

IMAGE RETRIEVAL

Môn: Xử lý ảnh và ứng dụng - CS406

Danh sách thành viên

1. Lê Trần Bảo Lợi	-	21522295
2. Ngô Đức Hoàng Hiệp	-	21520846
3. Bùi Đình Quân	-	21522487

Table of content

I	Giới thiệu bài toán	3
II	Dataset	8
III	Môi trường thực nghiệm	14
IV	Phương pháp thực hiện	17
V	Thực nghiệm & đánh giá	20
VI	Demo	41
VII	Tài liệu tham khảo	42

Giới thiệu bài toán

Giới thiệu bài toán

- Image retrieval là hệ thống tìm kiếm ảnh dựa trên **ảnh tham khảo**.
- Tính hiệu quả của hệ thống:
 - Tối ưu về không gian lưu trữ.
 - Tốc độ tìm kiếm nhanh.
 - Kết quả tìm kiếm phù hợp.
 - Tiện dụng.

Động lực thực hiện

- Tìm kiếm luôn là nhu cầu cần thiết của con người.
- Khối lượng dữ liệu ngày càng lớn, phức tạp và đa dạng.
- Ảnh là kiểu dữ liệu phổ biến có nhu cầu tìm kiếm cao.

Mô tả bài toán

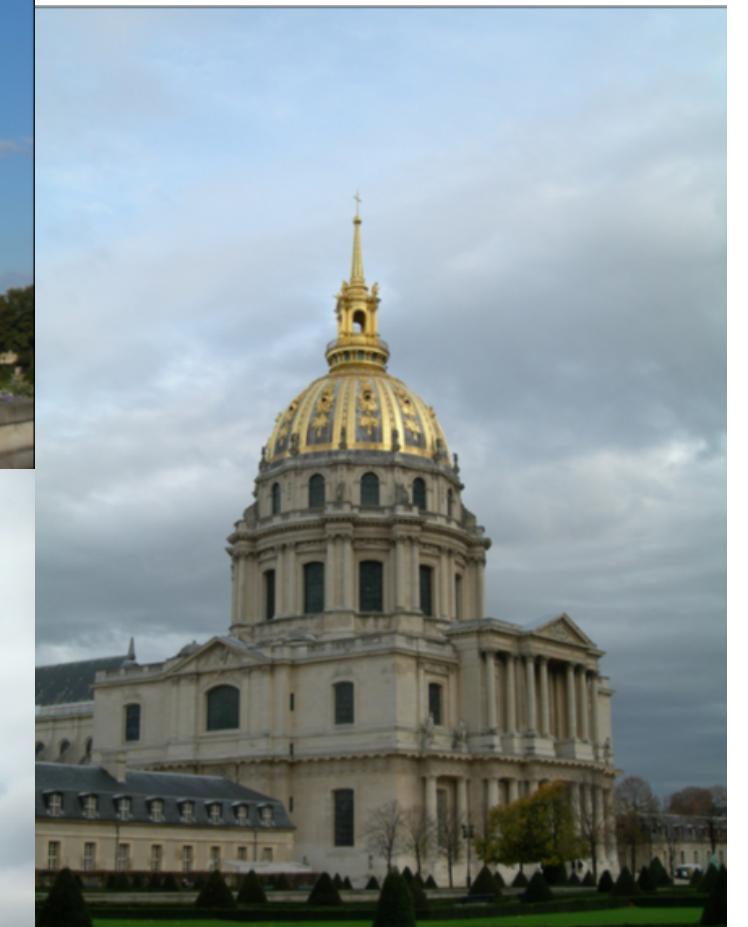
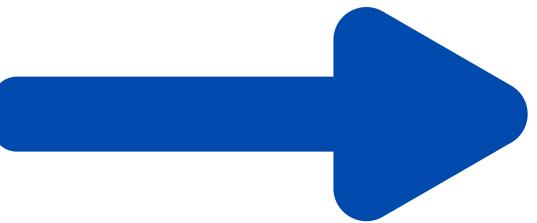
Input

- Một hình ảnh truy vấn.
- Tập dữ liệu hình ảnh mà hệ thống truy vấn sẽ tìm kiếm.

Output

- Danh sách xếp hạng các hình ảnh tương đồng hoặc liên quan đến hình ảnh truy vấn trong tập dữ liệu.

