

**BỘ GIÁO DỤC
VÀ ĐÀO TẠO**

**VIỆN HÀN LÂM KHOA HỌC
VÀ CÔNG NGHỆ VIỆT NAM**

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ



NGUYỄN MINH HẢI

**PHÁT TRIỂN THUẬT TOÁN TRA CỨU ẢNH
DỰA TRÊN NGỮ NGHĨA**

TÓM TẮT LUẬN ÁN TIẾN SĨ NGÀNH MÁY TÍNH

Ngành: Khoa học máy tính

Mã số: 9 48 01 01

TP. Hồ Chí Minh - 2024

Công trình được hoàn thành tại: Học viện Khoa học và Công nghệ,
Viện Hàn lâm Khoa học và Công nghệ Việt Nam

Người hướng dẫn khoa học:

1. Người hướng dẫn 1: PGS. TS. Trần Văn Lãng, Viện Cơ và Tin học Ứng dụng, Viện Hàn lâm Khoa học và Công nghệ Việt Nam
2. Người hướng dẫn 2: TS. Văn Thế Thành, Trường Đại học Sư phạm TP. Hồ Chí Minh

Phản biện 1: PGS.TS. Trần Công Hùng

Phản biện 2: PGS.TS. Huỳnh Tường Nguyên

Phản biện 3: TS. Lê Quang Minh

Luận án được bảo vệ trước Hội đồng đánh giá luận án tiến sĩ cấp Học viện họp tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam vào hồi giờ, ngày tháng năm

Có thể tìm hiểu luận án tại:

1. Thư viện Học viện Khoa học và Công nghệ
2. Thư viện Quốc gia Việt Nam

MỞ ĐẦU

1. Tổng quan

Thị giác máy tập trung chủ yếu vào ứng dụng của hình ảnh điện tử và máy tính hỗ trợ, đặc biệt trong việc phân tích hình ảnh thu thập từ các hệ thống như theo dõi người đi bộ, phân tích giao thông, giám sát an ninh, và khám phá không gian sâu [1]. Tuy nhiên, để đạt được kết quả tốt hơn trong ứng dụng thị giác máy, người ta ngày càng đẩy mạnh việc kết hợp học máy với các phương pháp xử lý hình ảnh trong quá trình phân tích [2]. Sự kết hợp này được coi là một cách để tạo ra các hệ thống thông minh hơn và đã trở nên phổ biến trong các ứng dụng truy xuất hình ảnh [3].

2. Tính cấp thiết của luận án

CBIR không phân tích ngữ nghĩa và không mô tả được nội dung hình ảnh bằng ngữ nghĩa cấp cao, dẫn đến hiệu suất tìm kiếm vẫn chưa đáp ứng được yêu cầu của người dùng. Phương pháp tiếp cận tìm kiếm ảnh theo ngữ nghĩa (SBIR) [4, 5] nhằm mô tả các khái niệm ngữ nghĩa được đề xuất kết hợp với CBIR để nâng cao hiệu quả tìm kiếm. Tuy nhiên, giải thích ngữ nghĩa một ảnh bằng máy tính là một nhiệm vụ rất phức tạp, do đó vấn đề chính trong SBIR là chuyển đổi ảnh có kích thước lớn thành các đặc trưng cấp thấp và liên kết các nội dung cấp thấp này với nội dung ngữ nghĩa cấp cao của ảnh. Từ đó, mô hình tra cứu ảnh theo ngữ nghĩa được tạo ra, trong đó máy tính có thể hiểu được ngữ nghĩa thực tế mong muốn của người dùng. Do đó, bài toán tìm kiếm ảnh theo ngữ nghĩa là một vấn đề được quan tâm rất nhiều bởi các nhà nghiên cứu trên toàn thế giới, và có hiệu suất tốt khi được giải quyết. Với mong muốn đóng góp một phương pháp truy xuất ảnh theo ngữ nghĩa hiệu quả, luận án thực hiện đề tài: “Phát triển thuật toán tra cứu ảnh dựa trên ngữ nghĩa”.

