc khác nhau,... Trong đó, ứng dụng vẽ lại ảnh theo những phong cách khác nhau cũng là một chủ đề đang được quan tâm. Một trong những kỹ thuật nổi bật nhất của ứng dụng vẽ lại ảnh theo phong cách là thuật toán Style transfer.

Style transfer lần đầu tiên được giới thiệu bởi Gatys và đồng nghiệp, với nhiệm vụ có thể tách nội dung từ một bức ảnh và kết hợp với phong cách của một bức ảnh khác<sup>1</sup>. Kỹ thuật này đã được sử dụng để tạo ra những tác phẩm nghệ thuật có tính thẩm mỹ cao. Nó không chỉ được sử dụng bởi những nhà phát triển nội dung hay các nhà thiết kế, mà còn được ứng dụng rất nhiều trong lĩnh vực giải trí và quảng cáo,...

Vậy nguyên lý hoạt động của thuật toán Style transfer ra sao, nó có những ưu nhược điểm gì, những ứng dụng của nó vào trong cuộc sống như thế nào,...? Những câu hỏi đó là động lực thôi thúc nhóm thực hiện lựa chọn đề tài “*Nghiên cứu thuật toán Style Transfer trong ứng dụng vẽ lại ảnh theo phong cách*”

### 2. Mục tiêu đề tài

Trong quá trình thực hiện đề tài, nhóm thực hiện đã đề ra và hoàn thành các mục tiêu của đề tài như sau:

- Nghiên cứu thuật toán Style transfer trong việc vẽ lại ảnh theo phong cách
- Đánh giá ưu điểm, nhược điểm của thuật toán
- Tìm hiểu một số phương pháp vẽ lại ảnh theo phong cách
- Tìm phương án cải thiện

### 3. Ý nghĩa đề tài

#### *Ý nghĩa khoa học:*

- Nghiên cứu kỹ thuật Style transfer và các kỹ thuật khác trong việc vẽ lại ảnh theo phong cách
- Cung cấp những kiến thức nhất định để ứng dụng, cải thiện các kỹ thuật đó

#### *Ý nghĩa kinh tế:*

- Có thể ứng dụng kỹ thuật vẽ lại ảnh theo phong cách trong các ứng dụng chỉnh sửa hình ảnh của lĩnh vực giải trí
- Tạo ra những hình ảnh đẹp, độc đáo cho mỹ thuật cũng như quảng cáo,...

### 4. Phương pháp nghiên cứu

Trong quá trình thực hiện đề tài, nhóm thực hiện đã sử dụng những phương pháp nghiên cứu sau:

- Nghiên cứu tài liệu: thu thập các tài liệu, các xuất bản liên quan đến kỹ thuật vẽ lại ảnh theo phong cách cũng như thuật toán Style Transfer
- Phân tích và so sánh: nhóm thực hiện đã so sánh các phương pháp khác nhau trong việc vẽ lại ảnh theo phong cách, từ đó tìm ra những ưu và nhược điểm của những thuật toán này
- Phát triển mô hình: dựa trên những cơ sở kiến thức đã thu thập, nhóm thực hiện đề xuất mô hình và thực nghiệm
- Thực nghiệm: nhóm tiến hành thực nghiệm
- Đánh giá kết quả: nhóm đã thu thập và đánh giá kết quả, cũng như đề xuất phương án trong tương lai

### 5. Đối tượng và phạm vi nghiên cứu

Đối tượng của đề tài “*Nghiên cứu thuật toán Style Transfer trong ứng dụng vẽ lại ảnh theo phong cách*” bao gồm :

- Các thuật toán chuyển đổi phong cách
- Mạng tích chập
- Một số ứng dụng của việc vẽ lại ảnh theo phong cách

Phạm vi nghiên cứu của đề tài bao gồm:

- Phạm vi lý thuyết: nghiên cứu các thuật toán vẽ lại ảnh theo phong cách và phân tích các thuật toán này
- Phạm vi thực nghiệm: thực nghiệm các thuật toán, đưa ra mô hình, đánh giá kết quả
- Phạm vi ứng dụng: tìm hiểu một số ứng dụng của việc vẽ lại ảnh theo phong cách