Ví dụ



Input

Output

Dataset

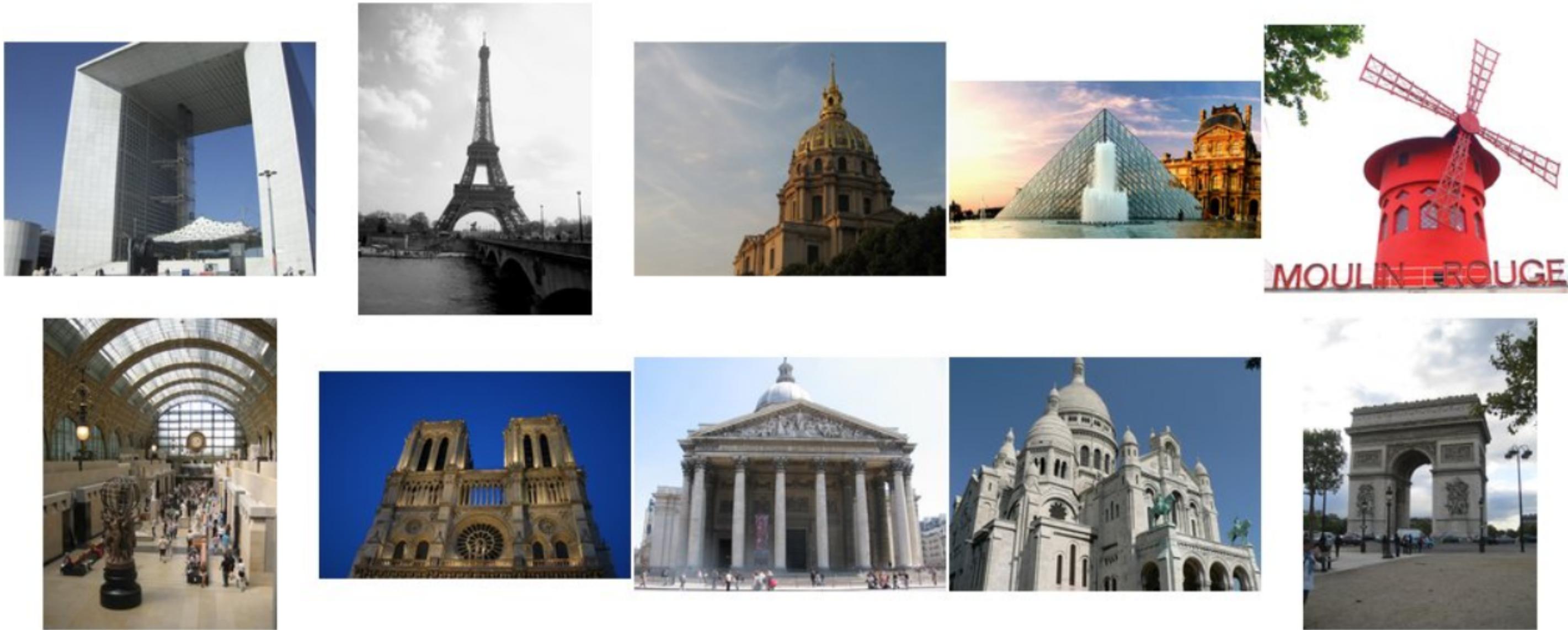
revisited Paris dataset

- Benchmark dataset cho việc đánh giá content-based image retrieval (CBIR) systems.
- Ảnh được thu thập từ Flickr của 11 địa danh của Pháp được đăng tải trên mạng xã hội.
- Được gán nhãn lại trên tập Paris giải quyết các trường hợp bị sai ở tập ground truth và mở rộng tập query.

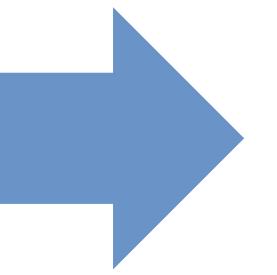
Cấu trúc dataset

- Dataset gồm: 6412 ảnh.
- Số lượng query: 70 ảnh query.
- Mỗi ảnh query được con người gán thủ công ground truth tương ứng.

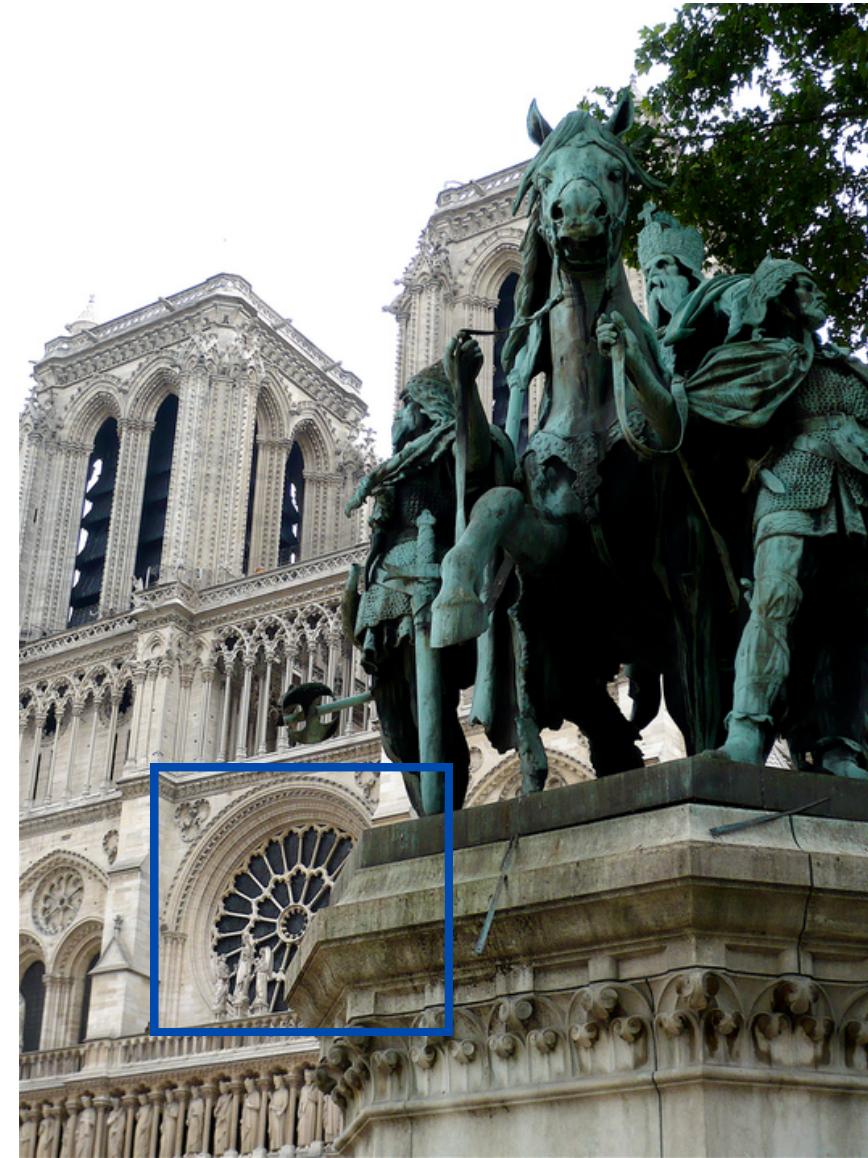
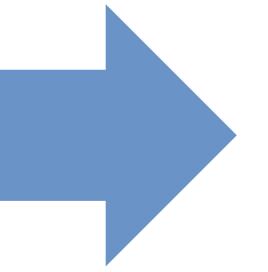
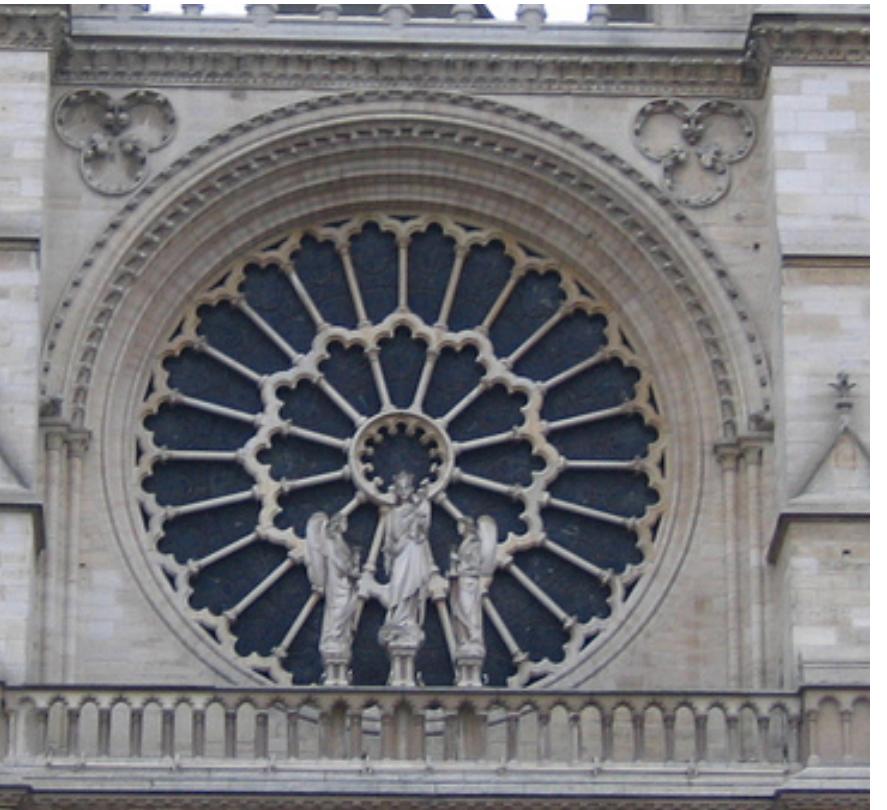
Một số hình ảnh trong tập dataset



