

Ngày nhận hồ sơ			
(Do CO quản lý ghi)			

THUYẾT MINH ĐỀ TÀI KHOA HỌC VÀ CÔNG NGHỆ CẤP SINH VIÊN 2017

A. THÔNG TIN CHUNG

A1. Tên đề tài

- Tên tiếng Việt: **Tìm hiểu và đánh giá một số mô hình học sâu cho bài** toán phát hiện vật thể trong tác phẩm nghệ thuật
- Tên tiếng Anh: **Research and evaluate some deep learning models for object detection in artwork**

A2. Loại hình nghiên cứu

(Tham khảo tiêu chuẩn đề tài đối với từng loại hình NC, chọn 01 trong 03 loại hình)

- ⊠ Nghiên cứu cơ bản
- □ Nghiên cứu ứng dụng
- □ Nghiên cứu triển khai

A3. Thời gian thực hiện

..06.. tháng (kể từ khi được duyệt).

A4. Tổng kinh phí

(Lưu ý tính nhất quán giữa mục này và mục B8. Tổng hợp kinh phí đề nghị cấp)

Tổng kinh phí: ...5.. triệu đồng, gồm

• Kinh phí từ Trường Đại học Công nghệ Thông tin: ..5.. triệu đồng

A5. Chủ nhiệm

Họ và tên: Trần Trung Anh.

Ngày, tháng, năm sinh: 01/05/2000 . Giới tính (Nam/Nữ): Nam

Số CMND: ; Ngày cấp: ; Nơi cấp:

Mã số sinh viên: 18520473

Số điện thoại liên lạc: 0869372167 .

Đơn vị (Khoa hoặc BM KH&KTTT): Khoa học máy tính

A7. Nhân lực nghiên cứu

TT	Ho tên	MSSV	Khoa/ Bộ Môn
1	Trần Trung Anh	18520473	KHMT
2	Nguyễn Phú Quốc	18520343	KHMT

B. MÔ TẢ NGHIÊN CỬU

B1. Giới thiệu về đề tài

(Ghi các ý về tổng quan tình hình nghiên cứu liên quan đến đề tài, lí do thực hiện đề tài, các thách thức) B1.1 Tổng quan tình hình nghiên cứu

Bài toán Phát hiện vật thể (Object Detection) là một trong những bài toán phổ biến, quan trọng và được nghiên cứu chuyên sâu trong lĩnh vực Thị giác máy tính. Một trong những bài toán còn nhiều thách thức của lĩnh vực này là bài toán Phát hiện một vật thể được biểu diễn dưới nhiều phong cách nghệ thuật khác nhau trên nhiều cách vẽ khác nhau [1]. Vấn đề này gọi là *cross-depiction* [2]: phát hiện vật thể bất kể chúng được mô tả như thế nào (ảnh chụp, tranh sơn màu, tranh vẽ tay,...).

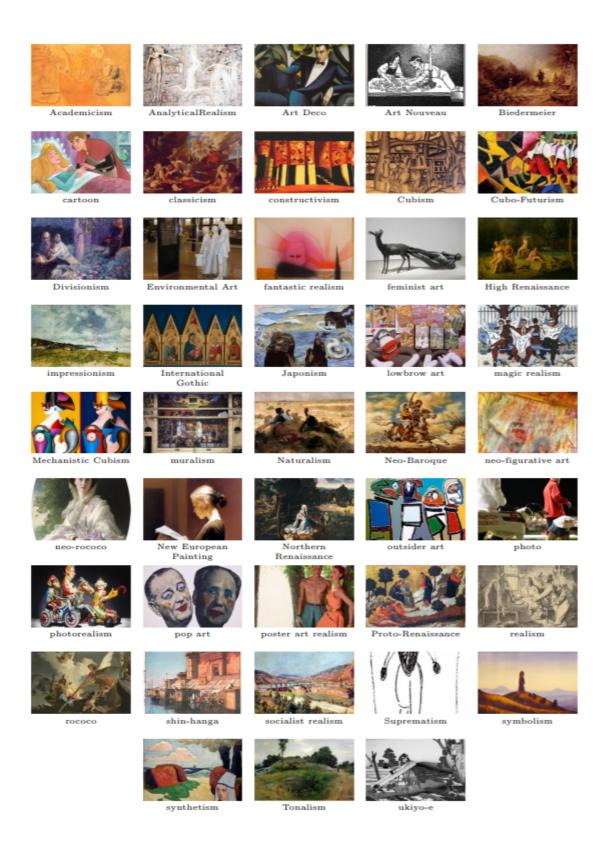


Hình 1. Cross-depiction problem

Hàng trên: Cross-depiction giữa các phong cách tranh vẽ Hàng dưới: Cross-depiction giữa các ảnh chụp

