

ĐẠI HỌC QUỐC GIA TP.HCM
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ THÔNG TIN



BÁO CÁO ĐỒ ÁN
NHẬN DẠNG THỊ GIÁC VÀ ỨNG DỤNG

GVHD: TS. Lê Đình Duy,
TS. Nguyễn Tân Trần Minh Khang

HVTH: CH1601015 – Lý Trọng Nhân

TP. Hồ Chí Minh, tháng 12 năm 2017

THÔNG TIN BÁO CÁO ĐỒ ÁN

- a. Địa chỉ github lưu source code và các báo cáo: https://github.com/vra-nhanlt/VRA-Nhanlt_FinalEssay
- b. Địa chỉ YouTube link đến video minh họa:
 - Link chính: <https://youtu.be/WLWmzAbnz-I>
 - Link backup: <https://youtu.be/j7-L6BQH-MM>
- c. Nội dung chi tiết báo cáo: đính kèm trong phần bên dưới của báo cáo này.

MỤC LỤC

CHƯƠNG 1. GIỚI THIỆU NỘI DUNG ĐOÁN	5
I. Mục tiêu	5
II. Tập dữ liệu (dataset)	6
2.1. Tập dữ liệu các tòa nhà tại ĐHQG-HCM	6
2.2. Tập dữ liệu các loài hoa phổ biến tại Anh	8
III. Thông tin kỹ thuật.....	8
CHƯƠNG 2. QUY TRÌNH XÂY DỰNG HỆ THỐNG TÌM KIẾM ẢNH	9
I. Tiền xử lý.....	10
1.1. Chọn lọc ảnh	10
1.2. Phân chia dữ liệu testing và training.....	11
II. Phát hiện và trích chọn đặc trưng ảnh	11
III. Gom cụm đặc trưng	18
IV. Xây dựng inverted index.....	19
4.1. Xây dựng inverted index cho dữ liệu training	19
4.2. Hỗ trợ bổ sung ảnh mới vào kho dữ liệu	19
V. Truy vấn ảnh	19
VI. Đánh giá hệ thống	21
CHƯƠNG 3. GIỚI THIỆU ỨNG DỤNG TRUY VẤN ẢNH.....	24
I. Giới thiệu giao diện ứng dụng	24
1.1. Màn hình chính	24
1.2. Truy vấn ảnh	24
1.3. Giao diện chi tiết ảnh trả về	27
1.4. Đánh giá kết quả	28
1.5. Thêm ảnh mới vào kho dữ liệu truy vấn.....	29
II. Demo ứng dụng	30

CHƯƠNG 4. KẾT QUẢ THỬ NGHIỆM	31
I. Kết quả thử nghiệm với bộ dữ liệu các tòa nhà tại ĐHQG-HCM	31
II. Kết quả thử nghiệm với bộ dữ liệu các loài hoa phổ biến tại Anh	34
CHƯƠNG 5. KẾT LUẬN	39
I. Kết quả đạt được	39
II. Những hạn chế	40
TÀI LIỆU THAM KHẢO	41

CHƯƠNG 1. GIỚI THIỆU NỘI DUNG ĐỒ ÁN

I. Mục tiêu

Đồ án này thuộc **đồ án dạng 4**:

- **Bài toán:** tìm kiếm ảnh (image retrieval).
- **Yêu cầu:** học viên tìm hiểu các phương pháp dùng BoW. Ví dụ, tham khảo matlab: <https://www.mathworks.com/help/vision/ug/image-retrieval-with-bag-of-visual-words.html>

Cụ thể, đồ án này hướng tới mục tiêu xây dựng một hệ thống tìm kiếm ảnh. Đôi tượng được hỗ trợ tìm kiếm bao gồm 2 nhóm đối tượng¹:

- Các tòa nhà tại các đơn vị, trường thành viên của Đại học Quốc gia thành phố Hồ Chí Minh. Bao gồm 6 trường thành viên: Đại học Bách Khoa (University of Technology - UT), Đại học Khoa học Tự nhiên (University of Science), Đại học Khoa học xã hội và Nhân văn (University of Social Sciences and Humanities - USSH), Đại học Kinh tế - Luật (University of Economics and Law - UEL), Đại học Quốc Tế (International University - IU), Đại học Công nghệ thông tin (University of Information Technology - UIT); Trung tâm giáo dục Quốc phòng và An ninh sinh viên (Center for National Defense Education); Ký túc xá (Hostel); Nhà điều hành và Thư viện trung tâm Đại học Quốc gia (Vietnam National University – Head quarters – VNU_HQT).
- Các hình ảnh về các loại hoa phổ biến tại vương quốc Anh, bao gồm 102 loại hoa khác nhau².

Các yêu cầu cụ thể của hệ thống như sau:

- Hệ thống cung cấp giao diện cho phép người dùng lựa chọn ảnh truy vấn. Bên cạnh việc sử dụng trực tiếp nội dung toàn bộ hình ảnh được chọn, hệ thống còn cho phép người dùng chọn một vùng ảnh để thực hiện truy vấn.
- Hệ thống hỗ trợ thêm ảnh mới vào kho dữ liệu ảnh sau khi đã xây dựng hệ thống mà không cần thực hiện lại quá trình trích chọn đặc trưng và xây dựng Bag of visual words.

¹ Trong đề tài này, tác giả hướng tới mục tiêu xây dựng hệ thống tìm kiếm ảnh các tòa nhà tại ĐHQG-HCM. Tuy nhiên, do số lượng ảnh thu thập được chưa lớn, chỉ khoảng hơn 700 ảnh, do đó, bên cạnh việc xây dựng hệ thống truy vấn ảnh liên quan đến các tòa nhà, tác giả xây dựng thêm hệ thống truy vấn ảnh các loài hoa lấy dataset có số lượng ảnh lớn (hơn 8000 ảnh) để thử nghiệm các features detectors và descriptors khác nhau và để chứng minh khả năng hoạt động của hệ thống với số lượng ảnh trong dataset lớn.

² <http://www.robots.ox.ac.uk/~vgg/data/flowers/102/index.html>

- Sau khi tìm kiếm, *các ảnh phù hợp* (nội dung giống/gần giống với ảnh ban đầu) *được hiển thị trong 3 trang kết quả*, mỗi trang hiển thị 9 bức ảnh. Mặc dù, hệ thống có thể trả về nhiều kết quả hơn nhưng thông thường người dùng chỉ sử dụng kết quả trong 3 trang đầu tiên nên trong đồ án này, tác giả giới hạn trả về 3 trang kết quả phù hợp nhất. Bên cạnh danh sách các hình ảnh phù hợp, hệ thống *dự đoán category phù hợp cho ảnh truy vấn* dựa vào danh sách kết quả trả về.
- Khi người dùng nhấn chọn vào một ảnh trong kết quả tìm kiếm, ứng dụng *hiển thị nội dung chi tiết của hình ảnh* gồm: tên, đường dẫn, score (chỉ số đo mức độ phù hợp so với ảnh query) và minh họa các điểm ảnh tương thích giữa ảnh query và ảnh trả về.
- *Hệ thống được đánh giá bằng* một hoặc một số chỉ số đánh giá hệ thống truy vấn như: Recall, Precision, Precision@k, R-Precision, Parameterized F Measure, MAP...

II. Tập dữ liệu (dataset)

Đồ án này sử dụng 2 tập dữ liệu: (1) tập dữ liệu các tòa nhà tại ĐHQG-HCM; (2) tập dữ liệu các loài hoa phổ biến tại Anh.

2.1. Tập dữ liệu các tòa nhà tại ĐHQG-HCM

Tập dữ liệu gồm 669 bức ảnh về các tòa nhà tại các đơn vị, trường thành viên thuộc ĐHQG-HCM. Các bức ảnh chủ yếu được thu thập từ nguồn Internet bao gồm:

- Tìm kiếm từ công cụ tìm kiếm Google.
- Các website của các trung tâm, trường thành viên ĐHQG:

Đơn vị/trường	Website	Đơn vị/trường	Website
VNU_HQT	http://vnuhcm.edu.vn/	US	https://www.hcmus.edu.vn/
Hostel	http://ktx.vnuhcm.edu.vn/	USSH	http://hcmussh.edu.vn/
CFNDE	http://www.ttgdpq.edu.vn/home/	IU	https://www.hcmiu.edu.vn/
UT	http://www.hcmut.edu.vn/	UEL	http://www.uel.edu.vn/
UIT	https://www.uit.edu.vn/		

- Fanpage, Facebook Group của các đơn vị, trường:

Đơn vị/trường	Fanpage, Facebook Group	Đơn vị/trường	Fanpage, Facebook Group
VNU_HQT	https://www.facebook.com/vnuhcm.edu/	US	https://www.facebook.com/DhKhtnTpHcm/
Hostel	<ul style="list-style-type: none"> - https://www.facebook.com/KTXDHQG/ - https://www.facebook.com/KTX-%C4%90HQG-TpHCM-Khu-B/199321570142660 - https://www.facebook.com/pages/KTX-%C4%90HQG-TpHCM-Khu-B/199321570142660 	USSH	https://www.facebook.com/uss.h.vnuhcm/
CFNDE	https://www.facebook.com/trungtamgiaoducquocphongdaihocquocgia/	IU	<ul style="list-style-type: none"> - https://www.facebook.com/pages/International-University/ - https://www.facebook.com/sob.iuvnu/
UT	https://www.facebook.com/bkquocte/	UEL	<ul style="list-style-type: none"> - https://www.facebook.com/pages/University-of-Economics-and-Law/158180297565982 - https://www.facebook.com/UniversityofEconomicsandLaw/ - https://www.facebook.com/pages/University-of-Economics-and-Law/158180297565982
UIT	https://www.facebook.com/UIT.Fanpage/		

Bộ dữ liệu có một số hạn chế như: số lượng ảnh nhỏ; một số ảnh có chất lượng thấp, kích thước ảnh nhỏ; các tòa nhà chưa có nhiều ảnh từ nhiều góc chụp khác nhau; ảnh được tải về từ mạng internet nên nhiều ảnh có nội dung trùng nhau, được tạo ra từ cùng một ảnh gốc với một số thao tác như chỉnh sửa kích thước, cắt ghép hoặc thay đổi màu sắc, độ sáng. Những hạn chế này có ảnh hưởng đến tính khách quan, chính xác khi đánh giá kết quả của hệ thống truy vấn ảnh.

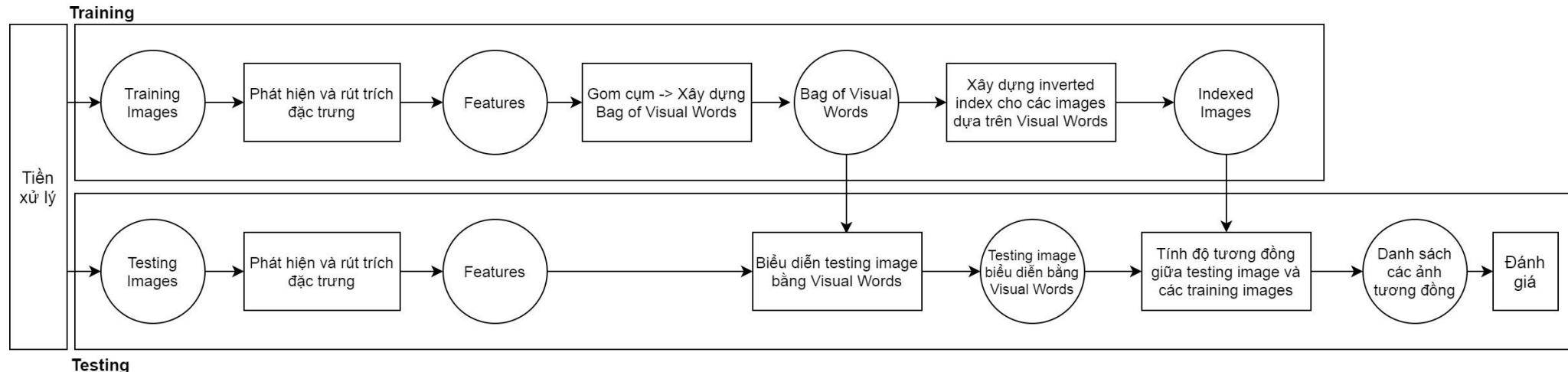
2.2. Tập dữ liệu các loài hoa phổ biến tại Anh

Tập dữ liệu gồm 8189 bức ảnh được gán nhãn tương ứng 102 loài hoa phổ biến tại vương quốc Anh. Mỗi loài hoa trong tập dữ liệu có ít nhất 40 ảnh và tối đa 258 ảnh. Bộ dữ liệu ảnh có sự đa dạng về kích thước, màu sắc, hình dáng các loài hoa và kích thước các bức ảnh. Thông tin chi tiết, đường dẫn tải tập dữ liệu, tập tin lưu danh sách gán nhãn được công bố tại: <http://www.robots.ox.ac.uk/~vgg/data/flowers/102/index.html>.

III. Thông tin kỹ thuật

- Giao diện của hệ thống tìm kiếm ảnh được xây dựng trên Matlab UI. Mặc dù, các UI controls và components do Matlab cung cấp có nhiều hạn chế so với việc xây dựng ứng dụng web, tuy nhiên các UI controls và components này vẫn đủ để đáp ứng việc xây dựng ứng dụng đáp ứng các yêu cầu của đồ án đã đề ra như: chọn ảnh, quét chọn vùng ảnh để tìm kiếm, hiển thị thông tin chi tiết ảnh kết quả và các điểm ảnh tương thích.
- Phiên bản Matlab được sử dụng: **R2017a**.
- Thư mục chứa ảnh và các tập tin lưu trữ Bag of Words và Images Index khi chạy hệ thống: E:\Google Drive\Study\Nam 2016\Master\CD_NDTGVUD\Essay\data

CHƯƠNG 2. QUY TRÌNH XÂY DỰNG HỆ THỐNG TÌM KIẾM ẢNH



Quy trình xây dựng hệ thống tìm kiếm ảnh bao gồm 2 pha: training và testing. Trong đó, bước tiền xử lý bộ ảnh nguồn cần được thực hiện trước quy trình của 2 pha này. Nhiệm vụ cụ thể của các bước như sau:

- **Tiền xử lý:** chọn lọc, chuyển định dạng ảnh, phân chia bộ ảnh testing và bộ ảnh training.
- **Phát hiện và trích chọn đặc trưng:** với mỗi ảnh cần thực hiện việc trích chọn đặc trưng ảnh sử dụng các Features Detectors (phương pháp phát hiện interest points) và Features descriptors (phương pháp rút trích và biểu diễn đặc trưng ảnh).
- **Gom cụm:** gom cụm features từ bộ ảnh training sử dụng K-means (hoặc thuật toán gom cụm khác) để xây dựng Bag of visual words.
- **Xây dựng inverted index:** đánh chỉ mục cho các bức ảnh trong bộ ảnh training dựa trên các visual words vừa được xây dựng.
- **Biểu diễn testing image bằng visual words:** sau khi rút trích các đặc trưng trên ảnh test, ảnh này sẽ được biểu diễn bằng các visual words trong Bag of visual words của bộ ảnh training.

- **Tính độ tương đồng:** tính độ tương đồng giữa ảnh test và các ảnh trong bộ ảnh training để tìm ra danh sách các ảnh tương đồng xếp theo thứ tự độ tương đồng từ cao đến thấp.
- **Đánh giá:** đánh giá hệ thống tìm kiếm dựa trên các chỉ số đánh giá như Recall, Precision, Precision@k, R-Precision, Parameterized F Measure, MAP...