3. Mục tiêu nghiên cứu của luận án

Mục tiêu nghiên cứu của luận án tập trung các vấn đề sau:

- (1) Phân tích và rút trích các đặc trưng ngữ nghĩa từ hình ảnh.
- (2) Hiểu và xử lý các thành phần trong hình ảnh để tạo ra một ngữ nghĩa cho ảnh.
- (3) Phát triển một hệ thống tìm kiếm và phân loại ảnh có khả năng hiểu và đáp ứng yêu cầu người dùng một cách hiệu quả.

4. Phương pháp nghiên cứu

Phương pháp lý thuyết: Tổng hợp các công trình liên quan đến truy vấn theo ngữ nghĩa sử dụng các phương pháp học máy và các cấu trúc lưu trữ dạng cây. Phân tích các ưu và nhược điểm của các công trình; nghiên cứu phương pháp làm giàu Ontology và phát triển mô hình tìm kiếm ảnh theo ngữ nghĩa dựa trên Ontology. Đề xuất các mô hình tìm kiếm ảnh dựa trên ngữ nghĩa; đánh giá thực nghiệm dựa trên mô hình đề xuất từ đó so sánh độ chính xác tìm kiếm ảnh với các công trình trong cùng lĩnh vực để có sự điều chỉnh và cải tiến phù hợp

Phương pháp thực nghiệm: Dựa trên các phương pháp và mô hình được đề xuất trong luận án, việc cài đặt chương trình thực nghiệm được triển khai trên máy có cùng cấu hình. Dữ liệu thực nghiệm được chọn là các bộ dữ liệu ảnh đáng tin cậy, được công bố rộng rãi và đã được sử dụng trong nhiều nghiên cứu trước đó để so sánh với các kết quả thực nghiệm từ các mô hình được đề xuất để minh chứng tính đúng đắn và hiệu quả của cơ sở lý thuyết

5. Đối tượng và phạm vi nghiên cứu

Đối tượng nghiên cứu: (1) Thuật toán tìm kiếm: Phương pháp và kỹ thuật phát triển thuật toán để tìm kiếm và phân loại (phân cụm và phân lớp) ảnh dựa trên thông tin ngữ nghĩa thay vì chỉ dựa vào các đặc trưng hình ảnh như màu sắc, hình dạng. Đồng thời nghiên cứu một số phương pháp học máy nhằm cải tiến hiệu quả tìm kiếm ảnh ; (2) Ngữ nghĩa trong hình ảnh: Các yếu tố ngữ nghĩa cần được hiểu và xác định trong ảnh để hỗ trợ việc tìm kiếm

chính xác; (3) Dữ liệu hình ảnh: Các bộ dữ liệu hình ảnh có chứa thông tin ngữ nghĩa cần thiết để huấn luyện và đánh giá thuật toán.

Phạm vi nghiên cứu: (1) Phạm vi ngữ nghĩa: Nhận diện một số loại ngữ nghĩa nhất định như vật thể (ví dụ: xe, người, động vật), hành động (ví dụ: chạy, nhảy), hoặc bối cảnh (ví dụ: ngoài trời, trong nhà) bằng cách dùng ontology và ngôn ngữ truy vấn SPARQL; (2) Không gian: Tập trung vào một hoặc một số bộ dữ liệu hình ảnh có chứa thông tin ngữ nghĩa cần thiết để huấn luyện và đánh giá thuật toán như bộ dữ liệu chuẩn như Wang, MS-COCO, ImageCLEF); (3) Giới hạn phương pháp: Tập trung vào phương pháp dựa trên ontology để liên kết ngữ nghĩa và hình ảnh.