Một số trường hợp phức tạp



Một số trường hợp phức tạp



Môi trường thực nghiệm

Môi trường thực nghiệm

- Google cung cấp một môi trường thực nghiệm trực tuyến miễn phí đó là Google Colab.
- Colab giúp làm việc nhóm hiệu quả và dễ sử dụng.



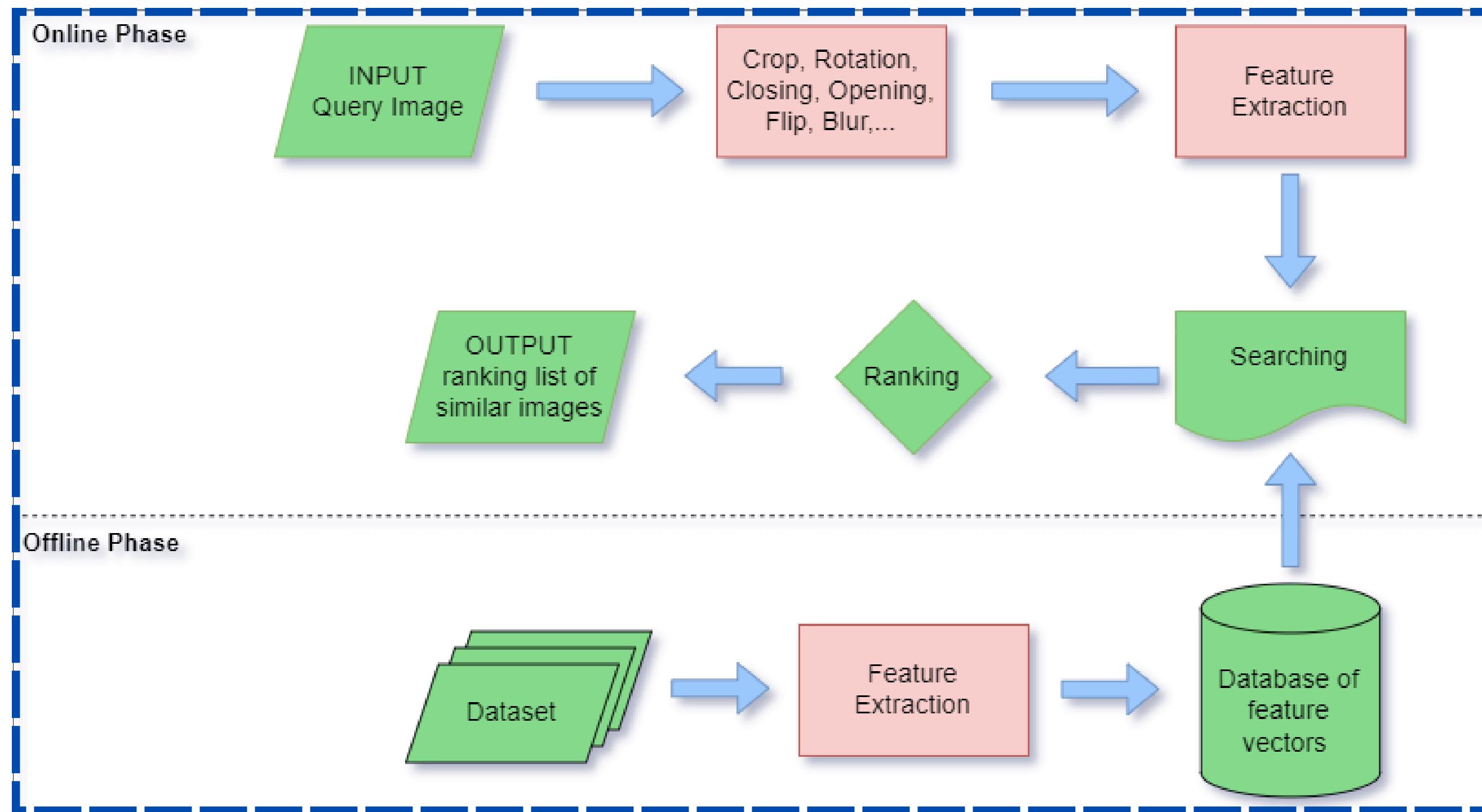
Độ đo đánh giá

- mAP

$$MAP = \frac{1}{Q} \sum_{q=1}^Q AP(q)$$

Phương pháp thực hiện

Phương pháp



Các phép biến đổi Input Query

- Crop
- Sharpen
- Smooth
- Blur
- Rotate
- Flip

Thực nghiệm & đánh giá

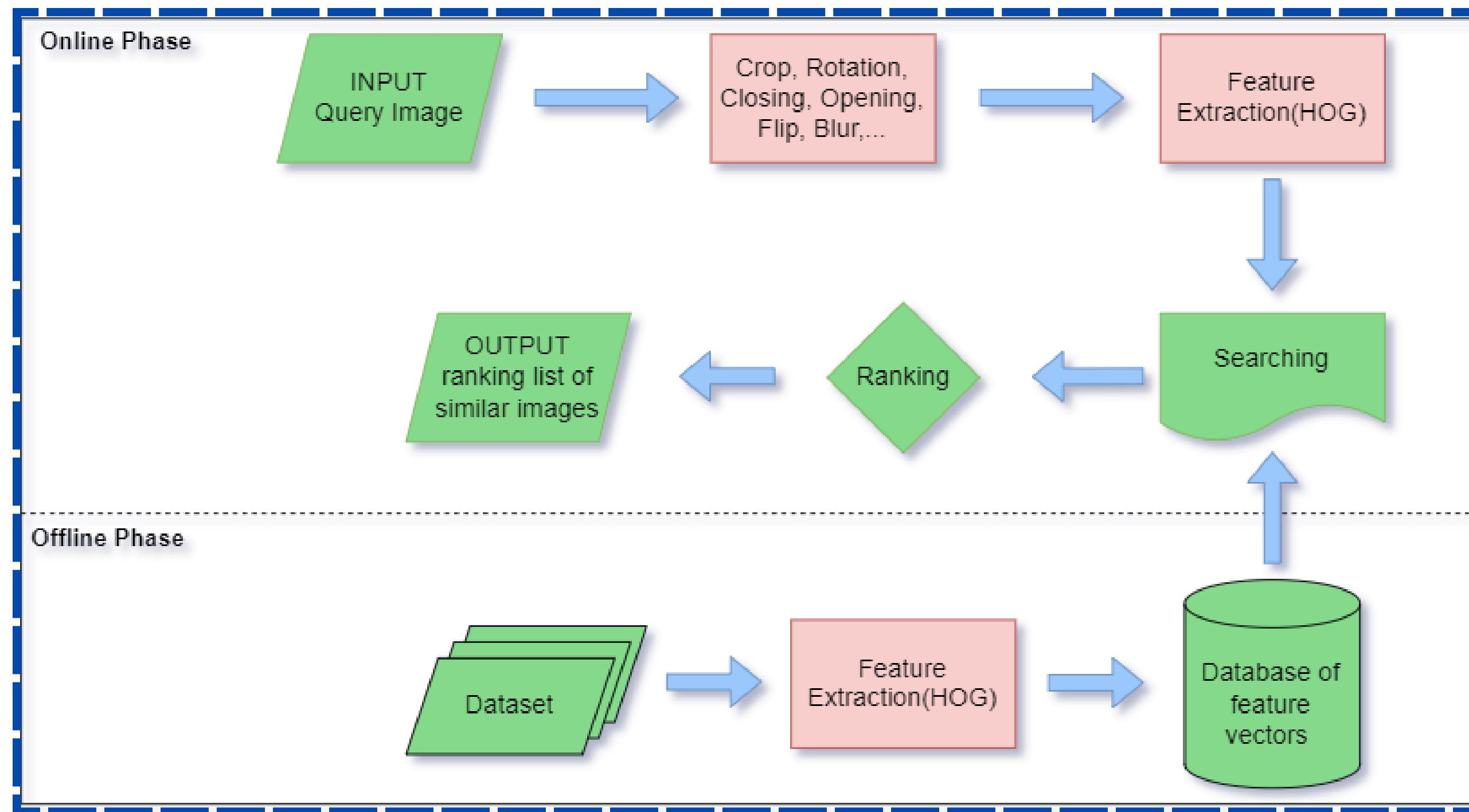
Một số thực nghiệm

1. Mở rộng ảnh truy vấn + đặc trưng HOG.
2. Mở rộng ảnh truy vấn + đặc trưng Resnet.
3. Tăng cường dữ liệu + đặc trưng Resnet.

Trích xuất đặc trưng với HOG

- Histogram of Oriented Gradient (HOG) : là một phương pháp mô tả đặc trưng
- HOG: tạo ra các bộ mô tả đặc trưng (feature descriptor) nhằm mục đích phát hiện đối tượng (object detection).

Mở rộng ảnh truy vấn + đặc trưng HOG



Mở rộng ảnh truy vấn + đặc trưng HOG

Tham số của HOG descriptor:

- Số lượng bin: 9.
- Kích thước cell: 32x32.
- Kích thước block: 1x1.
- Chuẩn hóa theo block: L2.

Kết quả

	ORIGINAL	CROP	RANDOM CROP	RANDOM ROTATE	HORIZONTAL FLIP	BLUR	SHARPEN	SMOOTH
mAP@5	0.5779	0.443	0.5378	0.4740	0.5708	0.5720	0.4593	0.5748
mAP@20	0.4884	0.3685	0.4242	0.4137	0.4870	0.4849	0.3725	0.4880
mAP@50	0.4117	0.2934	0.3489	0.3487	0.4057	0.4079	0.3395	0.4116

Trích xuất đặc trưng với Resnet50

- ResNet50 là một deep neural network.
- ResNet50 đã được công bố vào năm 2015 trong bài báo "Deep Residual Learning for Image Recognition" tại hội nghị CVPR.

Deep residual learning for image recognition

[K He, X Zhang, S Ren, J Sun - ... and pattern recognition, 2016 - openaccess.thecvf.com](#)