I. Tiền xử lý

Bộ dữ liệu ảnh các tòa nhà tại ĐHQG-HCM được xây dựng chủ yếu từ các hình ảnh được tải về từ internet nên bộ dữ liệu này tồn tại nhiều điểm hạn chế đã đề cập ở mục 2.1 trong chương 1. Các thao tác tiền xử lý cần được thực hiện trên bộ ảnh này gồm: chọn lọc ảnh; chuyển định dạng màu ảnh (về RGB) và phân chia dữ liệu testing và training.

1.1. Chọn lọc ảnh

Trong đồ án này, các ảnh chất lượng kém được hiểu là các ảnh không chứa nội dung liên quan đến các tòa nhà nên không mang nhiều ý nghĩa trong việc tìm kiếm ảnh về các tòa nhà và các ảnh có kích thước quá nhỏ. Ví dụ: các ảnh chụp cá nhân mà các bộ phận cơ thể người chiếm phần lớn khung hình; các bức ảnh chụp nền cỏ... Các bức ảnh này nên được loại đi trước khi thực hiện các bước phát hiện và rút trích đặc trưng khi xây dựng hệ thống.

Hình 1. Minh họa các bức ảnh chất lượng thấp



Bộ dữ liệu ban đầu thu thập được gồm hơn 700 bức ảnh. Sau khi loại bỏ các ảnh chất lượng thấp, bộ dữ liệu ảnh các tòa nhà tại ĐHQG-HCM có tổng số: **669** ảnh.

Đối với bộ dữ liệu ảnh các loài hoa phổ biến tại Anh thì do đây là bộ dữ liệu đã được xử lý và được công bố phục vụ mục đích nghiên cứu, do đó với các yêu cầu của đồ án này thì bộ dữ liệu đã rất phù hợp và không cần chọn lọc ảnh nữa.

1.2. Phân chia dữ liệu testing và training

Để việc đánh giá hệ thống đảm bảo tính khách quan, tránh lỗi testing bằng bộ dữ liệu training, mỗi bộ dữ liệu ban đầu cần được phân chia thành 2 nhóm: dữ liệu testing và dữ liệu training. Trong đồ án này, tác giả chọn tỉ lệ 2:8, tức 20% ảnh dùng để testing và 80% ảnh dùng để training.

Bộ dữ liệu	Số lượng ảnh Training	Số lượng ảnh Testing
Các tòa nhà ĐHQG-HCM	534	135
Các loài hoa phổ biến tại Anh	6582	1607

II. Phát hiện và trích chọn đặc trưng ảnh

Các tòa nhà, công trình có đặc điểm đặc trưng là hình dáng, các chi tiết góc, cạnh, hình khối. Việc khai thác các đặc điểm trên trong việc nhận dạng đối tượng tòa nhà, công trình sẽ giúp tăng hiệu quả cho hệ thống truy vấn. Trong khi đó, các loài hoa có đặc điểm nổi bật về thị giác là màu sắc và hình dáng.

Để thực hiện quá trình rút trích đặc trưng ảnh, ta cần xem xét, thử nghiệm 2 phương pháp quan trọng gồm: phương pháp phát hiện đặc trưng (features detectors) và phương pháp mô tả đặc trưng (features descriptors). Hai phương pháp này có ảnh hưởng trực tiếp đến kết quả của hệ thống truy vấn. Trong đồ án này, tác giả thử nghiệm 5 phương pháp phát hiện đặc trưng là SURF (Speeded Up Robust Features), MinEigen, FAST, GRID, Harris; và 2 phương pháp mô tả đặc trưng gồm **SURF** và **HoG** (Histogram of Oriented Gradients). Ngoài ra, với bộ dữ liệu các loài hoa, tác giả thử nghiệm thêm phương pháp trích chọn **đặc trưng màu sắc** để khai thác đặc điểm màu sắc của các loài hoa. Bên cạnh đó, tham số số lượng VocabularySize cũng có ảnh hưởng lớn tới kết quả của hệ thống truy vấn ảnh sử dụng phương pháp Bag of Visual Words. Do đó, đồ án này cũng đã thử nghiệm lần lượt 2 giá trị khác nhau cho tham số VocabularySize lần lượt là 500 (giá trị mặc định) và 1000.

Phương pháp mô tả đặc trưng SURF là phương pháp mô tả đặc trưng cục bộ dựa trên hướng tiếp cận tương tự với phương pháp trích chọn đặc trưng SIFT (Scale-Invariant Feature Transform). SURF sử dụng không gian scale để tìm điểm đặc trưng, các đặc trưng được mô tả dưới dạng vector và có kèm thêm hướng [1]. Phương pháp trích chọn đặc trưng SURF có tốc độ nhanh, phù hợp với các mục đích nhận dạng đối tượng, nhận dạng mặt người cũng như các tòa nhà, công trình. Mặt khác, phương pháp này dễ hiện thực khi xây dựng chương trình do Matlab đã xây dựng sẵn.

HOG (Histogram of Gradient) là một features descriptor được dùng phổ biến trong việc phát hiện đối tượng. Phương pháp này được đề xuất vào năm 2005 tại viện nghiên cứu INRIA bởi 2 nhà nghiên cứu Bill Triggs và Navel Dalal. HoG dùng đặc trưng sự phân bố về gradient (sự biến thiên về màu sắc/mức xám trong một vùng ảnh) và hướng của cạnh để biểu diễn hình dạng và trạng thái của vật thể. Đặc trưng này được phát triển dựa trên ý tưởng của SIFT, mỗi đặc trưng được tính dựa trên một vùng ảnh. Sự biến thiên về màu sắc/mức xám trong mỗi vùng ảnh sẽ có sự khác biệt với những vùng khác nên nó có khả năng thể hiện tính đặc trưng cho mỗi vùng ảnh thông qua một vector đặc trưng. Bên cạnh đó, để có được đặc trưng của toàn bộ cửa sổ (Window), ta phải kết hợp nhiều vùng lận cận lại với nhau [2].

Trong đồ án này, tác giả đã thử nghiệm kết hợp các phương pháp phát hiện đặc trưng với các phương pháp mô tả đặc trưng khác nhau. Mỗi thử nghiệm được đánh giá với 5 query khác nhau và ghi nhận kết quả đánh giá trung bình. Năm bức ảnh được chọn để thử nghiệm đánh giá kết quả trong bộ ảnh các tòa nhà tại ĐHQG-HCM gồm: (1) cfnde (24).jpg; (2) iu (65).jpg; (3) uel (20).jpg; (4) uit (56).jpg; (5) vnu_hqr (10).jpg.



Hình 2. Các hình ảnh được dùng để thử nghiệm trong bộ dữ liệu các tòa nhà tại ĐHQG-HCM

Chỉ số Precision@k được dùng để đánh giá các thử nghiệm. Việc lựa chọn các features detectors kết hợp với các features descriptors khác nhau để sử dụng cho hệ thống truy vấn được dựa vào kết quả trung bình của 5 ảnh query. Kết quả thử nghiệm trên bộ dữ liệu các tòa nhà tại ĐHQG-HCM được ghi nhận trong **Bảng 1**:

Bảng 1. Kết quả thử nghiệm kết hợp các features detectors với features descriptors trên bộ dữ liệu các tòa nhà tại ĐHQG-HCM

STT	Detector	Descriptor	VocabularySize = 1000					
			Hình #1	Hình #2	Hình #3	Hình #4	Hình #5	Trung bình
1	SURF	SURF	11,11	77,78	33,33	37,04	44,44	40,74
2	Color		x	x	x	x	x	x
3	MinEigen	SURF	7,41	55,56	22,22	33,33	44,44	32,592
4	FAST	SURF	11,11	59,26	22,22	25,93	25,93	28,89
5	GRID	SURF	0	25,93	14,81	14,81	22,22	15,554
6	Harris	SURF	11,11	40,74	22,22	37,04	37,04	29,63
7	SURF	HOG	11,11	77,78	22,22	37,04	33,33	36,296
7	FAST	HOG	3,7	62,96	29,63	44,44	40,74	36,294

STT	Detector	Descriptor	VocabularySize = 1000					
			Hình #1	Hình #2	Hình #3	Hình #4	Hình #5	Trung bình
1	SURF	SURF	11,11	77,78	33,33	37,04	44,44	40,74
2	Color		x	x	x	x	x	x
3	MinEigen	SURF	7,41	55,56	22,22	33,33	44,44	32,592
4	FAST	SURF	11,11	59,26	22,22	25,93	25,93	28,89
5	GRID	SURF	0	25,93	14,81	14,81	22,22	15,554
6	Harris	SURF	11,11	40,74	22,22	37,04	37,04	29,63
7	SURF	HOG	11,11	77,78	22,22	37,04	33,33	36,296
8	FAST	HOG	3,7	62,96	29,63	44,44	40,74	36,294

Tương tự, đối với bộ dữ liệu các loài hoa phổ biến tại Vương quốc Anh, do kích thước bộ dữ liệu quá lớn, để tiết kiệm thời gian thử nghiệm, tác giả chỉ chọn 30 loại hoa đầu tiên trong bộ dữ liệu, mỗi loài hoa chọn 30 bức ảnh ngẫu nhiên để thử nghiệm các phương pháp phát hiện và mô tả đặc trưng. Sau khi chọn ra các phương pháp phù hợp nhất, tác giả mới sử dụng bộ dữ liệu đầy đủ để trích chọn đặc trưng và xây dựng hệ thống hoàn chỉnh.

Năm hình ảnh được chọn để thử nghiệm lần lượt là: (1) label_1_8.jpg; (2) label_2_7.jpg; (3) label_5_4.jpg; (4) label_9_4.jpg; (5) label_17_4.jpg.

Hình 3. Các hình ảnh được dùng để thử nghiệm trong bộ dữ liệu các loài hoa phổ biến tại Vương quốc Anh



Kết quả thử nghiệm trên bộ dữ liệu các loài hoa phổ biến tại Vương quốc Anh được ghi nhận trong **Bảng 2**:

Bảng 2. Kết quả thử nghiệm kết hợp các features detectors với features descriptors trên bộ dữ liệu các loài hoa phổ biến tại Vương quốc Anh

STT	Detector	Descriptor	VocabularySize = 500					
			Hình #1	Hình #2	Hình #3	Hình #4	Hình #5	Trung bình
1	SURF	SURF	0	40,74	14,81	0	7,41	12,592
2	Color		44,44	40,74	51,85	37,04	3,7	35,554
3	MinEigen	SURF	29,63	59,26	3,7	33,33	7,41	26,666

4	FAST	SURF	7,41	55,56	40,74	0	7,41	22,224
5	GRID	SURF	33,33	51,85	3,7	11,11	0	19,998
6	Harris	SURF	14,81	66,67	18,52	7,41	7,41	22,964
7	SURF	HOG	3,7	18,52	0	0	7,41	5,926
8	FAST	HOG	14,81	44,44	18,52	0	0	15,554

STT	Detector	Descriptor	VocabularySize = 1000					
			Hình #1	Hình #2	Hình #3	Hình #4	Hình #5	Trung bình
1	SURF	SURF	14,81	55,56	0	3,7	18,52	18,518
2	Color		51,85	48,15	59,26	37,04	0	39,26
3	MinEigen	SURF	18,52	62,96	7,41	29,63	14,81	26,666
4	FAST	SURF	14,81	59,26	37,04	3,7	7,41	24,444
5	GRID	SURF	25,93	51,85	11,11	3,7	3,7	19,258
6	Harris	SURF	14,81	66,67	14,81	3,7	7,41	21,48
7	SURF	HOG	11,11	40,74	0	0	3,7	11,11
8	FAST	HOG	18,52	40,74	0	0	0	11,852

Để tăng độ chính xác cho hệ thống tìm kiếm, đồ án này **sử dụng hướng tiếp cận là kết hợp kết quả tìm kiếm từ các hệ thống sử dụng các phương pháp phát hiện và mô tả đặc trưng khác nhau để tìm ra các kết quả phù hợp nhất**. Cụ thể, dựa trên kết quả các thử nghiệm trên, tác giả chọn ra 3 cặp phương pháp phát hiện và mô tả đặc trưng cho kết quả tốt nhất tương ứng với mỗi bộ dữ liệu để kết hợp xây dựng 3 hệ thống mới. Khi thực hiện truy vấn ảnh, kết quả truy vấn từ 3 hệ thống sẽ được kết hợp lại với nhau để cho ra kết quả cuối cùng.

Bảng 3. Kết quả các phương pháp được lựa chọn để xây dựng hệ thống truy vấn ảnh

Bộ dữ liệu	Features detector	Features descriptor	VocabularySize
Các tòa nhà tại ĐHQG-HCM	FAST	HoG	500
	SURF	SURF	1000
	SURF	SURF	500

Các loài hoa phổ biến tại Vương quốc Anh	Color	1000
	Color	500
	MinEigen	SURF

Trong nội dung source code, bước phát hiện và rút trích đặc trưng được thực hiện trong các hàm³:

```

function [features, metrics] = extractFeaturesFAST_HoG(img)
    % convert RGB to Gray
    if size(img, 3) == 3
        img = rgb2gray(img);
    end

    % detect interested points using Fast
    interestedPoints = detectFASTFeatures(img);
    interestedPoints = selectStrongest(interestedPoints, 1000);

    % extract features
    [features] = extractHOGFeatures(img, interestedPoints);
    metrics = var(features, [], 2);
end

function [features, metrics, location] = extractFeaturesSURF(img)
    % convert RGB to Gray
    if size(img, 3) == 3
        img = rgb2gray(img);
    end

    % detect interested points using SURF
    interestedPoints = detectSURFFeatures(img);

    % extract features
    [features] = extractFeatures(img, interestedPoints, 'Method',
'SURF', 'Upright', true);
    location = interestedPoints.Location;
    metrics = var(features, [], 2);
end

% This function has been downloaded from
https://www.mathworks.com/help/vision/examples/image-retrieval-using-customized-bag-of-features.html
% I have not written it myself

```

³ Báo cáo chỉ liệt kê các phương pháp được lựa chọn để xây dựng hệ thống. Riêng các phương pháp được dùng để thử nghiệm và đánh giá nhưng không được lựa chọn sẽ không được trình bày trong báo cáo này, chi tiết các hàm được trình bày trong mã nguồn đính kèm báo cáo.

```
function [features, metrics] = extractColorFeatures(I)
[~,~,P] = size(I);

isColorImage = P == 3;

if isColorImage

    % Convert RGB images to the L*a*b* colorspace.
    Ilab = rgb2lab(I);

    % Compute the "average" L*a*b* color within 16-by-16 pixel
    % blocks. The
    % average value is used as the color portion of the image
    % feature. An
    % efficient method to approximate this averaging procedure
    % over
    % 16-by-16 pixel blocks is to reduce the size of the image by
    % a factor
    % of 16 using IMRESIZE.
    Ilab = imresize(Ilab, 1/16);

    % Note, the average pixel value in a block can also be
    % computed using
    % standard block processing or integral images.

    % Reshape L*a*b* image into "number of features"-by-3 matrix.
    [Mr,Nr,~] = size(Ilab);
    colorFeatures = reshape(Ilab, Mr*Nr, []);

    % L2 normalize color features
    rowNorm = sqrt(sum(colorFeatures.^2,2));
    colorFeatures = bsxfun(@rdivide, colorFeatures, rowNorm +
eps);

    % Augment the color feature by appending the [x y] location
    % within the
    % image from which the color feature was extracted. This
    % technique is
    % known as spatial augmentation. Spatial augmentation
    % incorporates the
    % spatial layout of the features within an image as part of
    % the
    % extracted feature vectors. Therefore, for two images to
    % have similar
    % color features, the color and spatial distribution of color
    % must be
    % similar.

    % Normalize pixel coordinates to handle different image
    % sizes.
```

```

xnorm = linspace(-0.5, 0.5, Nr);
ynorm = linspace(-0.5, 0.5, Mr);
[x, y] = meshgrid(xnorm, ynorm);

% Concatenate the spatial locations and color features.
features = [colorFeatures y(:) x(:)];

% Use color variance as feature metric.
metrics = var(colorFeatures(:,1:3),0,2);
else
    % Return empty features for non-color images. These features
    % are
    % ignored by bagOfFeatures.
    features = zeros(0,5);
    metrics = zeros(0,1);
end

function [features, metrics, location] =
extractFeaturesMinEigenFeaturesAndSURF(img)

% convert RGB to Gray
if size(img,3)== 3
    img = rgb2gray(img);
end

% detect interested points using MinEigenFeatures
interestedPoints = detectMinEigenFeatures(img);
interestedPoints = selectStrongest(interestedPoints, 1000);

% extract features
[features] = extractFeatures(img, interestedPoints, 'Method',
'SURF', 'Upright', true);
location = interestedPoints.Location;
metrics = var(features, [], 2);
end

```

Ngoài ra, trên thực tế còn có rất nhiều phương pháp trích chọn đặc trưng khác phù hợp cho bài toán truy vấn ảnh như LBP, histogram, FREAK, BRISK... [2] Tuy nhiên, do thời gian thực hiện đồ án có hạn nên tác giả chưa thể thử nghiệm, đánh giá các phương pháp này để chọn phương pháp phù hợp nhất cho bài toán truy vấn ảnh các tòa nhà và các loài hoa.