6. Các đóng góp của luận án

Phát triển thuật toán tìm kiếm ảnh theo ngữ nghĩa dựa trên các đặc trưng thị giác của ảnh dựa trên cấu trúc dữ liệu GP-Tree nhằm nâng cao độ chính xác tìm kiếm ảnh, gồm:

- (1) Xây dựng cấu trúc dữ liệu phân cụm phân cấp GP-Tree nhằm tổ chức lưu trữ các véc-tơ đặc trưng của ảnh
- (2) Phát triển cấu trúc GP-Tree dựa trên các thuật toán học có giám sát và bán giám sát nhằm tăng hiệu quả tìm kiếm ảnh
- (3) Xây dựng hệ tìm kiếm ảnh theo ngữ nghĩa dựa trên cấu trúc GP-Tree và ontology nhằm minh chứng hiệu quả tìm kiếm ảnh của các phương pháp đề xuất

7. Nội dung và bố cục của luận án

Cấu trúc của luận án bao gồm:

- **Chương 1:** Trình bày tổng quan về bài toán tìm kiếm ảnh, với hai hướng chính là tìm kiếm ảnh theo nội dung và tìm kiếm ảnh theo ngữ nghĩa. Các công trình nghiên cứu liên quan đã được khảo sát và phân tích nhằm xác định thách thức và hạn chế trong các phương pháp hiện có, từ đó đưa ra định hướng nghiên cứu cụ thể của luận án để khắc phục

những hạn chế này. Ngoài ra, phần này cũng trình bày chi tiết các phương pháp tổ chức thực nghiệm, bao gồm việc thiết lập môi trường, lựa chọn và sử dụng tập dữ liệu, cùng các tiêu chí đánh giá hiệu suất tìm kiếm.

- **Chương 2:** Trình bày các nghiên cứu liên quan đến việc sử dụng cấu trúc cây để lưu trữ và lập chỉ mục cho tập dữ liệu ảnh. Cấu trúc cây phân cụm GP-Tree được mô tả chi tiết, bao gồm các thao tác thêm, sửa, tách, và xóa phần tử. Một mô hình tìm kiếm ảnh theo ngữ nghĩa trên cây GP-Tree dựa trên ontology đã được đề xuất, với mục tiêu cải thiện hiệu suất và độ chính xác của việc tìm kiếm. Hệ thống tìm kiếm ảnh này đã được thử nghiệm trên các bộ dữ liệu phổ biến như Wang, MS-COCO và ImageCLEF, nhằm đánh giá kết quả và tính hiệu quả của mô hình đề xuất.
- **Chương 3:** Trình bày các phương pháp cải tiến cấu trúc cây phân cụm GP-Tree nhằm nâng cao hiệu quả tìm kiếm ảnh. Cụ thể, các phương pháp như đồ thị cụm Graph-GPTree và mạng kết hợp SgGP-Tree được giới thiệu để cải thiện khả năng lưu trữ và tìm kiếm các phần tử tương tự. Ngoài ra, phương pháp tìm kiếm ảnh theo ngữ nghĩa dựa trên ontology cũng được thảo luận, với cấu trúc SgGP-Tree được sử dụng để phân lớp đối tượng trên ảnh một cách chính xác hơn. Một mô hình tìm kiếm ảnh theo ngữ nghĩa, kết hợp giữa ontology và cấu trúc SgGP-Tree, đã được đề xuất và thử nghiệm trên các bộ dữ liệu phổ biến như Wang, MS-COCO, và ImageCLEF nhằm đánh giá hiệu quả của mô hình.
- **Kết luận và hướng phát triển:** Trình bày những kết quả đạt được định hướng phát triển tiếp theo của luận án.
- **Danh mục công trình của tác giả:** Liệt kê các công trình mà tác giả đã công bố trong quá trình thực hiện luận án.
- **Tài liệu tham khảo:** Liệt kê các tài liệu mà luận án đã tham khảo.

CHƯƠNG 1. TỔNG QUAN TÌM KIẾM ẢNH

1.1. Tổng quan về tìm kiếm ảnh

Bài toán tìm kiếm ảnh trình bày trong luận án được xác định là tìm ra tập ảnh có độ tương tự gần nhất với ảnh đầu vào dựa trên độ đo tương tự giữa các ảnh [6].