## 6. Bố cục đề tài

Bài báo cáo gồm 4 chương. Các chương được trình bày như sau:

### **Chương 1:** Giới thiệu bài toán

Chương này nêu tổng quan về đề tài, lịch sử phát triển và các ứng dụng của thuật toán

### **Chương 2:** Cơ sở lý thuyết

Chương này, nhóm khảo sát trình bày một số kỹ thuật mà nhóm khảo sát được

### **Chương 3:** Phương pháp đề xuất

Chương này nói về thuật toán được sử dụng trong đề tài nghiên cứu của nhóm thực hiện, những đóng góp của nhóm

### **Chương 4:** Kết quả thực nghiệm

Trong chương này, nhóm thực hiện trình bày, đưa ra các kết quả của quá trình nghiên cứu và tìm hiểu.

# PHẦN NỘI DUNG

## Chương 1. Giới thiệu bài toán

### 1. Giới thiệu đề tài

Style transfer là một kỹ thuật trong lĩnh vực xử lý ảnh và thị giác máy tính, nhằm chuyển một bức ảnh ban đầu thành một ảnh mục tiêu với những phong cách khác nhau mà vẫn giữ nguyên nội dung của nó. Phong cách của ảnh mục tiêu có thể được tạo ra bởi các thuật toán, hoặc cũng có thể truyền một bức ảnh mang phong cách muốn đạt được vào trong mô hình.

Style tranfer hiện nay đã đóng góp rất nhiều trong các lĩnh vực nghệ thuật, giải trí, thiết kế,... Kỹ thuật này đã giúp những nhà phát triển nội dung, nhà thiết kế mở rộng khả năng sáng tạo bởi kết quả của kỹ thuật style transfer luôn rất đa dạng và có tính linh hoạt. Những bức ảnh số có thể được vẽ lại theo phong cách vẽ của những nghệ sĩ, họa sĩ nổi tiếng như Van Gogh, Picasso.

Một ví dụ điển hình của việc ứng dụng kỹ thuật này là trong các ngành quảng cáo, tiếp thị. Những hình ảnh được trộn từ những phong cách độc đáo đã thu hút hàng ngàn người quan tâm. Hay trong thời trang, những họa tiết bắt mắt được tạo ra phù hợp với xu hướng làm đẹp của mọi người,...

Ngoài ra, style transfer cũng là một đề tài khoa học được các trường học, trung tâm nghiên cứu quan tâm và liên tục xuất bản những công trình nghiên cứu về đề tài này. Nhìn chung, style transfer vẫn đang là một chủ đề được quan tâm và có nhiều ứng dụng trong đời sống. Nó mở ra nhiều cơ hội và tiềm năng trong các lĩnh vực của đời sống hiện tại.

### 2. Lịch sử phát triển

Lịch sử phát triển của Style transfer có thể được chia làm hai giai đoạn chính, trong đó lấy 2015 làm gốc.

Trước năm 2015: các thuật toán style transfer chủ yếu sử dụng các thuật toán thủ công, không có tính tự động cao. Các thuật toán này, có thể sử dụng các bộ lọc truyền



thống như Gaussian, Sobel, hoặc là phân vùng ảnh. Những phương pháp này cũng có thể tạo ảnh nghệ thuật, tuy nhiên nó không có tính linh hoạt cao.

Năm 2015: với sự ứng dụng thành công của mạng tích chập vào đề tài vẽ ảnh nghệ thuật, đã tạo ra một bước đột phá với bài toán style transfer. Việc tạo ra một hình ảnh nghệ thuật trở nên dễ dàng hơn.

Từ năm 2015 trở đi: các cải tiến nhằm vào kỹ thuật sử dụng mạng tích chập trong style transfer. Những mục tiêu cải tiến bao gồm: chất lượng hình ảnh mục tiêu, tốc độ mô hình, tính sáng tạo của ảnh mục tiêu,...

Đến những năm 2020, Style transfer càng được ứng dụng trong thực tế. Các lĩnh vực như: nghệ thuật, thời trang, quảng cáo,... đều áp dụng thành tựu của kỹ thuật này.