III. Gom cụm đặc trưng

Việc gom cụm các đặc trưng nhằm xây dựng Bag of visual words cho hệ thống tìm kiếm. Thuật toán gom cụm được sử dụng trong đồ án này là thuật toán K-Means. Sau khi thử nghiệm tham số VocabularySize lần lượt là 500 (mặc định), 1000, tác giả chọn sử dụng các giá trị

tương ứng cho từng phương pháp trích chọn đặc trưng như đã trình bày ở phần II của chương này vì nó cho kết quả tốt nhất cho hệ thống tìm kiếm.

Trong nội dung source code, bước gom cụm đặc trưng, xây dựng Bag of visual words được thực hiện bằng hàm:

```
bag = bagOfFeatures(imgDataStore, 'VocabularySize',  
vocabularySize, 'CustomExtractor', @extractFeaturesSURF)
```

IV. Xây dựng inverted index

4.1. Xây dựng inverted index cho dữ liệu training

Xây dựng inverted index cho các ảnh trong bộ ảnh training dựa trên các visual words được tạo ra ở bước trước.

Trong nội dung source code, bước xây dựng inverted index cho các ảnh được thực hiện bằng hàm:

```
indexImages(imgDataStore, bag)
```

4.2. Hỗ trợ bổ sung ảnh mới vào kho dữ liệu

Để thêm một hoặc nhiều ảnh mới vào kho dữ liệu truy vấn, ta cần thực hiện theo các bước: tạo ImageDataStore từ các tập tin ảnh; đánh chỉ mục các ảnh mới và thêm vào bảng lưu trữ chỉ mục ảnh dựa trên visual words (imagesIndex). Các bước trên được thực hiện như sau:

```
imds = imageDatastore(filenames);  
addImages(imgsIndex, imds);  
save(savedImgIndex, 'imgsIndex');
```

V. Truy vấn ảnh

Trong pha testing, việc thực hiện truy vấn ảnh được thực hiện qua các bước: (1) rút trích đặc trưng ảnh (tương tự pha training); (2) bước biểu diễn testing image bằng visual words và (3) tính độ tương đồng để tìm ra danh sách ảnh tương đồng với ảnh truy vấn.

Trong nội dung source code, bước (2) và (3) được thực hiện đơn giản bằng hàm [4] [5]:

```
retrieveImages(queryImgCdata, imgsIndex, 'NumResults', numResults)
```

Trong đồ án này, với mỗi ảnh khi thực hiện truy vấn, nó sẽ được truy vấn lần lượt trên ba hệ thống sử dụng 3 phương pháp trích chọn đặc trưng khác nhau như đã trình bày ở bước trước. Kết quả trả về từ mỗi hệ thống bao gồm: ID, score (tạm hiểu là chỉ số đánh giá mức độ tương đồng với ảnh cần truy vấn) của các bức ảnh được trả về; các visual words biểu diễn bức ảnh truy vấn. Kết quả từ 3 hệ thống này sẽ được kết hợp lại với nhau dựa trên việc tính tổng score theo từng bức ảnh được tìm thấy ở 3 hệ thống. Kết quả cuối cùng được trả về cho người dùng sẽ là danh sách cách bức ảnh có tổng score từ cao đến thấp. Việc kết hợp kết quả trả về từ 3 hệ thống giúp cải thiện hiệu quả của hệ thống tìm kiếm ảnh. Tuy nhiên nó cũng đòi hỏi sự đánh đổi về thời gian tìm kiếm của hệ thống.

Trong source code, bước kết hợp kết quả từ 3 hệ thống được thực hiện bằng các hàm:

```

function combinedResult = combineResult(combinedResult,
imageIndex, foundImgIDs, scores, IRSysystemID, weight)

    % get result count
    resultCount = size(foundImgIDs, 1);
    currentResultCount = size(combinedResult, 1);

    % add new result to combinedResult
    for (i = 1:resultCount)
        path = imageIndex.ImageLocation{foundImgIDs(i)};
        combinedResult{currentResultCount + i, 1} = path;
        combinedResult{currentResultCount + i, 2} = foundImgIDs(i);
        combinedResult{currentResultCount + i, 3} = scores(i)*
weight;
        combinedResult{currentResultCount + i, 4} = scores(i);
        combinedResult{currentResultCount + i, 5} = IRSysystemID;

    end
end

function combinedResult =
getBestResultAfterCombining(combinedResult)

    % Calculate new score group by filename
    combineResultCount = size (combinedResult, 1);
    for (i = 1:(combineResultCount - 1))

```

```

for (j = (i + 1):combineResultCount)
    if (strcmp(char(combinedResult(i, 1)),
char(combinedResult(j, 1))))
        newScore = cell2mat(combinedResult(i, 3)) +
cell2mat(combinedResult(j, 3));
        combinedResult(i, 3) = num2cell(newScore);
        combinedResult(j, 3) = num2cell(0);
    end
end
% Sort by score
for (i = 1:(combineResultCount - 1))
    for (j = (i + 1):combineResultCount)
        if (cell2mat(combinedResult(i, 3)) <
cell2mat(combinedResult(j, 3)))
            tmp = combinedResult(i, :);
            combinedResult(i, :) = combinedResult(j, :);
            combinedResult(j, :) = tmp;
        end
    end
end end

```

VI. Đánh giá hệ thống

Để đánh giá hệ thống ta cần xác định danh sách các kết quả trả về phù hợp cho từng câu query thử nghiệm. Với bộ ảnh các tòa nhà tại ĐHQG-HCM, bộ ảnh testing và training được gán nhãn tương ứng thuộc 1 trong 7 category sau: cfnde (tương ứng Trung tâm giáo dục quốc phòng và An ninh sinh viên); hostel (Ký túc xá sinh viên); uit (trường ĐH CNTT; uel (trường ĐH Kinh tế - Luật; ut (trường ĐH Bách Khoa); us (trường ĐH KHTN); iu (trường ĐH Quốc tế); ussh (trường ĐH KHXH&NV); vnu_hqt (nhà điều hành và thư viện trung tâm ĐHQG-HCM). Với bộ ảnh các loài hoa phổ biến tại Vương quốc Anh, các bức ảnh được gán nhãn thuộc một trong 102 category tương ứng 102 loài hoa khác nhau.

Ngoài ra, hệ thống dự đoán category phù hợp cho một bức ảnh khi truy vấn bằng cách thống kê kết quả trả về. Một category sẽ được dự đoán cho một bức ảnh truy vấn khi danh

sách kết quả trả về chưa chứa ảnh thuộc category đó nhất. Kết quả này sẽ được dùng để đánh giá hệ thống thông qua các chỉ số đánh giá. Bên cạnh đó, nếu hệ thống dự đoán sai category, người dùng có thể tự chọn lại category phù hợp để hệ thống tính toán lại chỉ số đánh giá. Trong mã nguồn, hàm thống kê danh sách kết quả và dự đoán category được hiện thực như sau:

```

function [indexOfCategory, predictiveCategory] =
predictCategory(selectiveCategories, combinedResult)
maxResultCount = 27;
% count result group by category
numOfCategories = length(selectiveCategories);
resultCountGroupByCategory = zeros([1 numOfCategories]);
for (i=1:maxResultCount)
    for(j=1:numOfCategories)
        if (contains(combinedResult(i,1),
selectiveCategories(j)))
            resultCountGroupByCategory(j) =
resultCountGroupByCategory(j) + 1;
            break;
        end
    end
end

% find the position (index) of the category that finds the most
number of result
indexMax = 1;
for (i=1:numOfCategories)
    if
(resultCountGroupByCategory(i)>resultCountGroupByCategory(indexMa
x))
        indexMax = i;
    end
end

% return the category that finds the most number of result
predictiveCategory = selectiveCategories(indexMax);
indexOfCategory = indexMax;
end

```

Có nhiều chỉ số có thể được dùng để đánh giá một hệ thống truy vấn ảnh như Recall, Precision, Precision@k, R-Precision, Parameterized F Measure, MAP... Trong các hệ thống tìm kiếm ảnh, độ chính xác và thứ tự xuất hiện của các ảnh trong danh sách kết quả trả về là

các yếu tố được quan tâm nhiều hơn do thông thường người dùng chỉ chọn các kết quả trả về trong khoảng 3 trang kết quả đầu tiên. Do đó, các chỉ số liên quan tới độ chính xác giữ vai trò quan trọng hơn so với độ phủ (recall). Trong đồ án này, tác giả chọn chỉ số Precision@k để đánh giá hệ thống với $k = 27$ (số ảnh trả về trong 3 trang kết quả đầu tiên).

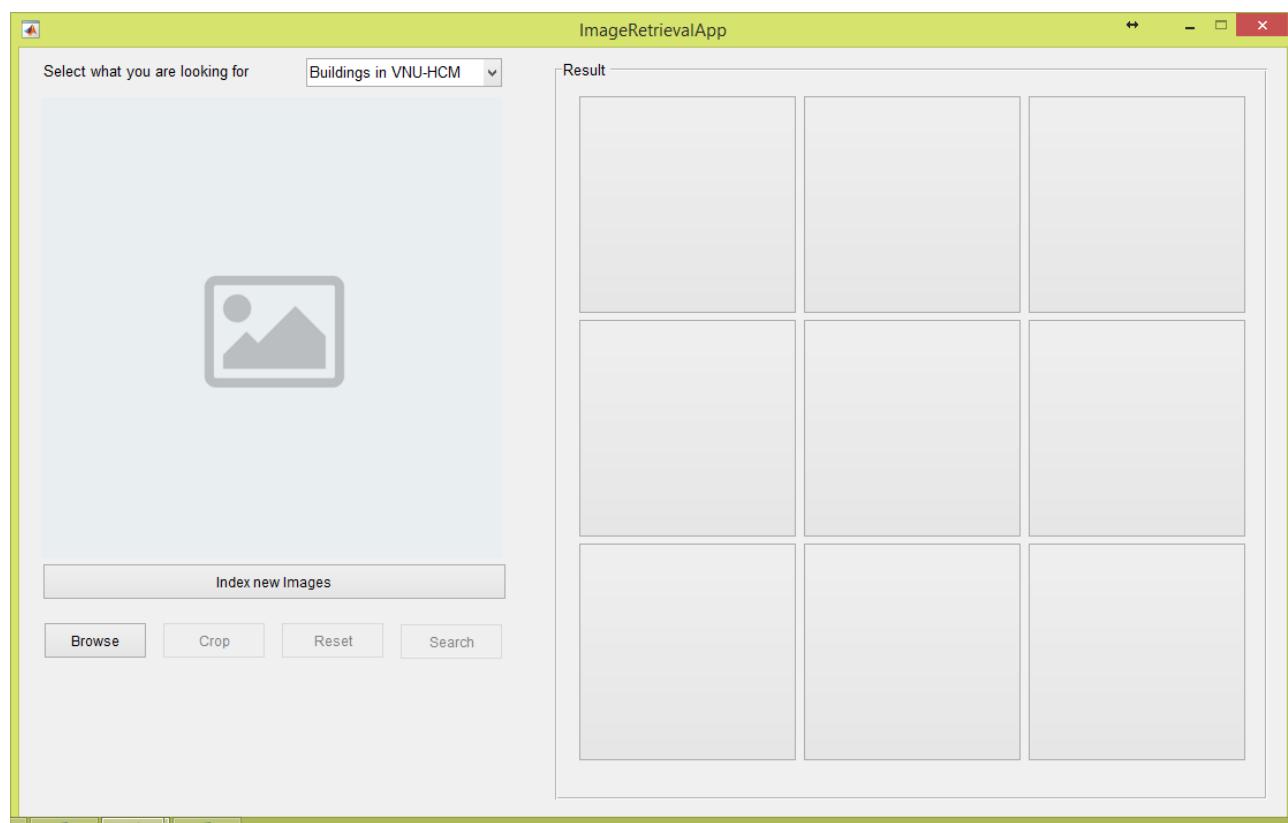
Trong nội dung source code, bước tính chỉ số Precision@k do tác giả tự xây dựng dựa theo công thức: $\text{Precision}@k = (\text{số kết quả phù hợp trong } k \text{ kết quả đầu tiên})/k$, với $k = 27$.

```
function precisionAtK = calculatePrecisionAtK(combinedResult,
category, k)
    rightResultCount = 0;
    % Calculate precision at k
    for (i = 1:k)
        foundImgPath = combinedResult(i, 1);
        if (contains(foundImgPath, category))
            rightResultCount = rightResultCount + 1;
        end
    end
    precisionAtK = rightResultCount/k;
end
```