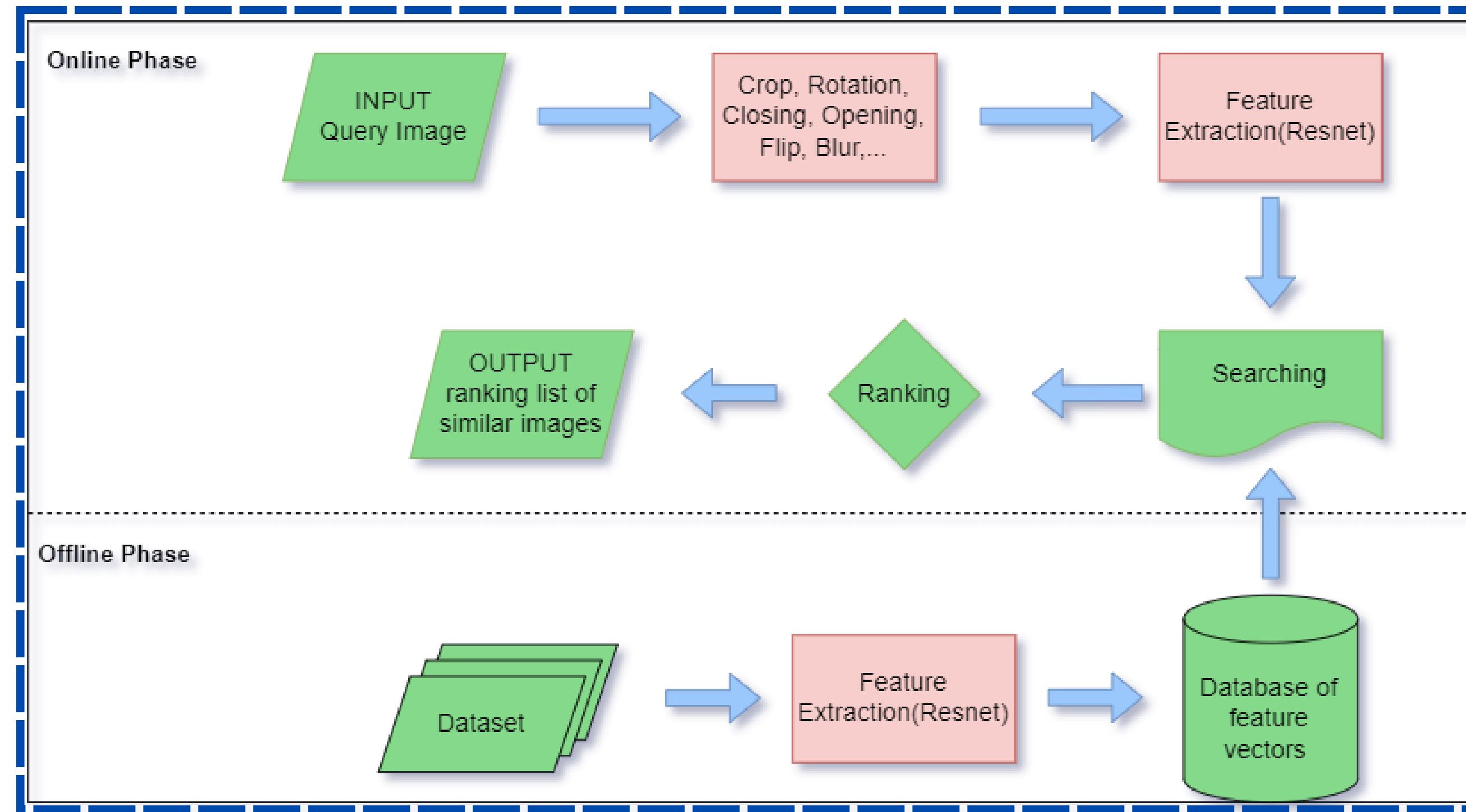
... Deeper neural networks are more difficult to train. We present a residual learning framework to ease the training of networks that are substantially deeper than those used previously.

☆ Save  Cited by 190973 Related articles All 76 versions 

Biến đổi ảnh truy vấn + đặc trưng Resnet

- Sử dụng các phép biến đổi để biến đổi ảnh cần truy vấn.
- Sử dụng pretrained Resnet50 trên tập dữ liệu Imagenet dùng để trích xuất đặc trưng.

Biến đổi ảnh truy vấn với đặc trưng Resnet

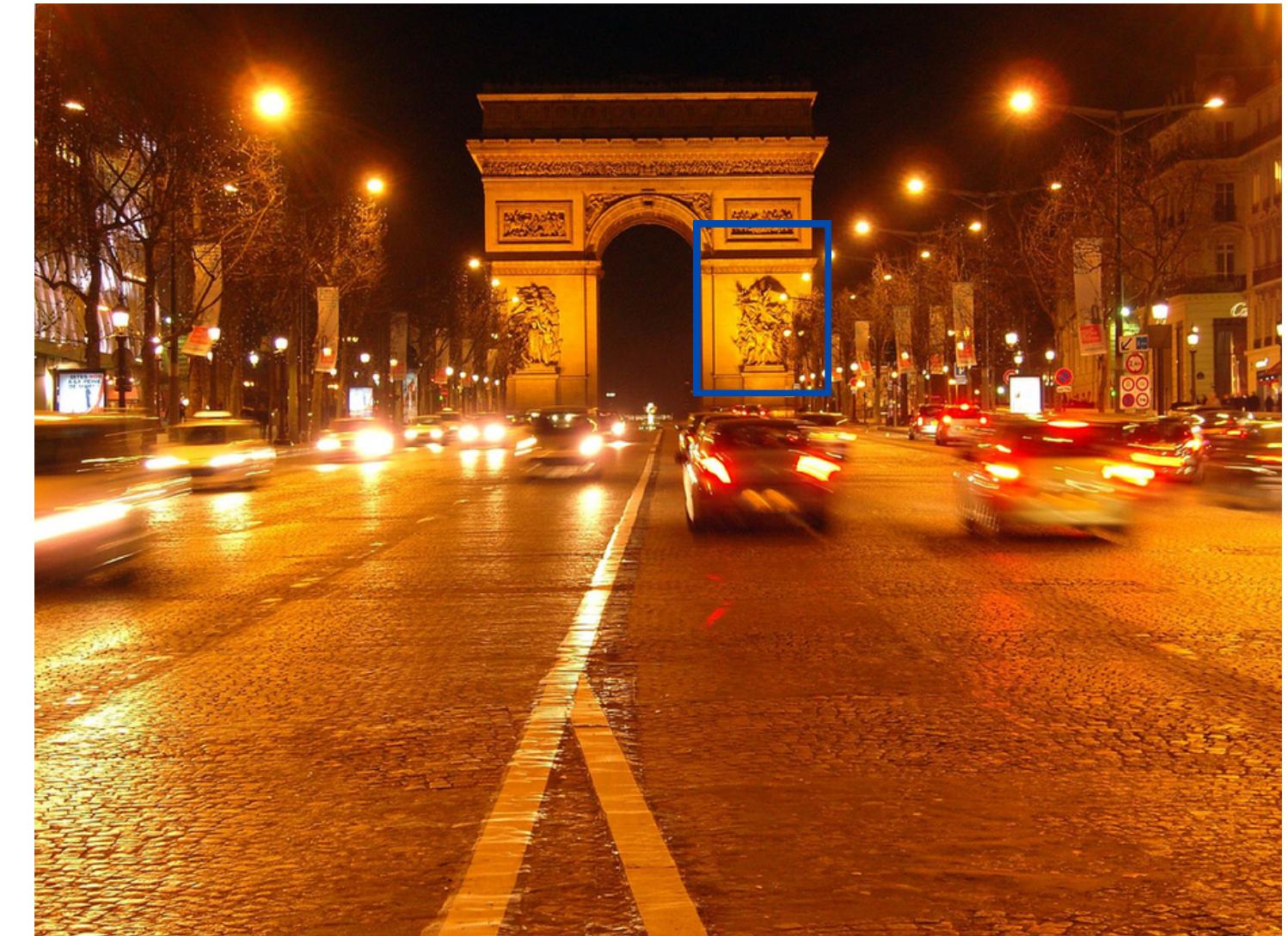
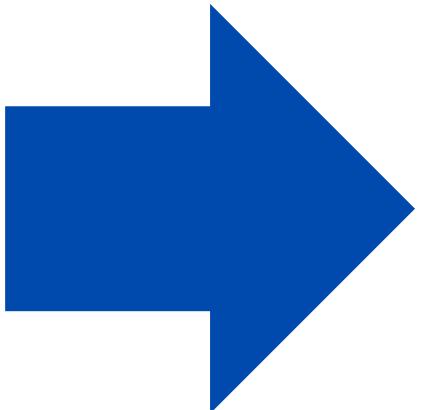