## Chương 2. Cơ sở lý thuyết

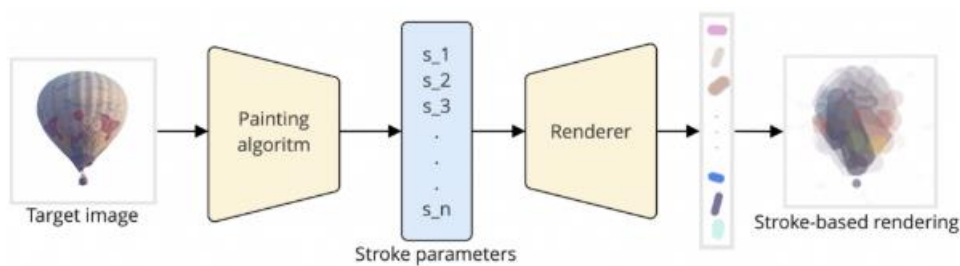
Vẽ ảnh nghệ thuật là một chủ đề đã xuất hiện từ rất lâu. Mỗi quốc gia, nền văn minh đều có những tác giả, tác phẩm với những phong cách vẽ khác nhau. Kể từ khi khoa học công nghệ phát triển, người ta đã tìm cách ứng dụng những sự phát triển đó vào trong nghệ thuật. Từ những năm 90 của thế kỷ trước cho đến hiện nay, đã có những kỹ thuật, thuật toán về đề tài nghệ thuật này, chúng có thể là những thuật toán đồ họa, thuật toán học máy hoặc học sâu,... Trong chương này, nhóm thực hiện trình bày một số kỹ thuật liên quan đến việc vẽ lại ảnh theo phong cách.

### 1. Không sử dụng mạng nơ-ron

Trước khi kỹ thuật nơ-ron phát triển, người ta đã nghĩ cách để vẽ lại những bức ảnh theo một phong cách cố định nào đó. Những kỹ thuật này thường chỉ áp dụng lên một ảnh đầu vào là ảnh 2D, chúng được gọi là kỹ thuật IB-AR (image-based artistic rendering)

#### 1.1. Stroke-Based Rendering

Stroke-Based Rendering là thuật toán cố gắng phác họa lại một bức ảnh dựa trên việc mô phỏng lại việc vẽ bằng tay của các họa sĩ. Thuật toán này được sử dụng để chuyển đổi ảnh kỹ thuật số thành các tác phẩm nghệ thuật với việc mô phỏng thành các nét vẽ chì hoặc sơn dầu <sup>2</sup>.



Hình 1: Mô tả thuật toán Stroke-Based Rendering

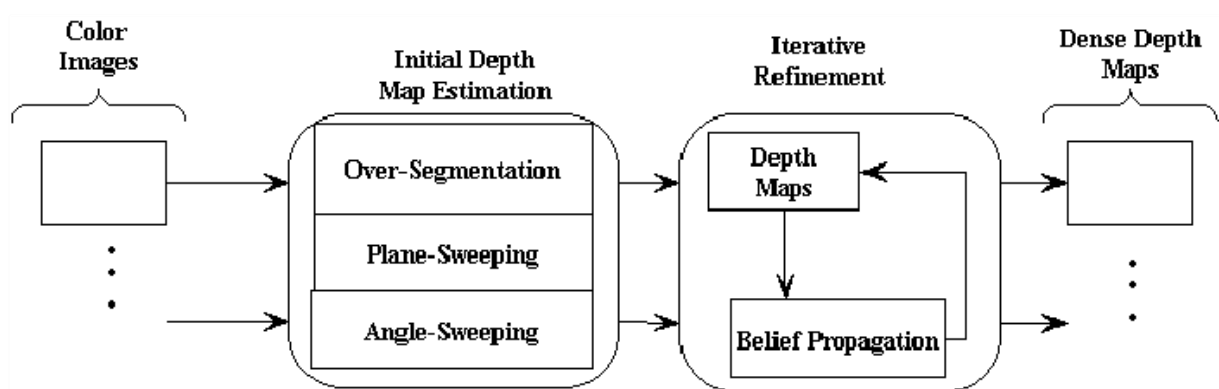
Nguyên lý hoạt động của kỹ thuật Stroke-Based Rendering là: đầu tiên, thuật toán sẽ phân tích ảnh đầu vào để lấy được đặc trưng các cạnh, màu sắc. Bước thứ hai, thuật toán sẽ mô phỏng các nét vẽ theo kỹ thuật vẽ tay. Cuối cùng là việc sắp xếp lại các nét vẽ theo cách tối ưu nhất.