CHƯƠNG 3. GIỚI THIỆU ỨNG DỤNG TRUY VẤN ẢNH

I. Giới thiệu giao diện ứng dụng

1.1. Màn hình chính

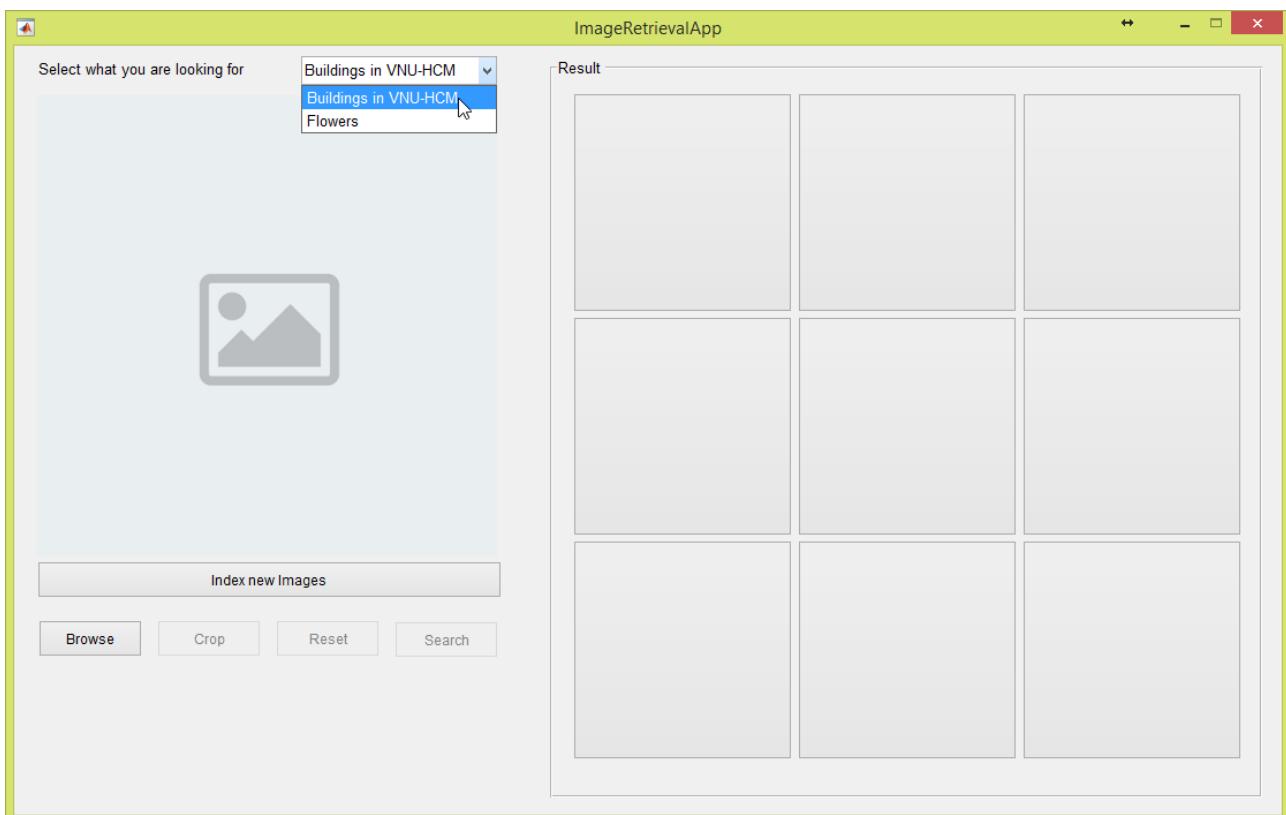


Hình 4. Màn hình chính

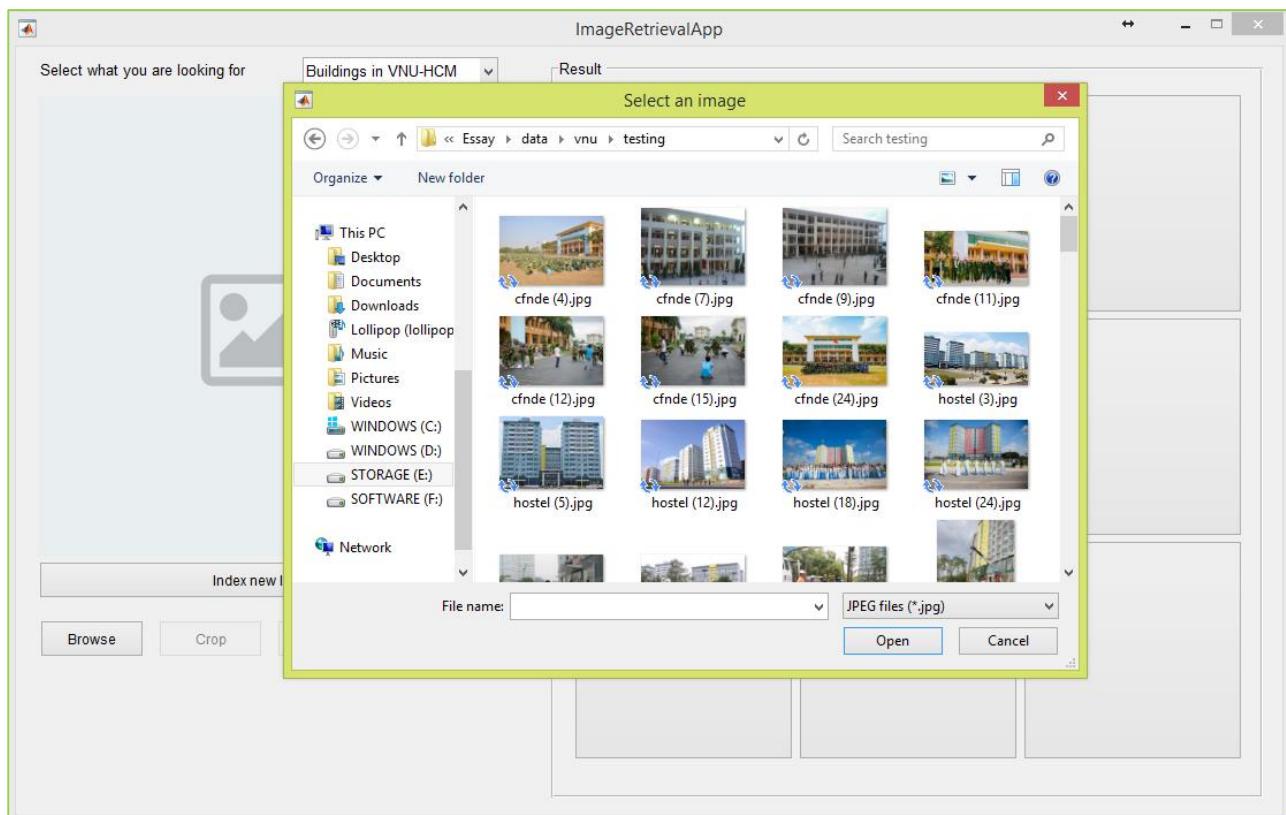
Màn hình chính bao gồm 2 phần: phần bên trái hỗ trợ chọn ảnh, vùng ảnh để truy vấn và chọn chủ đề cần truy vấn (tất cả nhà thuộc ĐHQG-HCM hay các loài hoa) và chức năng thêm ảnh mới vào kho dữ liệu ảnh; phần bên phải hiển thị kết quả trả về.

1.2. Truy vấn ảnh

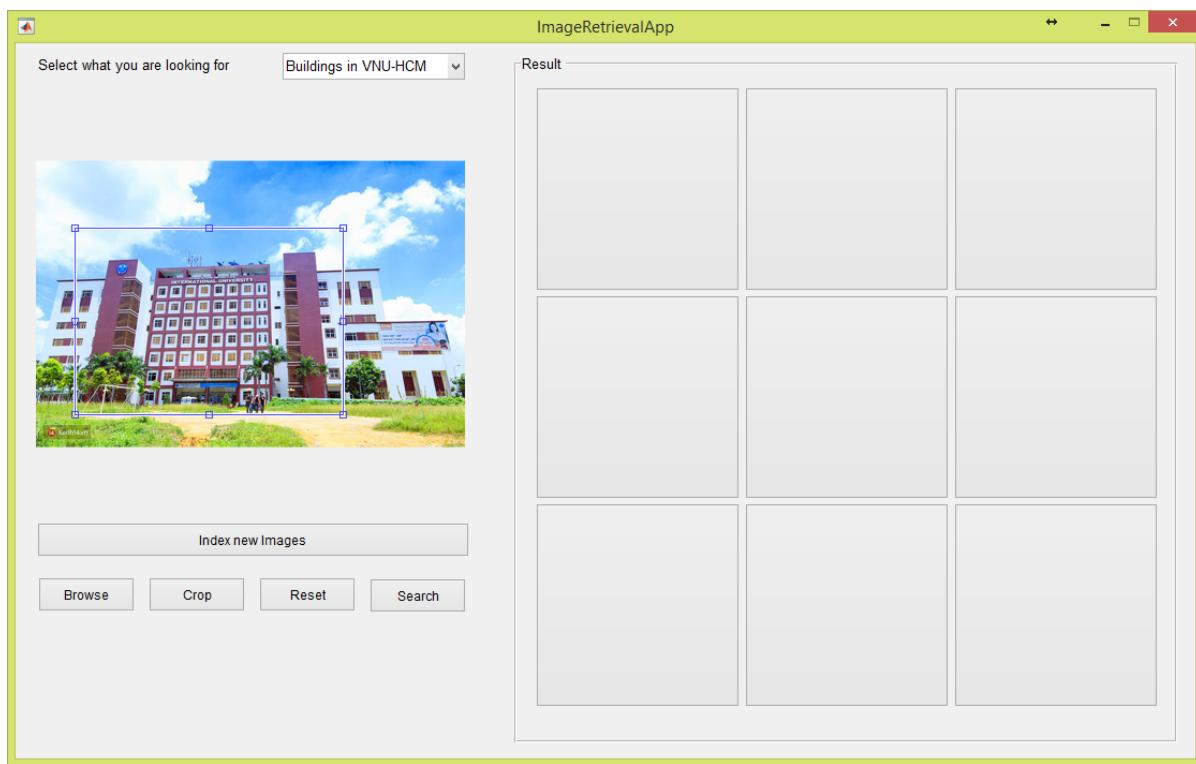
Giao diện hỗ trợ chọn chủ đề truy vấn, chọn ảnh, vùng ảnh để truy vấn thông qua các buttons: Browse, Crop, Reset.



Hình 5. Chọn chủ đề

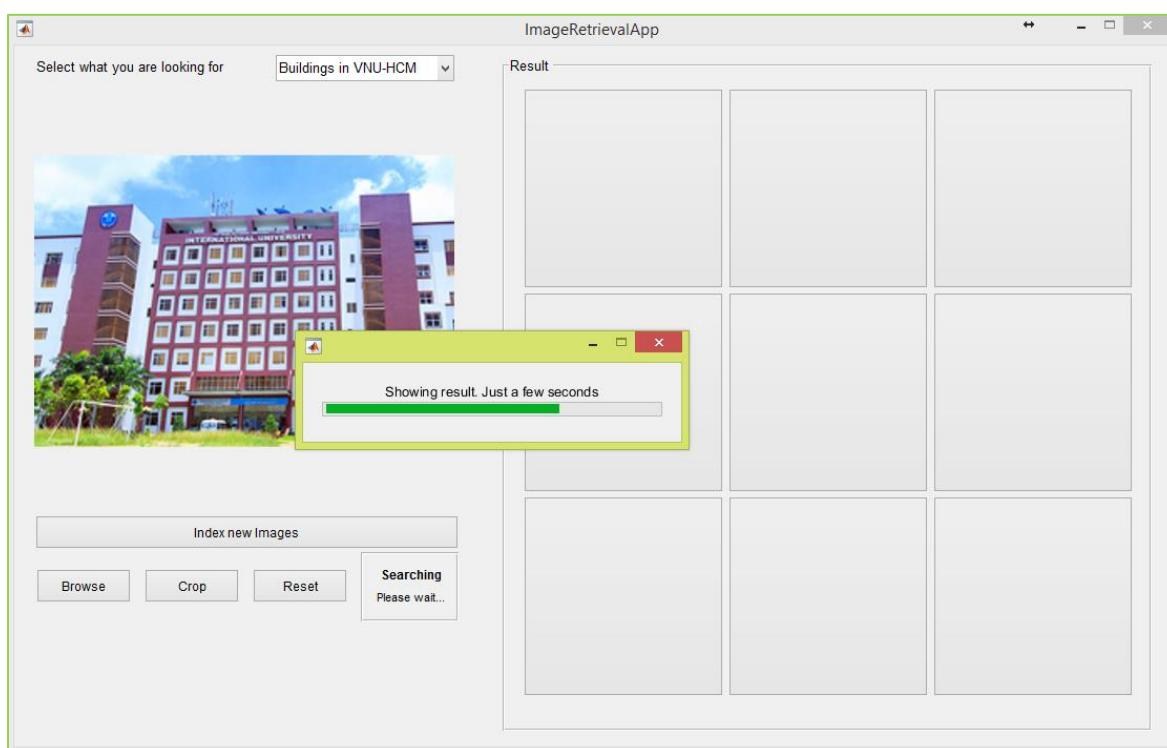


Hình 6. Chọn ảnh truy vấn

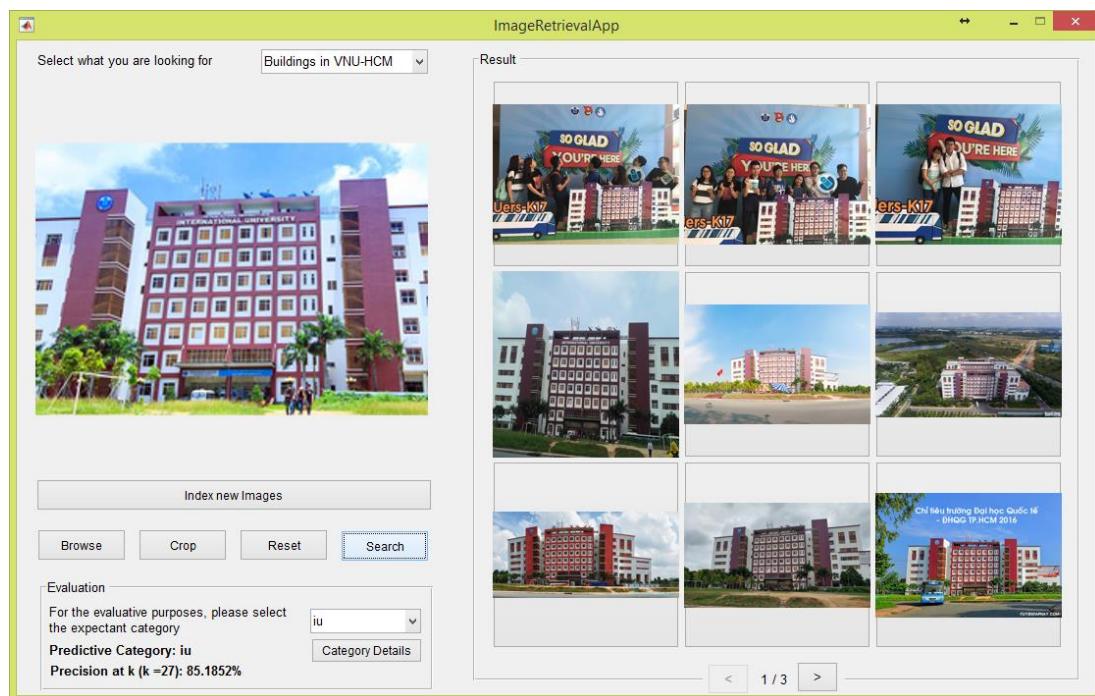


Hình 7. Chọn vùng ảnh truy vấn

Sau khi thực hiện truy vấn, kết quả từ 3 hệ thống được kết hợp để trả về kết quả tốt hơn. Kết quả trả về được phân trang thành 3 trang, mỗi trang hiển thị 9 ảnh.