Kết quả

	ORIGINAL	CROP	RANDOM CROP	RANDOM ROTATE	HORIZONTAL FLIP	BLUR	SHARPEN	SMOOTH
mAP@5	0.9519	0.9536	0.9627	0.9627	0.9627	0.9627	0.9627	0.9627
mAP@20	0.9213	0.9294	0.9300	0.9301	0.9301	0.9301	0.9301	0.9301
mAP@50	0.8914	0.9056	0.9087	0.9086	0.9044	0.9044	0.9044	0.9044

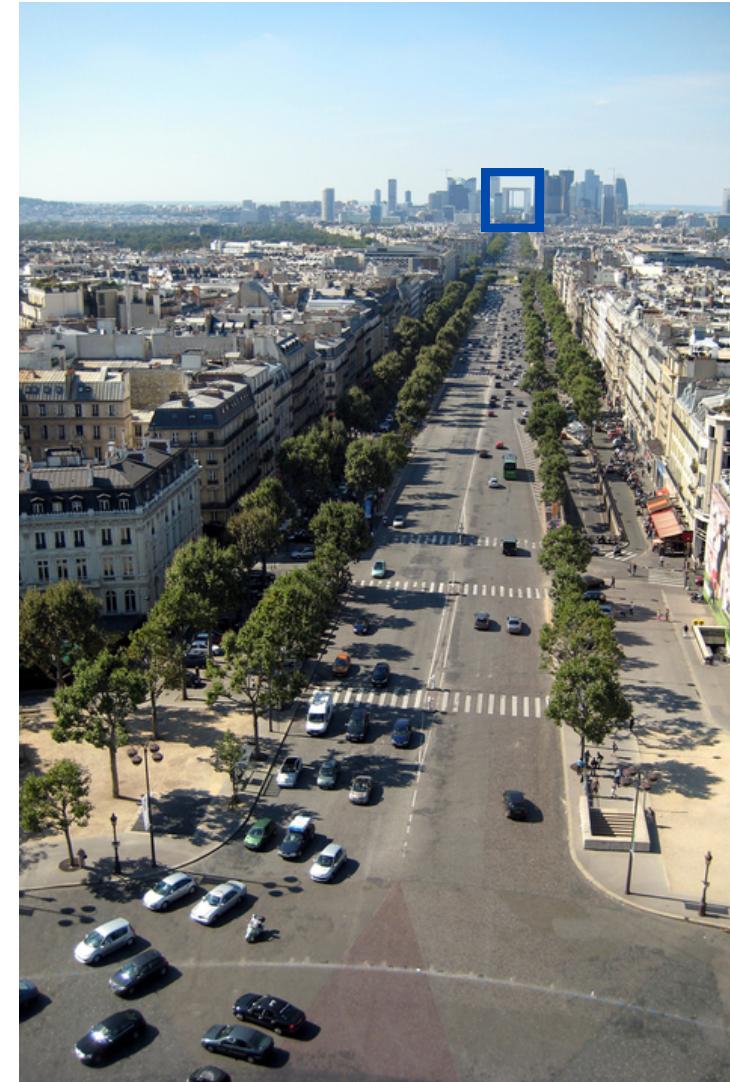
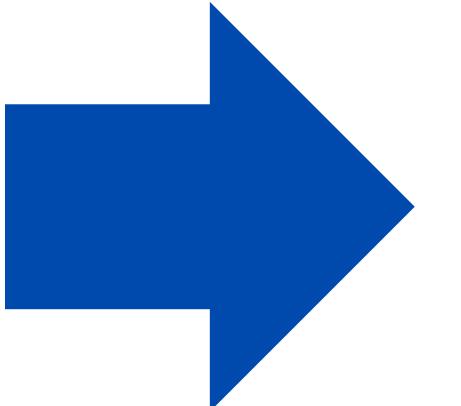
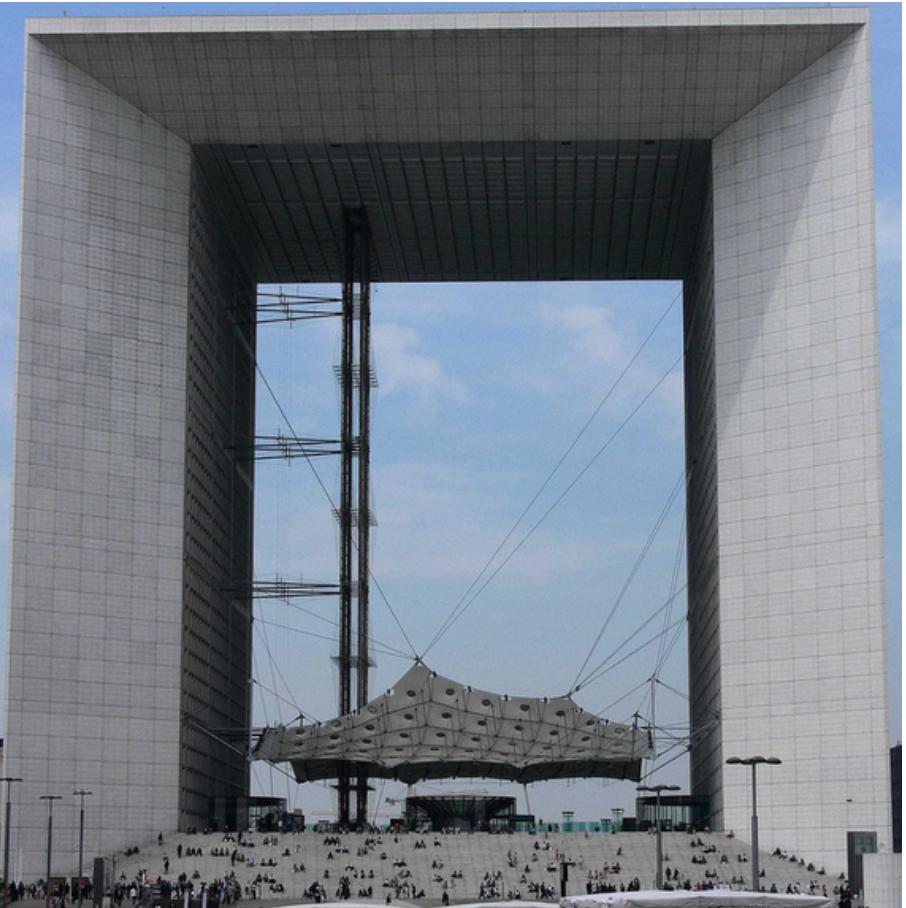
Một số kết quả chưa tốt