Ưu điểm: Kỹ thuật này hoạt động tốt khi mô tả một bức ảnh số theo phong cách sơn dầu hoặc vẽ chì.

Nhược điểm: Kỹ thuật Stroke-Based Rendering giới hạn số phong cách. Ngoài ra, thuật toán này cũng gây ra chi phí tính toán lớn.

## 1.2. Region-Based Techniques

Một thuật toán vẽ lại ảnh khác, đó là Region-Based Techniques. Ý tưởng của thuật toán này là phân bức ảnh số thành từng vùng nhỏ khác nhau, khai thác thông tin của các vùng ấy để tạo nét vẽ<sup>3</sup>.



Hình 2: Mô tả thuật toán Region-Based Techniques

Region-Based Techniques ban đầu sẽ phân ảnh thành các vùng nhỏ hơn dựa theo màu sắc hoặc kết cấu bằng các phương pháp phân đoạn. Sau đó, các vùng này sẽ được phân tích thông tin để tạo nét vẽ riêng. Do đó, các đặc điểm của đối tượng có các nét vẽ được xây nhiều hơn tại các vùng biên của phân đoạn.

Ưu điểm của phương pháp:

- Từng vùng được xử lý riêng lẻ điều này giúp cho ảnh kết quả độc đáo hơn, do ảnh được điều chỉnh phù hợp cho từng phần
- Với mỗi cách phân vùng khác nhau, sẽ tạo ra những phong cách khác nhau, làm tăng tính đa dạng của kết quả

Nhược điểm:

- Do phải thực hiện công việc phân vùng, nên sẽ tăng thêm chi phí tính toán cho thuật toán

### 1.3. Example-Based Rendering

Để có thể kiểm soát được phong cách vẽ của ảnh nghệ thuật, Hertzmann và đồng nghiệp đã đề xuất và xây dựng 1 framework tên là “*image analogies*”, nhằm ánh xạ từ 1 cặp ảnh nguồn và đích <sup>4</sup>. Kỹ thuật này là một phương pháp học có giám sát, nó cố gắng học các phong cách của các ảnh mục tiêu, sau đó áp dụng vào những ảnh đầu vào để tạo ra kết quả như mong muốn.

Thuật toán bao gồm các bước: đầu tiên cần chuẩn bị bộ dữ liệu để mô hình có thể lấy đặc trưng phong cách. Bước tiếp theo, mô hình sẽ lấy phong cách của ảnh dữ liệu. Cuối cùng là áp dụng phong cách lên ảnh đầu vào để thu được kết quả như mong muốn.



Hình 3: Mô tả kết quả của kỹ thuật example-based rendering

Ưu điểm:

- Example-based rendering có thể tạo ra các kết quả tự nhiên và chân thực
- Linh hoạt theo nhiều phong cách, do nó có thể học theo nhiều mẫu khác nhau
- Có thể mở rộng và áp dụng nhiều phong cách khác nhau bằng cách đưa thêm nhiều dữ liệu

Nhược điểm:

- Example-based rendering chỉ tận dụng được đặc trưng của ảnh cấp thấp, do đó không mô tả hết về phong cách nghệ thuật
- Chi phí tính toán cao trong quá trình khớp với phong cách gốc

### 1.4. Tổng kết

Dựa trên những tìm hiểu trên, nhóm thực hiện đưa ra một số nhận xét về các thuật toán IB-AR như sau.

Ưu điểm:

- Phản ánh trung thực khi mô tả với các nét vẽ sơn dầu, vẽ cọ,...
- Dễ dàng áp dụng vào các ứng dụng đồ họa máy tính
- Dễ dàng thử nghiệm và điều chỉnh

Nhược điểm:

- Một số kỹ thuật đã cố gắng đa dạng phong cách, nhưng vẫn còn rất hạn chế
- Không trích xuất cấu trúc phong cách một cách hiệu quả
- Khả năng bị giới hạn trong tính toán cục bộ

Để khắc phục những nhược điểm trên, mạng nơ-ron ra đời giúp quá trình tạo ảnh nghệ thuật đạt được tính hiệu quả cao hơn.