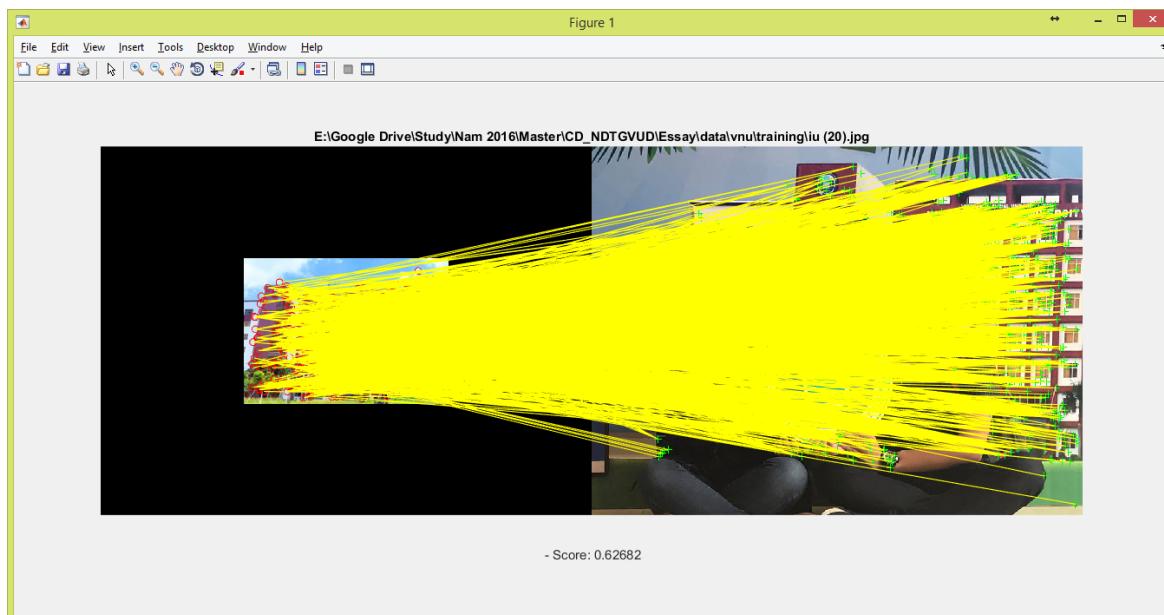


Hình 8. Giao diện tìm kiếm

**Hình 9.** Hiển thị kết quả tìm kiếm

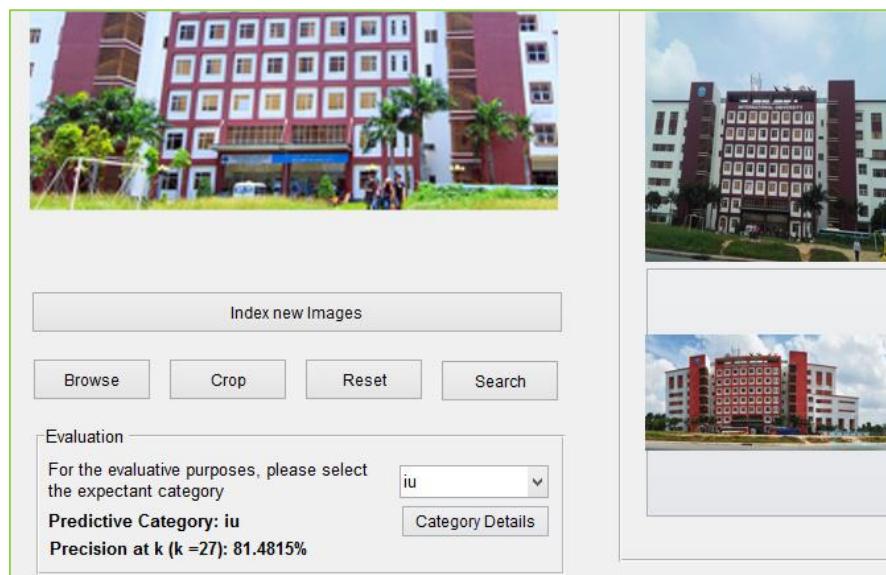
1.3. Giao diện chi tiết ảnh trả về

Nội dung chi tiết về các ảnh trong kết quả truy vấn được hiển thị khi người dùng nhấn chọn một ảnh trong danh sách ảnh trả về. Nội dung chi tiết bao gồm: đường dẫn đến tập tin ảnh, score (mức độ tương đồng) và minh họa vị trí các điểm ảnh tương thích giữa ảnh truy vấn và ảnh kết quả.

**Hình 10.** Nội dung chi tiết của ảnh trả về

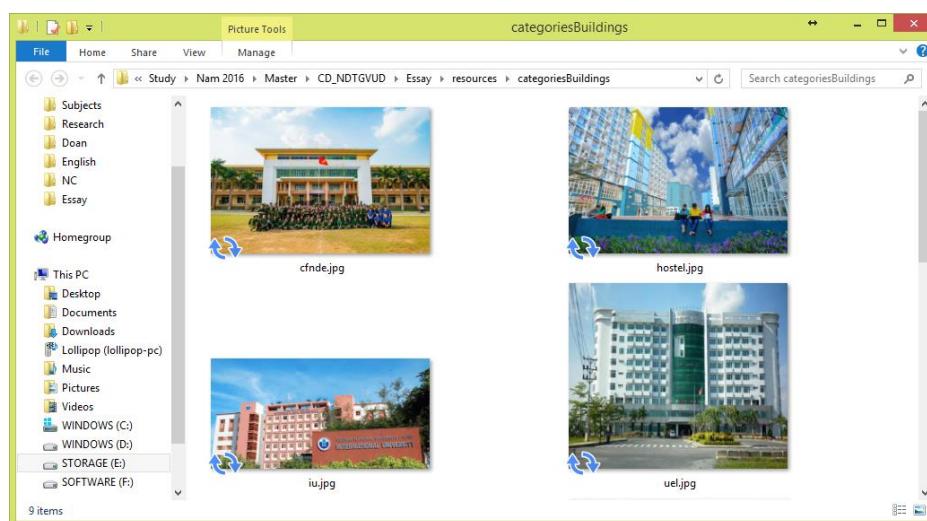
1.4. Đánh giá kết quả

Hệ thống sử dụng chỉ số Precision@k để đánh giá hệ thống truy vấn với $k = 27$ là tổng số lượng ảnh trả về trong 3 trang đầu tiên. Hệ thống dự đoán category phù hợp cho bức ảnh truy vấn dựa trên danh sách kết quả trả về. Kết quả dự đoán category sẽ được dùng để tính chỉ số Precision@k một cách tự động. Bên cạnh đó, nếu hệ thống dự đoán không chính xác category, người dùng có thể tự lựa chọn lại category phù hợp để hệ thống tính lại chỉ số Precision@k.

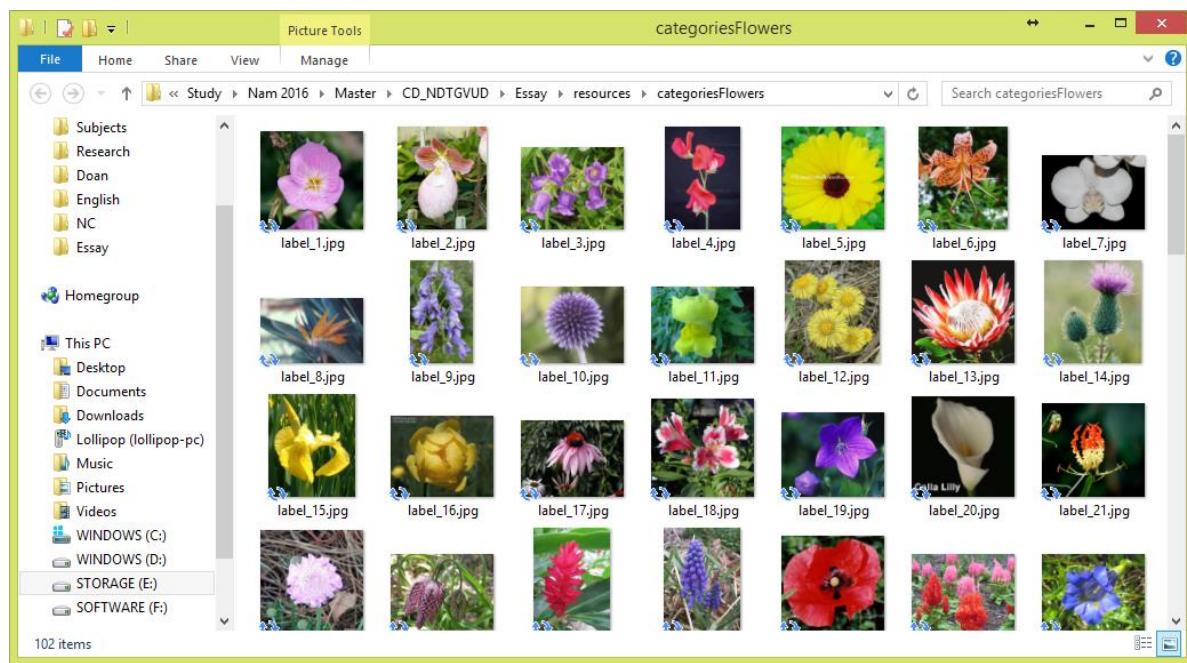


Hình 11. Kết quả đánh giá Precision@k

Nếu người dùng gặp khó khăn trong việc xác định category cho bức ảnh truy vấn, hệ thống hỗ trợ hiển thị danh sách các ảnh mẫu cho từng category khi nhấn vào Categories Details.



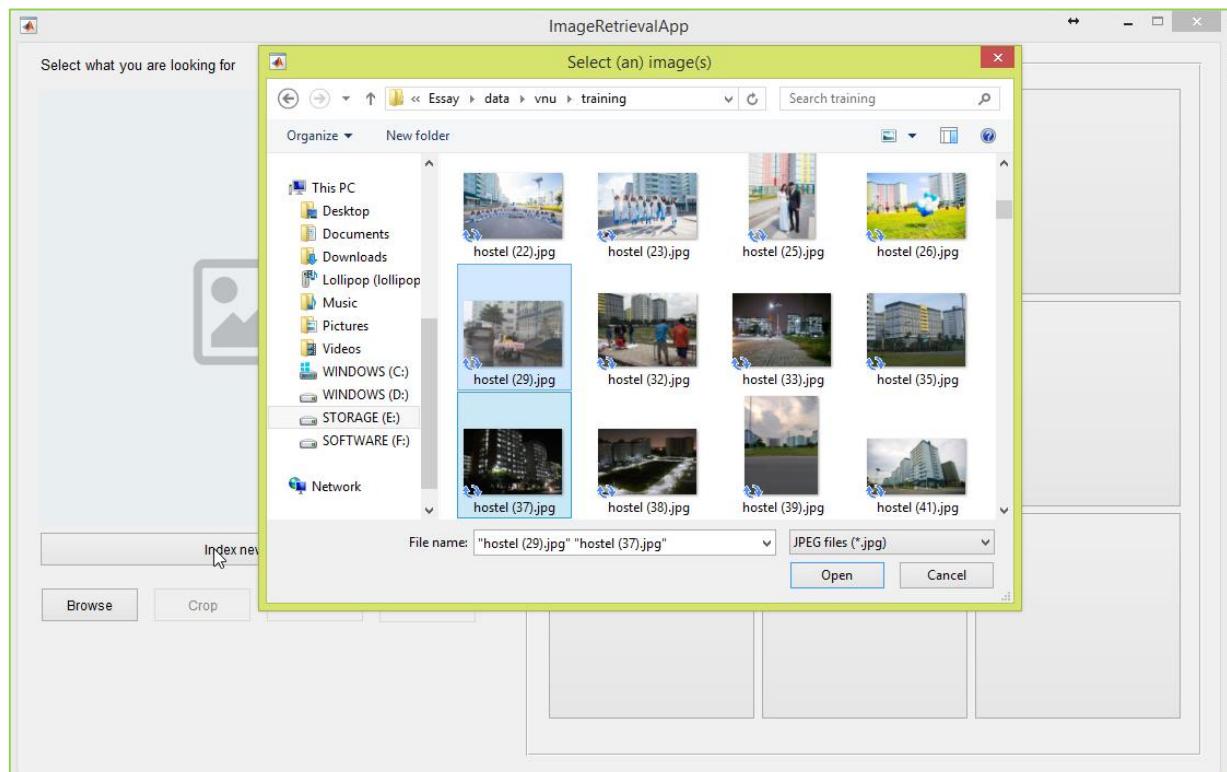
Hình 12. Hình mẫu các categories tòa nhà tại ĐHQG-HCM



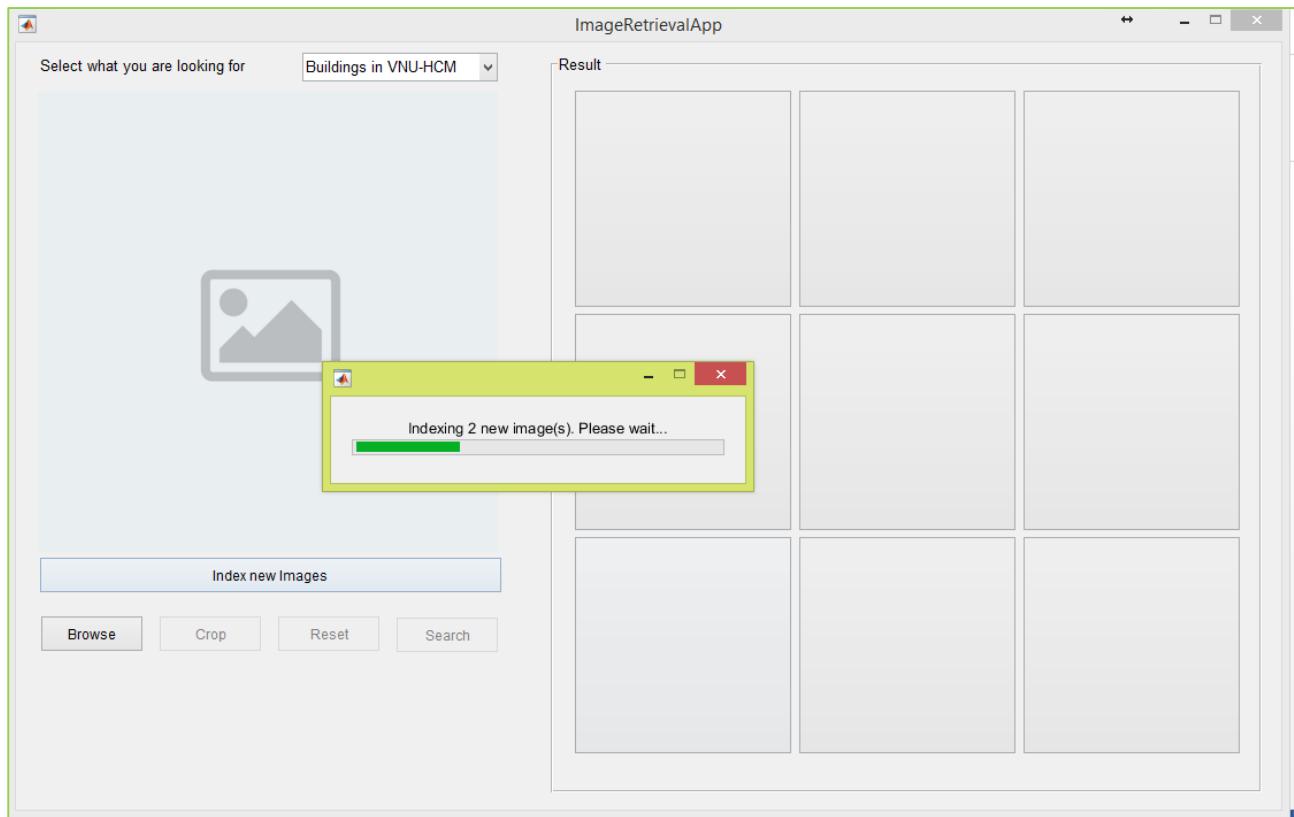
Hình 13. Hình mẫu các categories các loài hoa phổ biến tại Anh

1.5. Thêm ảnh mới vào kho dữ liệu truy vấn

Hệ thống hỗ trợ thêm một hoặc nhiều ảnh cùng lúc vào kho dữ liệu ảnh để truy vấn mà không cần trích chọn đặc trưng và xây dựng lại Bag of words từ đầu.



Hình 14. Chọn nhiều ảnh mới thêm vào kho dữ liệu



Hình 15. Hệ thống thêm ảnh mới vào kho dữ liệu

II. Demo ứng dụng

Để demo ứng dụng, tác giả đã thực hiện một video ngắn (khoảng 2 phút) để demo các thao tác truy vấn ảnh, xem chi tiết kết quả và đánh giá hệ thống.