Ảnh mục tiêu bị nhiễu bởi yếu tố môi trường



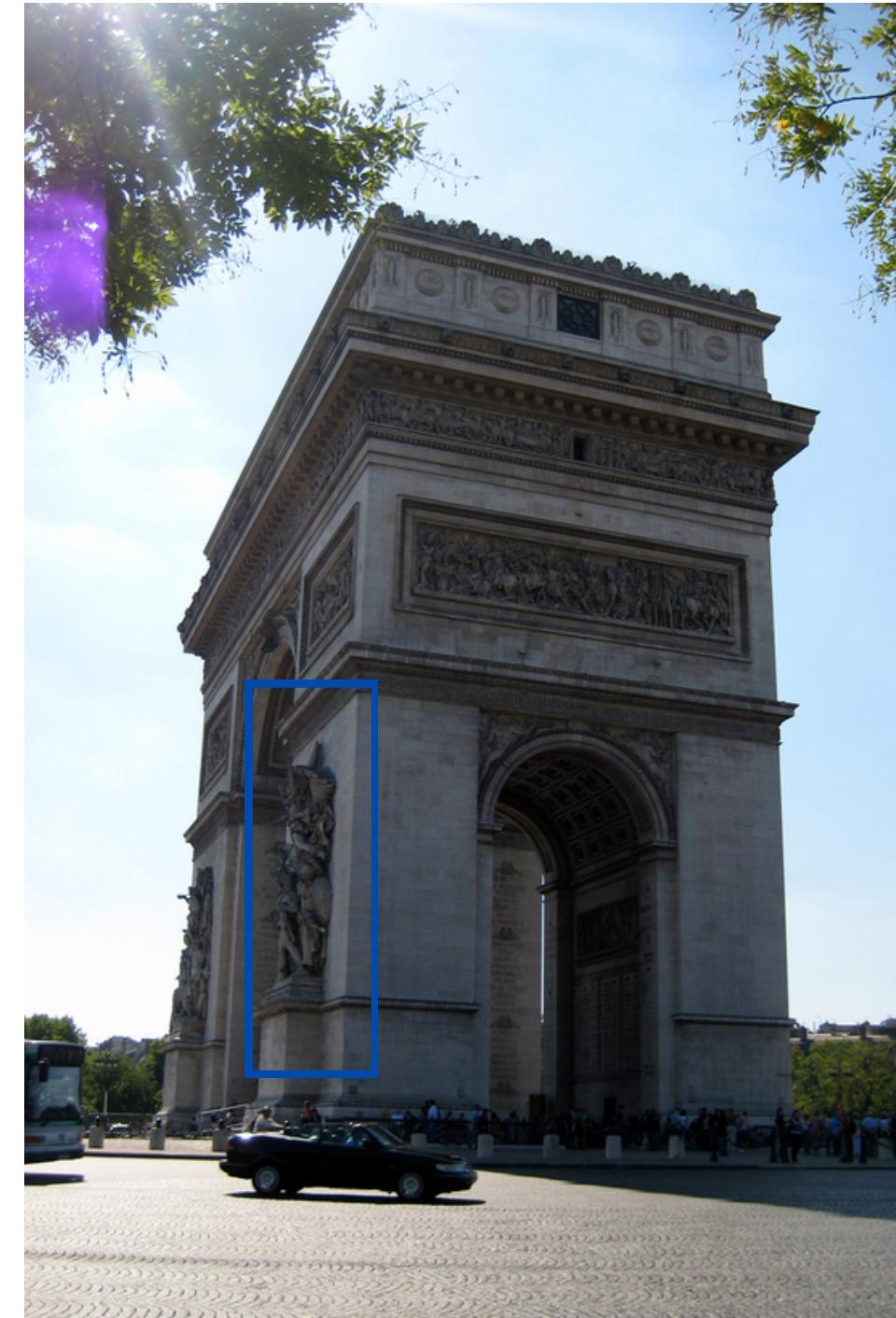
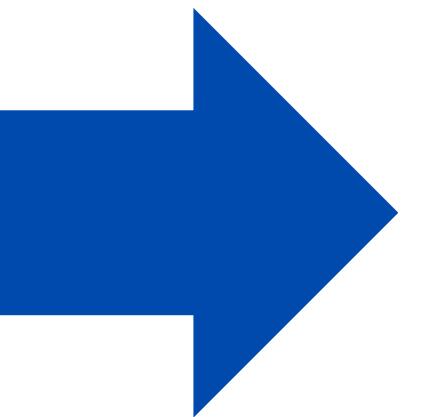
Một số kết quả chưa tốt

Ảnh mục tiêu bị nhiễu bởi yếu tố môi trường



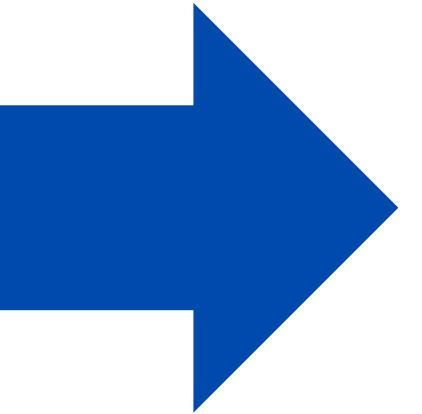
Một số kết quả chưa tốt

Biến đổi góc nhìn



Một số kết quả chưa tốt

Vùng truy vấn bị che khuất



Mục tiêu

Tìm được thêm những đáp án đúng trong trường hợp ảnh
truy vấn là một **vùng nhỏ** trên ảnh mục tiêu.