## 2. Sử dụng mạng nơ-ron

Kể từ khi mạng nơ-ron tích chập thực nghiệm thành công ở các lĩnh vực khác nhau, các nhà nghiên cứu đã cố gắng sử dụng mạng tích chập trong việc vẽ lại ảnh theo phong cách. Việc sử dụng mạng nơ-ron đã tạo ra những kết quả ấn tượng với các nhà khoa học, các nhà mỹ thuật về những gì mạng nơ-ron làm được. Các kỹ thuật lúc này đã có thể nhận đầu vào bao gồm hai hoặc nhiều ảnh hơn, trong đó có ảnh nội dung và ảnh phong cách. Các mô hình được xây dựng nhằm lấy ngữ cảnh của ảnh nội dung và hòa trộn với phong cách nghệ thuật của ảnh phong cách. Mô hình sử dụng mạng nơ-ron đầu tiên đã được Gatys giới thiệu <sup>1</sup>, cho đến nay đã có nhiều cải tiến từ mô hình này.



Hình 4: Vẽ lại ảnh theo phong cách

Mô hình của thuật toán style transfer nhận đầu vào là một ảnh nội dung (ảnh content) và một ảnh phong cách (ảnh style) Nhiệm vụ của mô hình là tạo ra một ảnh đầu ra có sự kết hợp của hai ảnh đầu vào. Do đó, mô hình học sâu cần thực hiện các nhiệm vụ sau:

- Lấy được nội dung của ảnh content
- Lấy được phong cách của ảnh style
- Kết hợp và tạo ra ảnh phong cách

Trong phần này, nhóm thực hiện sẽ trình bày các phương pháp thực hiện các nhiệm vụ của thuật toán style transfer.

## 2.1. Lấy nội dung của ảnh

### 2.1.1. Sử dụng bộ lọc truyền thống

Để trích xuất đặc trưng của một ảnh đầu vào, có thể sử dụng các bộ lọc truyền thống như Gaussian, Sobel,... Những bộ lọc này do các kỹ sư hoặc chuyên gia xử lý ảnh, qua quá trình và kinh nghiệm làm việc đã lấy được một bộ tham số phù hợp để áp dụng lên các ảnh đầu vào nhằm lấy ra đặc trưng của ảnh.

-1	0	+1
-2	0	+2
-1	0	+1

Gx

+1	+2	+1
0	0	0
-1	-2	-1

Gy

Hình 5: Ma trận của bộ lọc sobel

Ưu điểm: Phương pháp này, sẽ cho kết quả khá chính xác với trường hợp cụ thể của bài toán, nơi mà các tham số của các bộ lọc được chứng minh là hiệu quả. Ngoài ra, nó còn xử lý rất nhanh mà không cần phải thực hiện nhiều tính toán.

Nhược điểm: Tuy nhiên, điểm yếu của phương pháp này là nó chỉ hiệu quả với những trường hợp cụ thể, chứ không mang tính chất tổng quát cho mọi bài toán.

Trích chọn đặc trưng với sobel:

```

import cv2
import numpy as np

def sobel_feature_extraction(image):
    # Chuyển ảnh về ảnh đen trắng
    gray_image = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)

    # Áp dụng filter Sobel cho ảnh
    sobel_x = cv2.Sobel(gray_image, cv2.CV_64F, 1, 0, ksize=3)
    sobel_y = cv2.Sobel(gray_image, cv2.CV_64F, 0, 1, ksize=3)

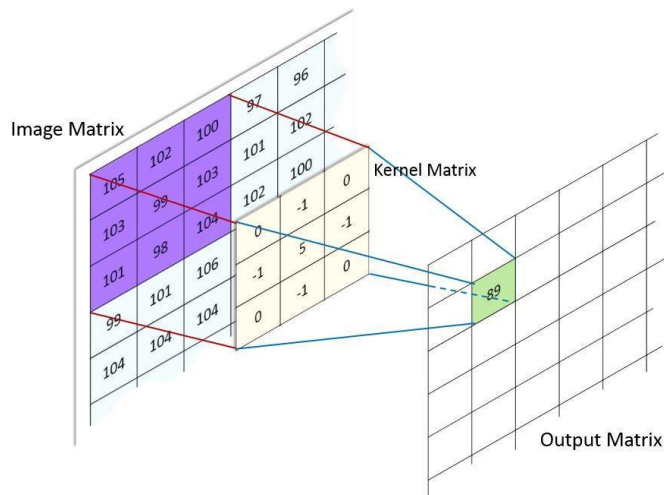
    # Tính độ lớn gradient và hướng gradient
    gradient_magnitude = np.sqrt(sobel_x**2 + sobel_y**2)
    gradient_direction = np.arctan2(sobel_y, sobel_x)

    return gradient_magnitude, gradient_direction