- Link chính: <https://youtu.be/WLWmzAbnz-I>
- Link backup: <https://youtu.be/j7-L6BQH-MM>

CHƯƠNG 4. KẾT QUẢ THỬ NGHIỆM

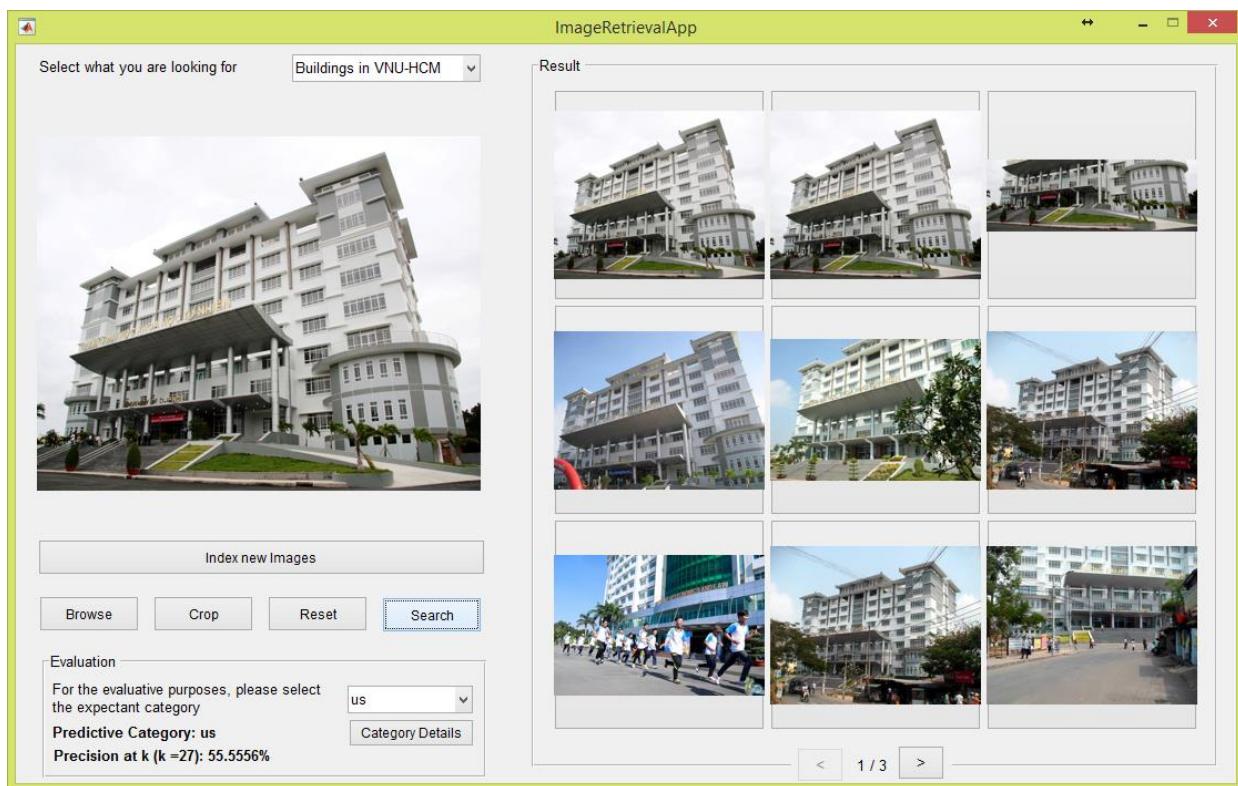
Để đánh giá hệ thống, với mỗi bộ dữ liệu, đồ án thực hiện 10 thử nghiệm. Các thử nghiệm được thực hiện trên 5 ảnh test ngẫu nhiên trong bộ ảnh testing. Mỗi ảnh sẽ được thử nghiệm truy vấn toàn bộ nội dung ảnh và truy vấn một vùng ảnh. Hệ thống được đánh giá qua chỉ số Precision@k ($k=27$).

I. Kết quả thử nghiệm với bộ dữ liệu các tòa nhà tại ĐHQG-HCM

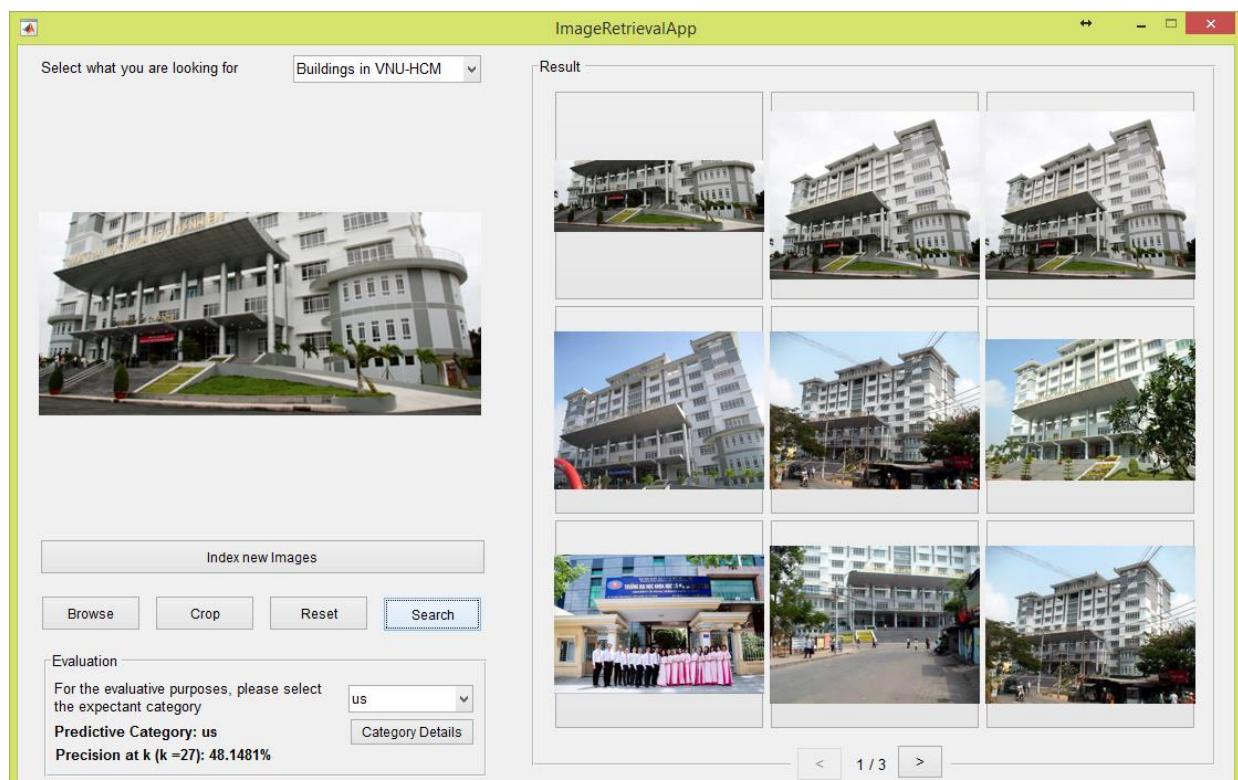
Bảng 4. Kết quả đánh giá hệ thống với bộ dữ liệu các tòa nhà tại ĐHQG-HCM

STT	Thử nghiệm	Ảnh test	Precision@k	Dự đoán category
1	Truy vấn ảnh	iu (51).jpg	55,56%	IU
2	Truy vấn vùng ảnh		81,48%	IU
3	Truy vấn ảnh	uit (17).jpg	33,33%	Hostel (sai)
4	Truy vấn vùng ảnh		25,93%	UIT
5	Truy vấn ảnh	us (66).jpg	55,56%	US
6	Truy vấn vùng ảnh		48,15%	US
7	Truy vấn ảnh	uss (21).jpg	33,33%	US (sai)
8	Truy vấn vùng ảnh		18,52%	US (sai)
9	Truy vấn ảnh	ut (28).jpg	25,93%	Hostel (sai)
10	Truy vấn vùng ảnh		37,04%	UT
Kết quả trung bình			41,48%	

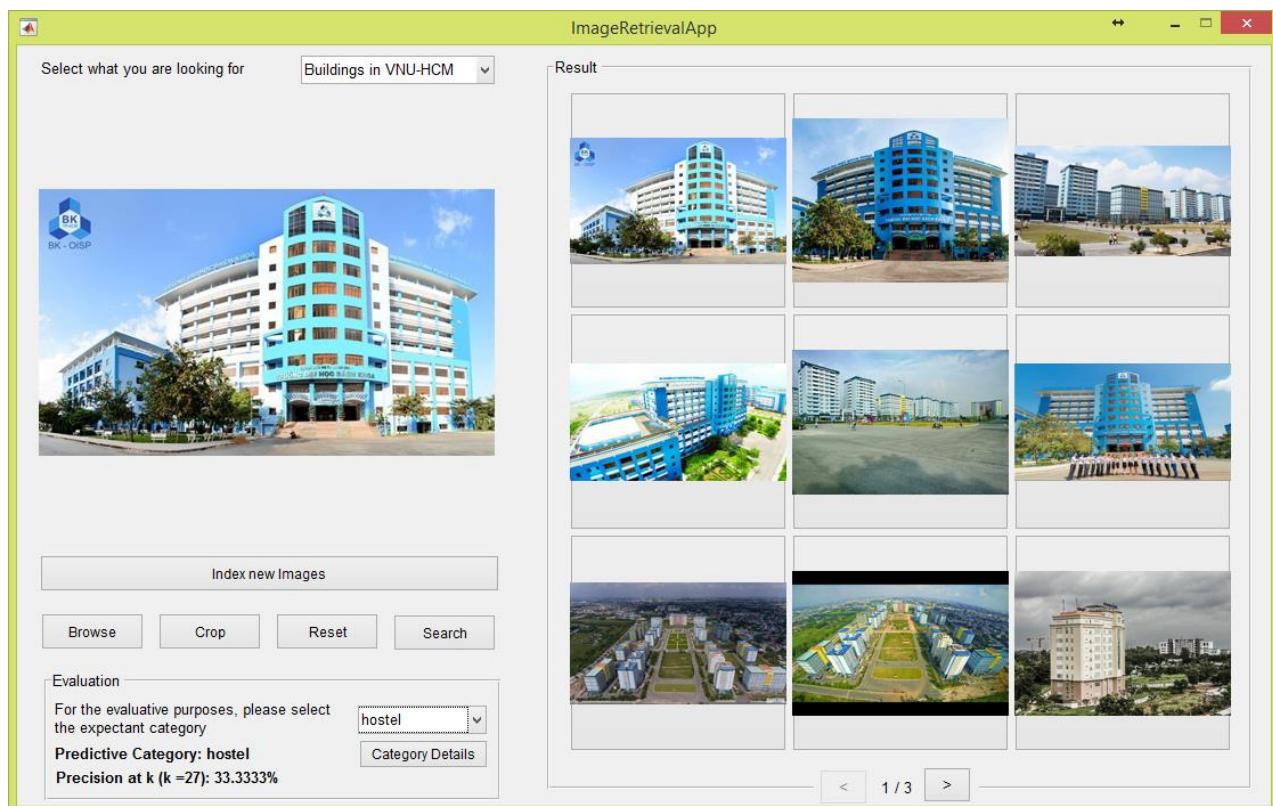
Một số hình ảnh thử nghiệm hệ thống:



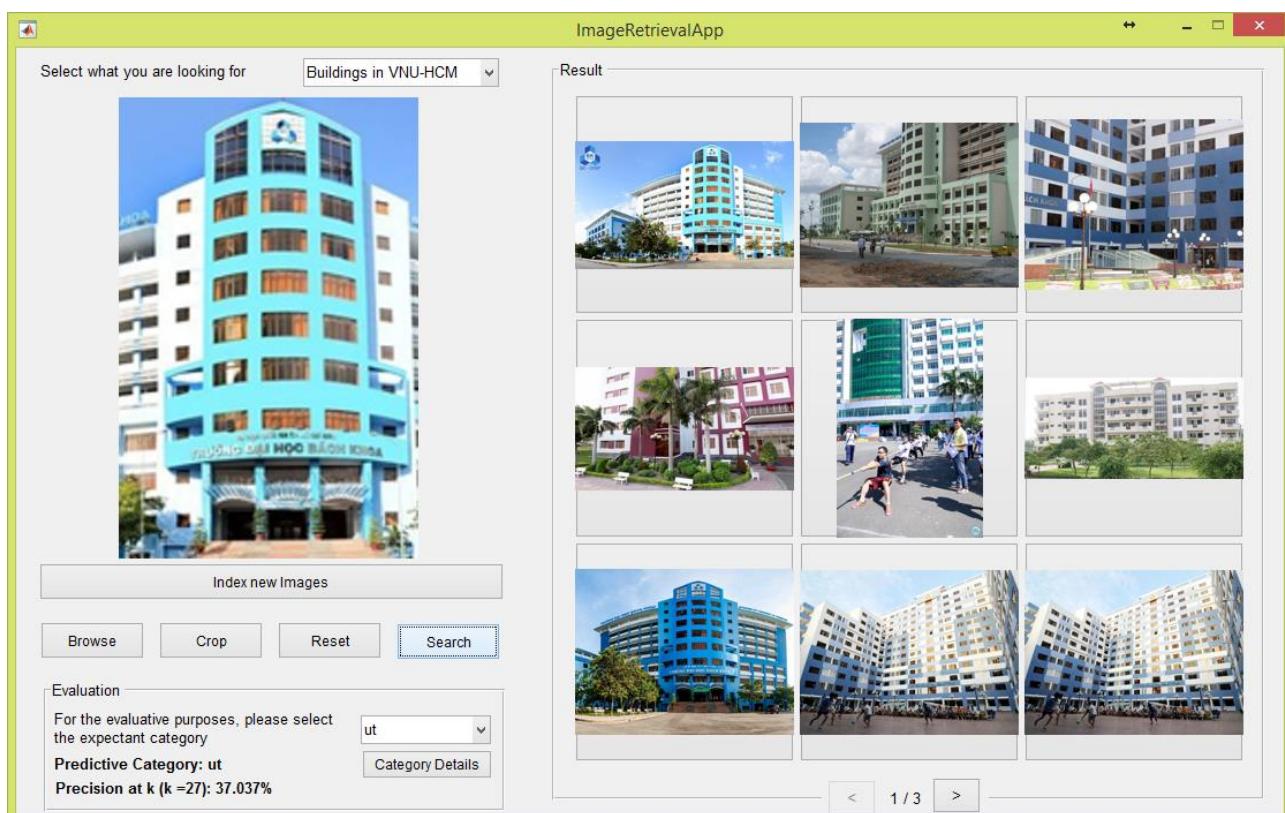
Hình 16. Kết quả truy vấn ảnh us (66).jpg