Phương pháp thực hiện

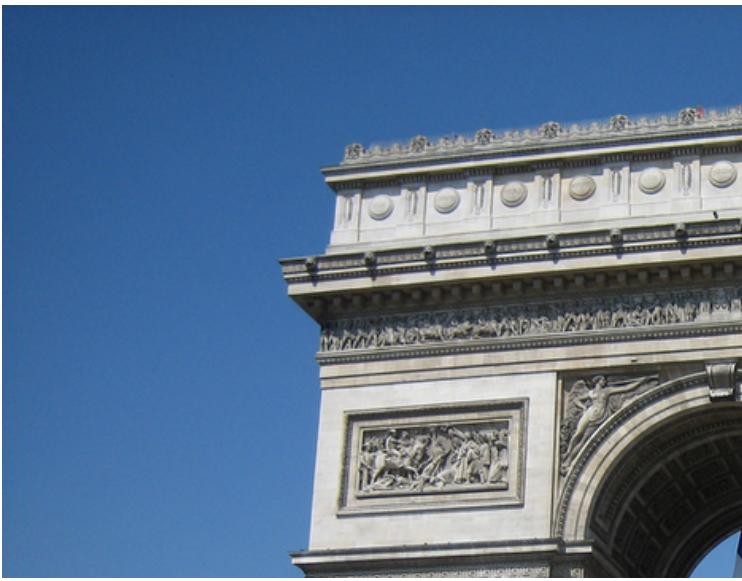
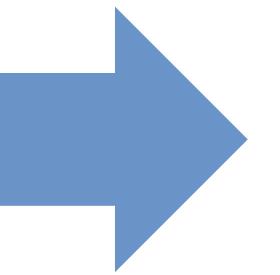
- Giảm độ nhiễu của môi trường của ảnh mục tiêu đối với vùng truy vấn.
- Tập trung vào một vùng ảnh **quan trọng**, loại bỏ thông tin không cần thiết.

Phương pháp thực hiện

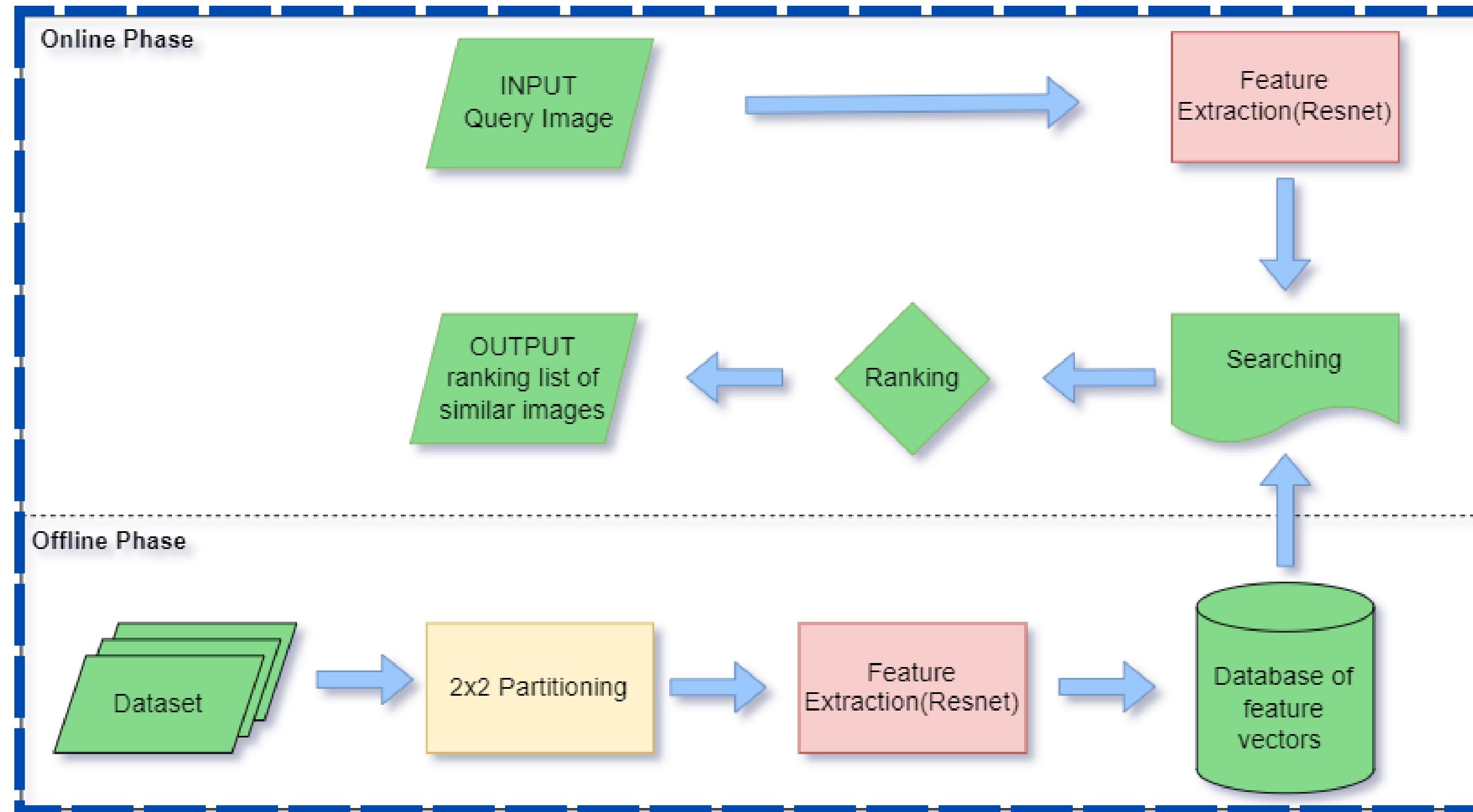
Tăng cường dữ liệu:

- Chia ảnh lớn thành các phần nhỏ hơn.
- Tìm kiếm trên các đặc trưng của các phần được phân chia.

2x2 Partitioning



Tăng cường dữ liệu



Nhận xét

Ưu điểm

- Cho kết quả tốt nếu vùng truy vấn nằm gọn trong một phần phân chia.

Nhược điểm

- Kết quả sẽ không cải thiện khi vùng truy vấn bị chia tách.
- Mất thêm không gian lưu trữ.

Kết quả

	ORIGINAL DATABASE	AUGMENTATED DATABASE
mAP@5	0.9543	0.9613
mAP@20	0.9297	0.9328
mAP@50	0.9058	0.9062
mean Precision@50	0.8274	0.8346
Recall@Overall	0.4855	0.5232
Recall@Hard	0.2954	0.3206
MRR@Overall	0.9659	0.9659

Demo

Tài liệu tham khảo

Tài liệu tham khảo

1. <https://arxiv.org/pdf/1803.11285.pdf>
2. <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9239268>
3. <https://www.robots.ox.ac.uk/~vgg/data/parisbuildings/>
4. <https://arxiv.org/pdf/1512.03385.pdf>

Thank You For Listening