```

### 2.1.2. Sử dụng convolution

Để khắc phục nhược điểm của những bộ lọc truyền thống, các nhà nghiên cứu đã sử dụng các bộ lọc của mạng tích chập nhằm lấy ra đặc trưng của hình ảnh.



Hình 6: Phép toán convolution trên ảnh

Một mạng tích chập CNN bao gồm nhiều khối khác nhau tùy vào thiết kế của mạng. Trong đó, có các kernel được khởi tạo là những ma trận số dùng để lấy ra đặc trưng của ảnh đầu vào. Ma trận này có thể sử dụng những filter truyền thống như Sobel, Gaussian,... hoặc cũng có thể khởi tạo ngẫu nhiên. Khối kernel sẽ trượt theo bức ảnh, tính tích vô hướng tại thành phần mà nó trượt qua. Kết quả được tổng hợp lại gọi là bản đồ đặc trưng (output matrix) chứa đặc trưng của bức ảnh.

Mỗi kiến trúc mạng lại có nhiều khối kernel khác nhau, cho phép phân tích ảnh theo đa chiều. Mặt khác các phép tính này lại tự động học đặc trưng ảnh, bao gồm các

đặc trưng cấp thấp và cấp cao. Nó cho phép lấy được những đặc trưng phức tạp của bức ảnh. Những layer đầu tiên của mạng sẽ lấy ra đặc trưng cơ bản nhất như cạnh, góc,... Càng layer về sau, những đặc trưng trừu tượng hơn sẽ được lấy ra.

## 2.2. Lấy phong cách của ảnh

### 2.2.1. Sử dụng tham số thống kê

Một phương pháp lấy đặc trưng của phong cách thường được sử dụng trong các mô hình thuật toán của style transfer là sử dụng các layer của CNN. Giống như việc lấy đặc trưng của ảnh nội dung, việc lấy đặc trưng cho phong cách cũng có thể lợi dụng các kernel của kiến trúc mạng tích chập. Tuy nhiên, việc áp dụng kernel để lấy đặc trưng cần có sự biến đổi để phù hợp với nội dung bài toán.

Mục tiêu của style transfer là trộn ảnh nội dung và phong cách với nhau để tạo ra một bức ảnh nghệ thuật. Trong đó, thành phần nội dung của ảnh đầu ra không nhất định phải sao chép hoàn toàn, chỉ cần nội dung theo ngữ cảnh của bức ảnh. Thành phần phong cách sẽ lấy đặc trưng từ ảnh phong cách và không nhất thiết phải giữ đúng vị trí của nó như ban đầu. Các thành phần phong cách có mặt trong ảnh đầu ra có thể ở một vị trí bất kỳ nào đó.

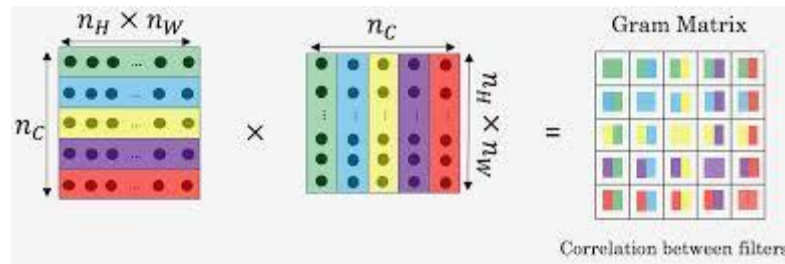
Do đó, các nghiên cứu về style transfer đã giới thiệu một hàm mất mát là Gram-matrix nhằm lấy ra đặc trưng về phong cách trong ảnh style. Công thức Grammatrix trong mục đích lấy đặc trưng phong cách được mô tả như sau:

$$G(F^l(I_s)') = [F^l(I_s)'] [F^l(I_s)']^T \quad (1)$$

Trong đó,  $F^l(I_s)$  được hiểu là feature map của ảnh style tại layer thứ  $l$  của mô hình.  $F^l(I_s)'$  là một phiên bản reshape của  $F^l(I_s)$ . Và  $[F^l(I_s)']^T$  là ma trận chuyển vị của  $F^l(I_s)'$ .