Hình 17. Kết quả truy vấn một vùng ảnh us (66).jpg



Hình 18. Hệ thống dự đoán sai category khi truy vấn ảnh ut (28).jpg



Hình 19. Hệ thống dự đoán đúng category khi truy vấn một vùng ảnh ut (28).jpg

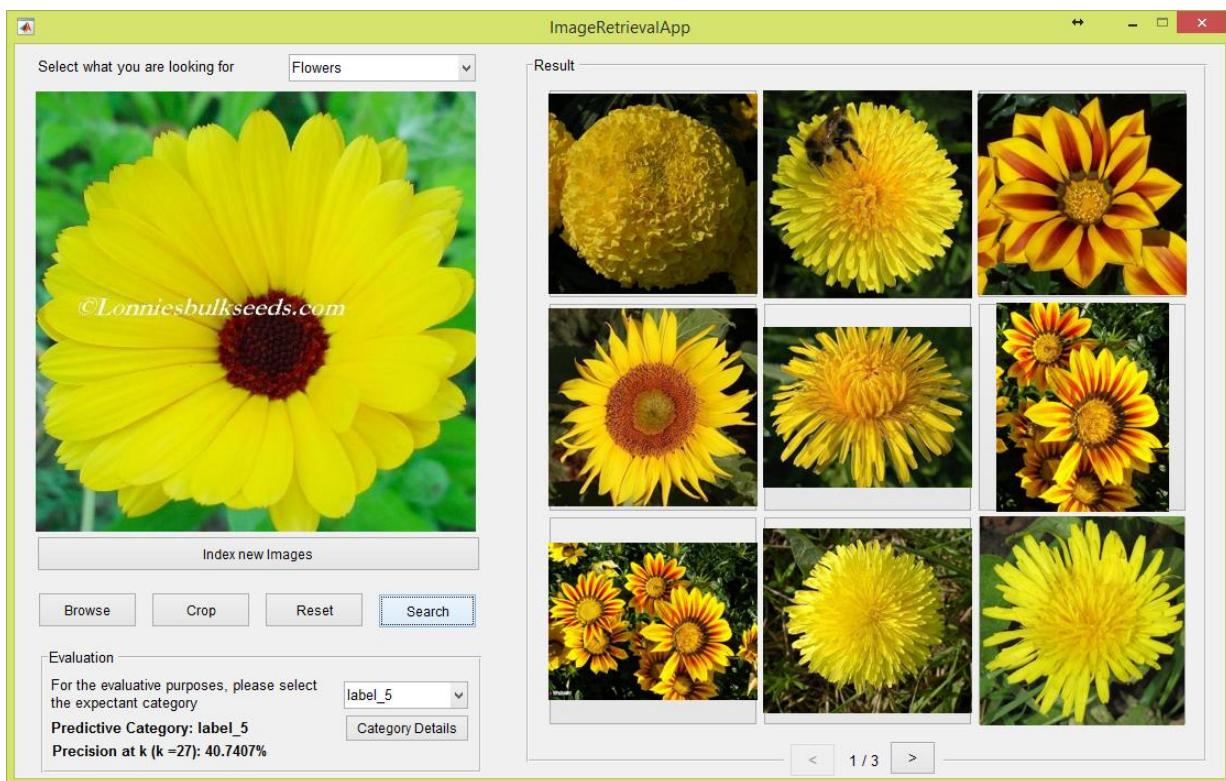
Nhân xét: Kết quả thử nghiệm cho thấy hệ thống truy vấn ảnh cho kết quả khá tốt trên bộ dữ liệu các tòa nhà tại ĐHQG-HCM. Chỉ số Precision@k trung bình cho 10 query đạt 41,48% (cao hơn các hệ thống chỉ sử dụng 1 phương pháp trích chọn đặc trưng riêng lẻ đã thử nghiệm trong phần trước). Kết quả này cho thấy hướng tiếp cận kết hợp kết quả từ các hệ thống tìm kiếm là hướng tiếp cận khả thi và mang đến hiệu quả. Bên cạnh đó, chức năng dự đoán category cũng hoạt động khá tốt mặc dù có một số trường hợp dự đoán sai. Để cải tiến hệ thống, nâng cao độ chính xác, ta có thể thử nghiệm thêm việc sử dụng các features detectors và descriptors để chọn ra những phương pháp trích chọn đặc trưng phù hợp nhất cho bài toán truy vấn ảnh liên quan đến các tòa nhà.

II. Kết quả thử nghiệm với bộ dữ liệu các loài hoa phổ biến tại Anh

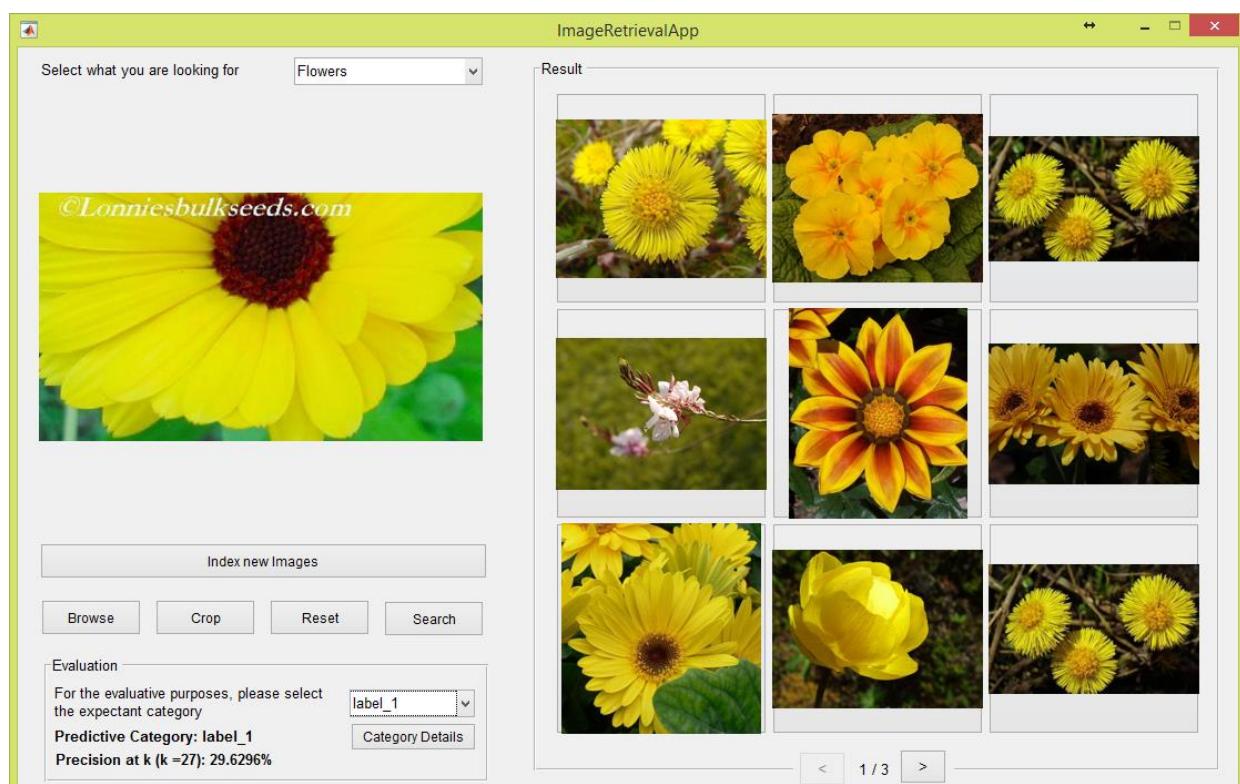
Bảng 5. Kết quả đánh giá hệ thống với bộ dữ liệu các loài hoa phổ biến tại Anh

STT	Thử nghiệm	Ảnh test	Precision@k	Dự đoán category
1	Truy vấn ảnh	label_5_1.jpg	40,74%	Label_5
2	Truy vấn vùng ảnh		22,22%	Label_1 (sai)
3	Truy vấn ảnh	label_10_1.jpg	11,11%	Label_1 (sai)
4	Truy vấn vùng ảnh		14,81%	Label_2 (sai)
5	Truy vấn ảnh	label_66_7.jpg	44,44 %	Label_6 (sai)
6	Truy vấn vùng ảnh		29,63%	Label_3 (sai)
7	Truy vấn ảnh	label_77_34.jpg	22,22%	Label_7 (sai)
8	Truy vấn vùng ảnh		51,85%	Label_7 (sai)
9	Truy vấn ảnh	label_102_3.jpg	0%	Label_7 (sai)
10	Truy vấn vùng ảnh		7,41%	Label_8 (sai)
Kết quả trung bình			24,44%	

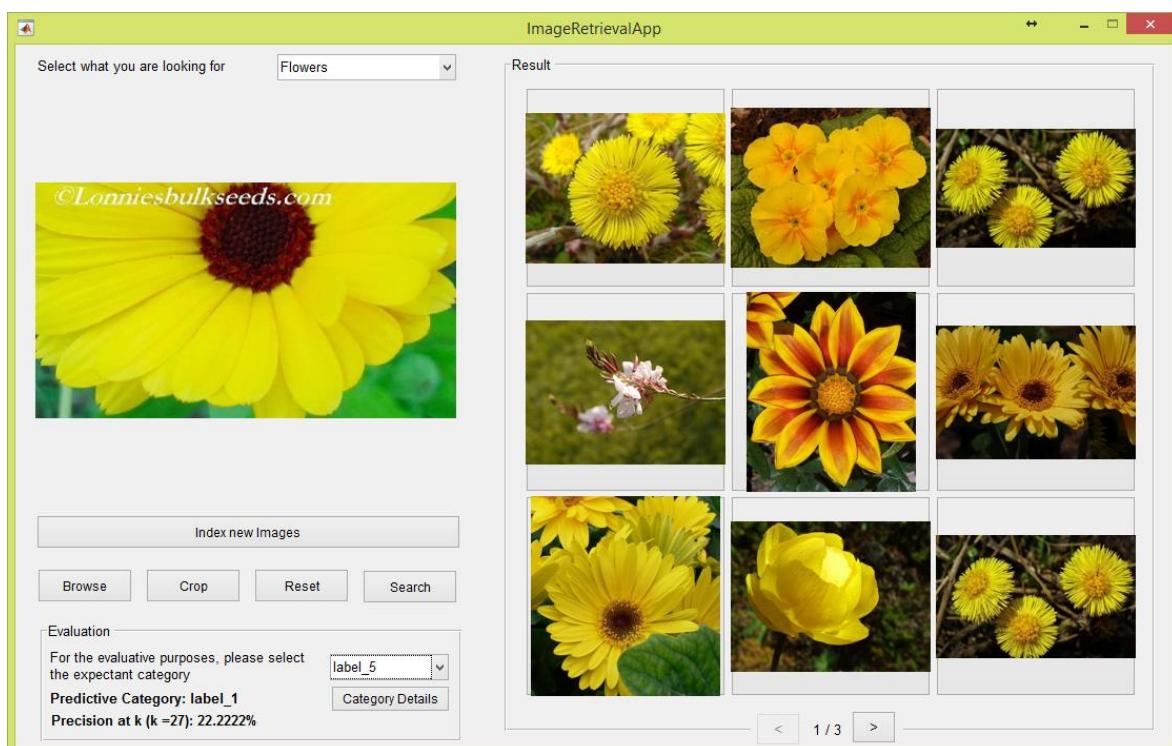
Một số hình ảnh thử nghiệm trên bộ dữ liệu ảnh các loài hoa:



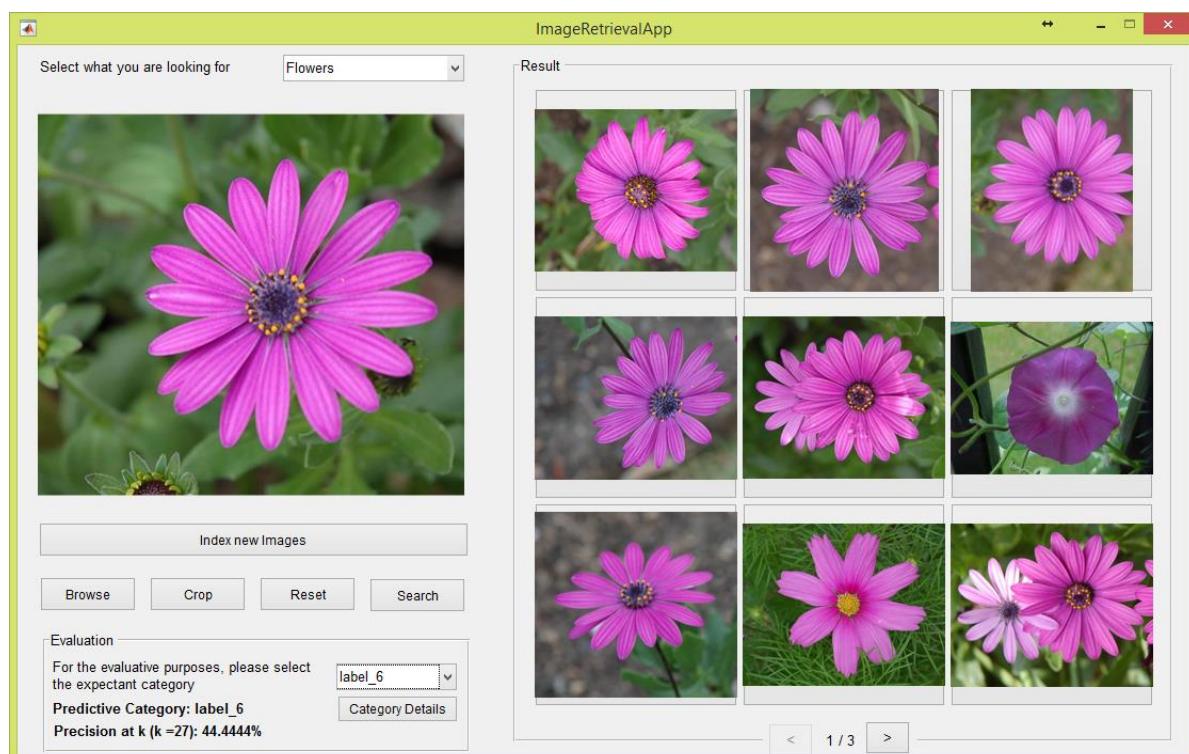
Hình 20. Kết quả truy vấn ảnh label_5_1.jpg



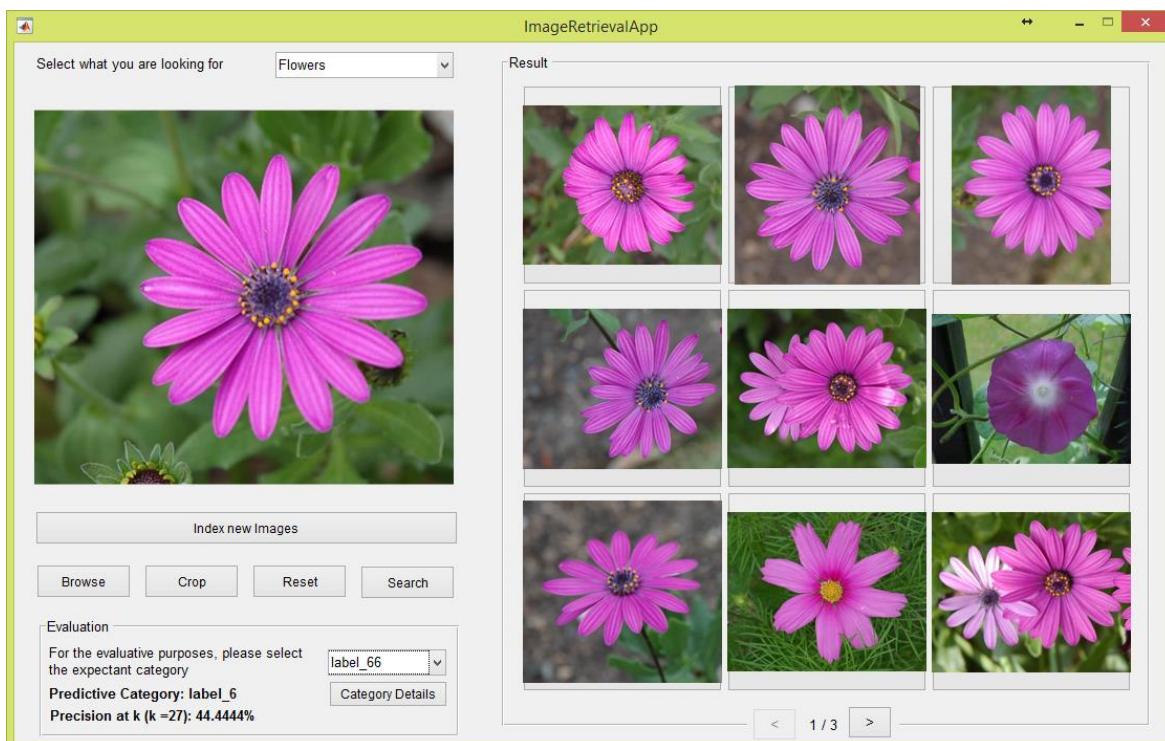
Hình 21. Hệ thống dự đoán sai category khi truy vấn một vùng ảnh label_5_1.jpg