Giải quyết được bài toán phát hiện vật thể trên nhiều phong cách thể hiện của chúng có thể hỗ trợ cho các ứng dụng thực tế, đặc biệt là tìm kiếm hình ảnh trên internet. Trên thế giới hiện nay có một khối lượng khổng lồ các hình ảnh, tranh vẽ được biểu diễn trên nhiều phong cách khác nhau. Trong khi đó, đa số các phương pháp Thị giác máy tính lại chỉ tập trung xử lý trên dữ liệu là ảnh chụp thực tế, dẫn đến các hệ thống phát hiện vật thể bị giới hạn hiệu suất trên những dữ liệu tranh ảnh nghệ thuật. Bên cạnh đó, hiện nay với sự gia tăng về nhu cầu sử dụng công nghệ phát hiện vật thể trong các viện bảo tàng, thư viện tranh vẽ điện tử [3] thì việc có một hệ thống giúp người xem có thể tìm kiếm tranh vẽ bằng một đối tượng cụ thể là vấn đề cấp thiết. Một ví dụ khác, một hệ thống tìm kiếm đa dạng các loại hình ảnh có thể giúp cho các hoạ sĩ, nhà sáng tạo nội dung tạo ra những ý tưởng, cách mô tả hình ảnh mới dựa trên những phong cách mô tả hình ảnh đã có.

Trong những nghiên cứu gần đây, mặc dù đã có những bước tiến bộ lớn trong việc giải quyết bài toán phát hiện vật thể dựa trên các kỹ thuật Học sâu thay vì chỉ sử dụng các đặc trưng cấp thấp, nhưng vẫn chưa có nhiều nghiên cứu, đánh giá hiệu suất của chúng trong việc phát diện vật thể trong nhiều loại hình nghệ thuật khác nhau. Trong bài nghiên cứu này, chúng tôi tiến hành thực nghiệm một số mô hình Học sâu, đã đạt được nhiều kết quả đáng kể trên các bộ dữ liệu ảnh chụp thực tế, trên một bộ dữ liệu tranh vẽ, cụ thể là *PeopleArt* [1]. Đây là bộ dữ liệu bao gồm các tranh ảnh vẽ con người thuộc 43 trường phái nghệ thuật khác nhau nên phù hợp để thực nghiệm, đánh giá các mô hình Học sâu để giải quyết vấn đề cross-depiction. Ngoài ra, vì số lượng tranh ảnh thuộc mỗi trường phái nghệ thuật còn trong bộ dữ liệu còn ít và có nhiều tranh vẽ rất khó nhận ra là con người nên chúng tôi đề xuất phương pháp tăng cường dữ liệu bằng cách sử dụng các mô hình Style Transfer [4], theo đó chuyển các hình ảnh con người trên thực tế thành các tranh ảnh nghệ thuật thuộc 43 trường phái trong bộ dữ liệu gốc.



Hình 2. Bộ dữ liệu PeopleArt

B1.2 Các thách thức

- Các đặc trưng của con người hoặc các vật thể trong tranh vẽ có thể không giống đặc trưng của con người hoặc vật thể trên thực tế như màu da, khuôn mặt, mắt,...



Hình 3. Ẩnh thực tế (Bên phải) và Chân dung tự vẽ (Bên trái) của Pablo Picasso

Sự khác biệt lớn về màu da, hình dạng mắt, mũi, miệng của tranh vẽ và người trên thực tế sẽ khó khăn trong việc lúc quá trình huấn luyện.

- Bộ dữ liệu PeopleArt chỉ có dữ liệu về ảnh người trong tranh vẽ nên số lượng đối tượng không đa dạng, dẫn đến hiệu quả không cao nếu được sử dụng để xây dựng hệ thống tìm kiếm tranh ảnh.