Trong đó, công thức Grammatrix được định nghĩa trong toán học như là tích vô hướng của các vector trong không gian O-clit.





Hình 7: Mô tả công thức Gram-matrix

Công thức Gram-matrix có thể được sử dụng để hỗ trợ việc lấy đặc trưng của phong cách, vì nó có thể lấy giá trị của đặc trưng, Gram matrix không quan tâm đến vị trí cụ thể của chúng trên bức ảnh.

Triển khai gram-matrix trên Python:

```
import tensorflow as tf

def gram_matrix(features):
    # Tính toán số lượng feature maps và kích thước của mỗi feature map
    batch_size, height, width, channels = tf.shape(features)

    # Reshape features để có kích thước batch_size * (height * width) * channels
    features = tf.reshape(features, (batch_size, height * width, channels))

    # Tính ma trận Gram bằng cách nhân ma trận tích vô hướng của features với chuyển vị của nó
    gram = tf.matmul(features, features, transpose_a=True)

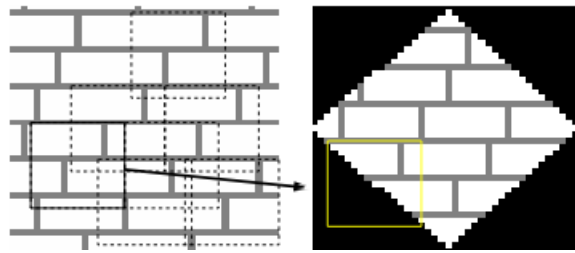
    # Chuẩn hóa ma trận Gram bằng cách chia cho số phần tử trong mỗi feature map
    gram /= tf.cast(height * width * channels, tf.float32)

    return gram
```

### 2.2.2. Không sử dụng tham số

Ngoài phương pháp sử dụng tham số để lấy đặc trưng style. Các nhà nghiên cứu cũng giới thiệu một cách tiếp cận khác không sử dụng tham số dựa trên mô hình Markov Random Fields. Họ cho rằng, trong một bức ảnh, mỗi điểm ảnh lại được xác định bởi những điểm ảnh xung quanh nó.

Do đó, một thuật toán tổng hợp từng điểm ảnh bằng cách tìm kiếm các vùng lân cận và gán cho điểm ảnh đó được Efros và Leung đề xuất <sup>5</sup>.



Hình 8: Quá trình tổng hợp điểm ảnh bằng cách tìm lân cận

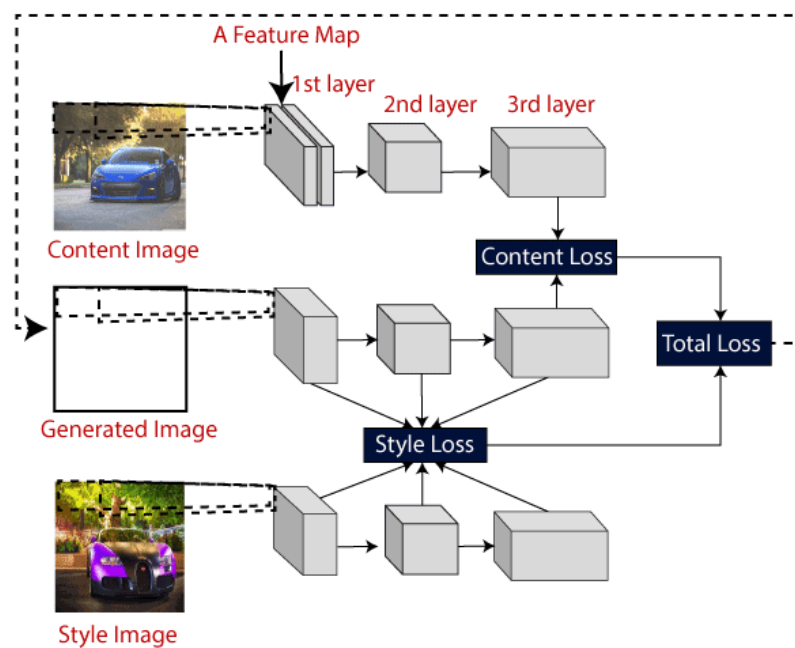
Ưu điểm:

- Phương pháp này không yêu cầu tham số, nên có thể giảm bớt gánh nặng cho mô hình
- Không cần thông tin trước của phong cách cụ thể, do nó lấy thông tin dựa vào những điểm ảnh xung quanh, giúp giảm khó khăn khi triển khai mô hình

Nhược điểm:

- Không thể lấy đặc trưng có tính trừu tượng cao
- Khả năng khớp mẫu kém

### 2.3. Tái tạo hình ảnh



Hình 9: Kiến trúc một mô hình thuật toán Style transfer

Sau khi lấy được thông tin về nội dung và phong cách, công việc tiếp theo là trộn hai thông tin đó lại với nhau để lấy được hình ảnh đầu ra là một ảnh nghệ thuật có chứa nội dung của bức ảnh content và phong cách của bức ảnh style. Có hai phương pháp tái tạo hình ảnh, đó là tái tạo ảnh dựa trên tối ưu hóa hình ảnh và tái tạo ảnh dựa trên tối ưu hóa mô hình.

### 2.3.1. Tái tạo ảnh dựa trên tối ưu ảnh

Mục tiêu: thuật toán cố gắng tối ưu hóa hàm mục tiêu dựa vào hình ảnh đầu vào.

Nội dung thuật toán: Ảnh đầu vào thông qua các vòng lặp có nhiệm vụ tối ưu hóa hình ảnh đầu ra sao cho thỏa mãn hàm mất mát. Hàm mất mát sẽ mô tả, đo lường sự khác biệt giữa ảnh tái tạo và ảnh thực. Việc tối ưu hóa tham số có thể sử dụng các kỹ thuật tối ưu như gradient descent hoặc adam,...

Ưu điểm:

- Chất lượng hình ảnh cao
- Có tính linh hoạt do mạng tích chập có tính tự học cao
- Các tham số được cập nhật liên tục

Nhược điểm:

- Phụ thuộc vào mô hình có phù hợp với dữ liệu không
- Tốn tài nguyên thực thi
- Có thể bị tối ưu cục bộ

### 2.3.2. Tái tạo ảnh dựa trên tối ưu mô hình

Mục tiêu:

- Xây dựng một mô hình và tối ưu các tham số của mô hình trước khi sử dụng với hình ảnh
- Huấn luyện mô hình với bộ dữ liệu lớn để lấy bộ tham số, qua đó sử dụng nó cho những dữ liệu thực cụ thể

Nội dung thuật toán: thuật toán trải qua hai giai đoạn là xây dựng mô hình và tái tạo hình ảnh. Ở bước xây dựng mô hình, các tham số, layer được lựa chọn sao cho phù hợp nhất. Sau đó, thực hiện huấn luyện mô hình trên bộ dữ liệu lớn để có một bộ tham số có tính tổng quát cho các trường hợp. Quá trình tái tạo ảnh, sẽ sử dụng mô hình được xây

dựng trước đó cùng với bộ tham số đã thu được nhằm cải thiện tốc độ cho quá trình xử lý ảnh

Ưu điểm:

- Hiệu suất được cải thiện do việc xây dựng mô hình và huấn luyện được thực hiện trước đó
- Có tính ổn định: do tham số được khởi tạo có tính tổng quát cao sẽ ảnh hưởng đến quá trình thực hiện của thuật toán

Nhược điểm:

- Dữ liệu huấn luyện: việc xây dựng bộ dữ liệu phải đủ lớn, đa dạng và có tính chất tổng quát cao
- Không linh hoạt: với trường hợp hình ảnh thực không có nhiều sự tương đồng với bộ dữ liệu đã huấn luyện trước đó, việc tái tạo sẽ khó khăn về mặt chất lượng hình ảnh