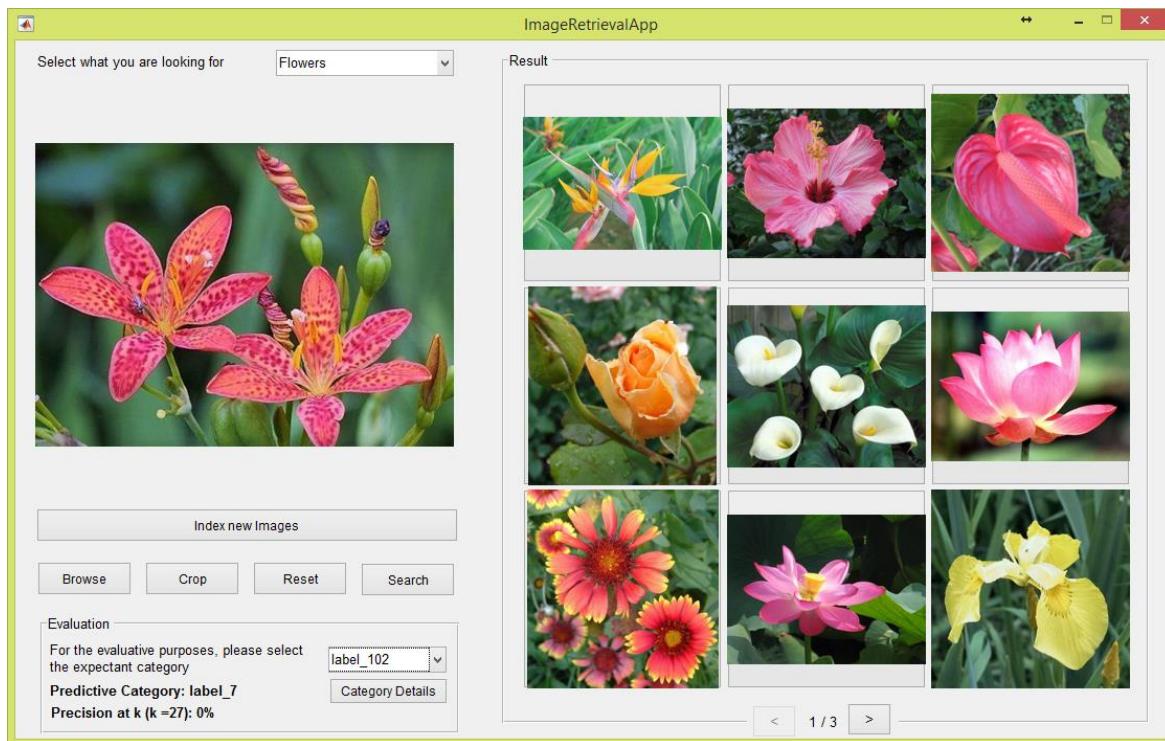
Hình 22. Hệ thống tính toán lại Precision@k sau khi người dùng chọn lại category đúng cho vùng ảnh label_5_1.jpg



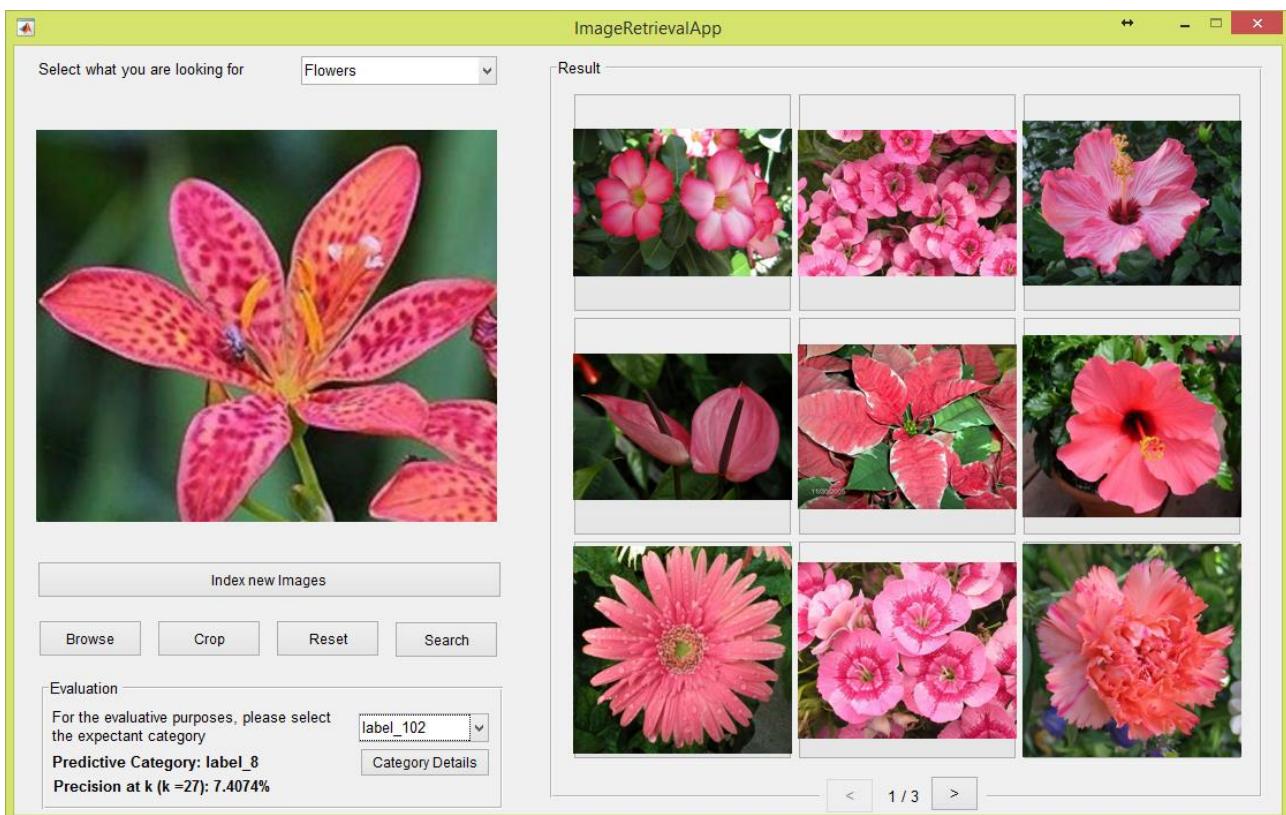
Hình 23. Hệ thống dự đoán sai category khi truy vấn ảnh label_66_7.jpg do hai category label_6 và label_66 có cùng số lượng kết quả trả về



Hình 24. Hệ thống tự tính toán lại thông số Precision@k sau khi người dùng chọn lại category đúng cho ảnh label_66_7.jpg



Hình 25. Hệ thống dự đoán sai category khi truy vấn ảnh label_102_3.jpg do trong danh sách kết quả trả về không chứa ảnh nào thuộc category label_102



Hình 26. Hệ thống trả về kết quả tốt hơn trong trường hợp truy vấn một vùng ảnh
label_102_3.jpg

Nhận xét:

- Nhìn chung hệ thống khi thực hiện truy vấn trên bộ dữ liệu các loài hoa cho kết quả có độ chính xác thấp hơn so với khi truy vấn trên bộ dữ liệu các tòa nhà. Nguyên nhân có thể giải thích là do các tòa nhà có các đặc trưng hình dáng, góc cạnh mang đặc điểm phân biệt giữa các tòa nhà phù hợp với các phương pháp trích chọn đặc trưng trong đồ án này hơn. Bộ dữ liệu hình ảnh về các loài hoa có độ đa dạng cao trong màu sắc, hình dáng hoa, có nhiều hoa cũng màu sắc và hình dáng tương tự rất khó phân biệt ngay cả khi phân loại thủ công bằng mắt người. Hệ thống sử dụng kết hợp kết quả từ ba phương pháp trích chọn đặc trưng cho bộ dữ liệu ảnh các loài hoa, trong đó 2/3 phương pháp sử dụng đặc trưng màu sắc nhưng trong bộ dữ liệu có rất nhiều loài hoa có màu sắc giống nhau nên khi truy vấn với các ảnh test thuộc các loài hoa có màu sắc phô biến như đỏ, tím thường cho kết quả có độ chính xác thấp hơn các truy vấn với các loại hoa có màu sắc không phô biến.
- Bên cạnh đó, hệ thống sau khi kết hợp 3 phương pháp trích chọn đặc trưng với nhau thì kết quả trả về của hệ thống nhìn chung thấp hơn so với khi sử dụng một phương pháp trích chọn đặc trưng như phần trình bày thử nghiệm ở phần II chương 2. Nguyên nhân được giải thích là do thử nghiệm ban đầu chỉ được thực hiện trên bộ dữ liệu nhỏ (gồm 30

categories) nên khi thử nghiệm trên bộ dữ liệu lớn (102 categories) thì xác suất hệ thống trả về các kết quả không phù hợp sẽ phải tăng lên.

- *Chức năng dự đoán category cho kết quả có độ chính xác thấp hơn* so với khi thử nghiệm trên bộ dữ liệu các tòa nhà. Một trong những nguyên nhân dẫn đến kết quả dự đoán chưa tốt là do trong nhiều trường hợp có hai hoặc nhiều category có số lượng kết quả trả về bằng nhau trong danh sách kết quả trả về thì hệ thống mặc định ưu tiên category đứng trước (theo thứ tự bảng chữ cái). Để cải tiến trường hợp này, ta có thể thay đổi việc lựa category dự đoán khi các category có cùng số lượng kết quả trả về dựa trên thứ tự xuất hiện của các kết quả, tức là category nào có nhiều ảnh xuất hiện ở các vị trí đầu danh sách kết quả trả về (dựa trên mức độ phù hợp với ảnh truy vấn) nhiều hơn sẽ được chọn là kết quả dự đoán category. Giải pháp này sẽ giúp hệ thống dự đoán khách quan và công bằng hơn giữa các category với nhau.

CHƯƠNG 5. KẾT LUẬN

I. Kết quả đạt được

Đồ án đã xây dựng được một hệ thống tìm kiếm ảnh đáp ứng các yêu cầu cần thiết của một hệ thống tìm kiếm ảnh liên quan đến các tòa nhà và các loài hoa, cụ thể như sau:

- Hệ thống cung cấp giao diện người dùng đáp ứng các yêu cầu: chọn ảnh, vùng ảnh truy vấn; hiển thị danh sách kết quả (có phân trang); hiển thị chi tiết ảnh kết quả với các thông tin độ tương đồng và minh họa các điểm ảnh tương đồng với ảnh truy vấn.
- Hệ thống hỗ trợ truy vấn ảnh từ một trong 2 chủ đề: các tòa nhà của các đơn vị, trường thành viên thuộc ĐHQG-HCM và các loài hoa phổ biến tại Vương quốc Anh.
- Về chức năng, hệ thống đáp ứng yêu cầu truy vấn trả về các ảnh tương đồng trong bộ ảnh training và dự đoán category cho ảnh truy vấn với độ chính xác khá tốt. Hệ thống được đánh giá thông qua chỉ số Precision@k phù hợp để đánh giá các hệ thống truy vấn ảnh.
- Hệ thống thử nghiệm và sử dụng nhiều phương pháp phát hiện và trích chọn đặc trưng khác nhau (SURF, HoG, Color). Việc kết hợp kết quả từ 3 hệ thống sử dụng các phương pháp trích chọn đặc trưng khác nhau để tìm ra kết quả trả về phù hợp góp phần nâng cao độ chính xác của hệ thống truy vấn.
- Hệ thống hỗ trợ thêm ảnh mới vào kho dữ liệu ảnh sau khi đã xây dựng hệ thống mà không cần thực hiện lại quá trình trích chọn đặc trưng và xây dựng Bag of visual words giúp tiết kiệm thời gian cải tiến hệ thống.

II. Những hạn chế

Do giới hạn về thời gian, kiến thức cũng như kinh nghiệm, đồ án vẫn còn tồn tại các hạn chế:

- Bộ dữ liệu về các tòa nhà tại ĐHQG-HCM có số lượng chưa đủ lớn, chưa mang tính đa dạng về góc chụp, thời gian chụp, một số ảnh chất lượng thấp ảnh hưởng đến kết quả truy vấn của hệ thống.
- Hình ảnh về các tòa nhà có những đặc điểm đặc trưng như hình dáng, góc cạnh. Việc tìm ra phương pháp trích chọn đặc trưng phù hợp với bài toán truy vấn ảnh về các tòa nhà sẽ cải thiện đáng kể hệ thống tìm kiếm. Đồ án chỉ thử nghiệm kết hợp một số features detectors với một số features descriptor phổ biến, việc thử nghiệm đánh giá thêm các features descriptors khác như LBP, histogram, FREAK, BRISK... có thể giúp tìm ra các phương pháp phù hợp hơn cho bài toán truy vấn ảnh.
- Việc kết hợp kết quả từ 3 hệ thống tìm kiếm sử dụng các phương pháp trích chọn đặc trưng khác nhau dù góp phần tăng tính chính xác của hệ thống nhưng hướng tiếp cận này khá thô sơ và yêu cầu đánh đổi về thời gian tìm kiếm.
- Hàm dự đoán category tương ứng với hình ảnh truy vấn dựa trên danh sách các kết quả trả về chưa xử lý tốt trường hợp các category có cùng số lượng kết quả trả về. Điều này dẫn đến việc thiếu công bằng trong việc dự đoán giữa các category (các category có vị trí đứng trước theo thứ tự bảng chữ cái sẽ được ưu tiên hơn). Giải pháp khắc phục hạn chế này đã được trình bày trong phần nhận xét kết quả thử nghiệm trên bộ dữ liệu ảnh các loài hoa phổ biến tại Anh của chương 4.
- Việc đánh giá hệ thống còn khá đơn giản, chỉ sử dụng một chỉ số đánh giá là Precision@k và thực hiện trên số lượng query ít. Vì vậy để đảm bảo tính khách quan và chính xác, hệ thống cần bổ sung thêm các chỉ số đánh giá khác như R-Precision, Parameterized F Measure, MAP... và thực hiện trên một số lượng lớn query thử nghiệm.

TÀI LIỆU THAM KHẢO

- [1] N. T. Hoàn, "Phương pháp trích chọn đặc trưng ảnh trong thuật toán học máy tìm kiếm ảnh áp dụng vào bài toán tìm kiếm sản phẩm," Hà Nội, 2010.
- [2] N. X. T. Huy and N. H. Vũ, "Tìm hiểu kỹ thuật phát hiện người trong video giám sát lớp học," Hà Nội, 2010.
- [3] "Local Feature Detection and Extraction," [Online]. Available: <https://www.mathworks.com/help/vision/ug/local-feature-detection-and-extraction.html>. [Accessed 01 Dec 2017].
- [4] [Online]. Available: <https://www.mathworks.com/help/vision/ref/retrieveimages.html>. [Accessed 02 Dec 2017].
- [5] "Image Retrieval Using Customized Bag of Features," [Online]. Available: <https://www.mathworks.com/help/vision/examples/image-retrieval-using-customized-bag-of-features.html>. [Accessed 03 Dec 2017].
- [6] "Create a Custom Feature Extractor," [Online]. Available: <https://www.mathworks.com/help/vision/ug/create-a-custom-feature-extractor.html>. [Accessed 02 Dec 2017].