B2. Mục tiêu, nội dung, kế hoạch nghiên cứu

B2.1 Mục tiêu

Trong phạm vi đề tài nghiên cứu này, chúng tôi sẽ:

- Tìm hiểu tổng quan bài toán và các cách tiếp cận bài toán phát hiện vật thể trong các tác phẩm nghệ thuật.
- Cài đặt và đánh giá hiệu suất của các mô hình Học sâu cho bài toán trên.
- Xây dựng ứng dụng tìm kiếm hình ảnh để minh hoạ.

B2.2 Nội dung và phương pháp nghiên cứu

Nội dung 1: Tìm hiểu một số mô hình Học sâu liên quan

Phương pháp:

- Tìm hiểu một số mô hình Học sâu có thể áp dụng cho bài toán, cụ thể là Faster R-CNN, Mask R-CNN và YOLOv5. Các mô hình này đã được thực nghiệm và đạt hiệu quả cao trong các bài toán Phát hiện vật thể thực tế, tuy nhiên chưa có nhiều đánh giá trên dữ liệu tranh ảnh nghệ thuật.

Dự kiến kết quả:

- Báo cáo kỹ thuật về các nghiên cứu liên quan đến đề tài.
- Tài liệu kỹ thuật cho các mô hình Học sâu đã tìm hiểu.

Nội dung 2: Tìm hiểu và thử nghiệm các phương pháp giải quyết vấn đề cross-depiction

Phương pháp:

- Tìm hiểu một số hướng giải quyết vấn đề cross-depiction. Trong đó, chúng tôi dự định sẽ tăng cường dữ liệu bằng cách sử dụng Style Transfer lên những hình ảnh con người thực tế để có thêm dữ liệu tranh ảnh nghệ thuật cho quá trình huấn luyện.
- Cài đặt, thử nghiệm các hướng giải quyết.

Dự kiến kết quả:

- Báo cáo kết quả tổng quan cho quá trình tìm hiểu các phương pháp.
- Đánh giá các phương pháp phát hiện đối tượng đối với tác phẩm nghệ thuật.

Nội dung 3: Cài đặt, thực nghiệm và so sánh các mô hình Học sâu đã chọn

Phương pháp:

- Cài đặt, thực nghiệm với các mô hình trên với bộ dữ liệu PeopleArt, sau khi
 đã tăng cường dữ liệu.
- Đánh giá, so sánh hiệu suất của các mô hình trên các thông số average precision: AP, AP^{50} .

Dự kiến kết quả:

- Chương trình minh họa cho mỗi mô hình.
- Tài liệu kỹ thuật của các mô hình đã sử dụng.
- Báo cáo hiệu suất của các mô hình đã sử dụng.

B3. Kết quả nghiên cứu

Nghiệm thu đề tài với các kết quả sau:

- Tài liệu báo cáo tìm hiểu về bài toán và các cách tiếp cận.
- Báo cáo tổng hợp về các kết quả, đánh giá các mô hình đã sử dụng.
- Xây dựng chương trình minh hoạ.

B4. Tài liệu tham khảo

- [1] Westlake, Nicholas & Cai, Hongping & Hall, Peter. (2016). Detecting People in Artwork with CNNs. 9913. 825-841. 10.1007/978-3-319-46604-0_57.
- [2] Cai, Hongping & Wu, Qi & Corradi, Tadeo & Hall, Peter. (2015). The Cross-Depiction Problem: Computer Vision Algorithms for Recognising Objects in Artwork and in Photographs.
- [3] B. Ciecko, "AI Sees What? The Good, the Bad, and the Ugly of Machine Vision for Museum Collections," in Museums and the Web 2020. Online: Museums and the Web, 2020.

- [3] L. A. Gatys, A. S. Ecker, and M. Bethge. A neural algorithm of artistic style. arXiv preprint arXiv:1508.06576, 2015.
- [5] Adrian Lecoutre, Benjamin Negrevergne, Florian Yger. Recognizing Art Style Automatically with deep learning. Proceedings of Machine Learning Research, PMLR, 2017.

Ngày __ tháng __ năm 20_ Chủ nhiệm đề tài (Ký và ghi rõ họ tên) Ngày __ tháng __ năm 20_ Giảng viên hướng dẫn (Ký và ghi rõ họ tên)