1.2. Các đặc trưng phổ biến trong tìm kiếm ảnh

Một đặc trưng được định nghĩa là việc xác định một thuộc tính trực quan của hình ảnh [7]. Nhìn chung, các đặc trưng hình ảnh có thể là toàn cục hoặc cục bộ [8]

1.3. Các công trình nghiên cứu liên quan về tìm kiếm ảnh

1.3.1. Tìm kiếm ảnh dựa trên văn bản

Phương pháp dựa trên văn bản là một phương pháp tìm kiếm đơn giản theo từ khóa truyền thống. Các hình ảnh được lập chỉ mục theo nội dung, như chú thích của hình ảnh; tên tệp, tiêu đề của trang web và thẻ thay thế.... và được lưu trữ trong cơ sở dữ liệu

1.3.2. Tìm kiếm ảnh theo nội dung

Trong hệ thống tìm kiếm ảnh dựa trên nội dung, hiệu quả của quá trình tính toán phụ thuộc vào khả năng trích xuất đặc trưng cấp thấp và đo độ tương tự. Trong phương pháp tìm kiếm ảnh dựa trên nội dung (CBIR), các đặc trưng trực quan cấp thấp như: màu sắc, hình dạng, kết cấu và bố cục không gian,...

1.3.3. Tìm kiếm ảnh dựa ngữ nghĩa

Trong lĩnh vực tìm kiếm ảnh, thách thức chính đó là việc chuyển đổi ảnh sang các đặc trưng cấp thấp mà máy tính có thể tính toán được và liên kết chúng với các khái niệm bậc cao để giảm khoảng cách ngữ nghĩa. Nhiều phương pháp tìm kiếm ảnh theo ngữ nghĩa (SBIR) đã được đề xuất để giảm khoảng cách ngữ nghĩa [9] như: tìm kiếm ảnh theo ngữ nghĩa dựa trên kỹ thuật học máy, tìm kiếm ảnh theo ngữ nghĩa dựa trên ontology.

1.4. Các phương pháp tổ chức thực nghiệm và đánh giá

Các mô hình đề xuất được xây dựng và thực nghiệm dựa trên nền tảng dotNet Framework 4.5 và Python 3.7

Các bộ dữ liệu được sử dụng trong thực nghiệm của luận án là các bộ dữ liệu phổ biến và được sử dụng rộng rãi trong các công trình nghiên cứu về tìm kiếm ảnh, gồm: Wang, ImageCLEF, MS-COCO.

Các kết quả cần được đánh giá hiệu suất gồm: Kết quả phân lớp hình ảnh và độ chính xác tìm kiếm ảnh. Các giá trị đánh giá hiệu suất của phân lớp và tìm kiếm ảnh được sử dụng trong luận án bao gồm P (precision), độ phủ R (recall) và độ dung hoà F_m (F-measure).

1.5. Tiểu kết chương

Chương này cung cấp một cái nhìn tổng quan chi tiết về các phương pháp tìm kiếm ảnh hiện đại, bao gồm hai hướng tiếp cận chính: tìm kiếm ảnh theo nội dung và tìm kiếm ảnh theo ngữ nghĩa.

Chương này cũng trình bày cụ thể các phương pháp tổ chức thực nghiệm, bao gồm việc thiết lập môi trường thực nghiệm, lựa chọn và chuẩn bị tập dữ liệu thực nghiệm, cũng như các thước đo đánh giá hiệu suất của các phương pháp tìm kiếm.

Chương tiếp theo sẽ trình bày một cấu trúc dữ liệu mới, đó là cây phân cụm phân cấp, được đề xuất nhằm tối ưu hóa việc lưu trữ và lập chỉ mục tập dữ liệu ảnh. Cấu trúc này không chỉ hỗ trợ quá trình truy xuất nhanh chóng mà còn giúp tăng cường hiệu quả tìm kiếm ảnh theo ngữ nghĩa, cho phép hệ thống có thể xử lý những tập dữ liệu ảnh lớn nhanh và hiệu quả.

CHƯƠNG 2. CẤU TRÚC GP-TREE ĐỂ TÌM KIẾM ẢNH THEO NGỮ NGHĨA

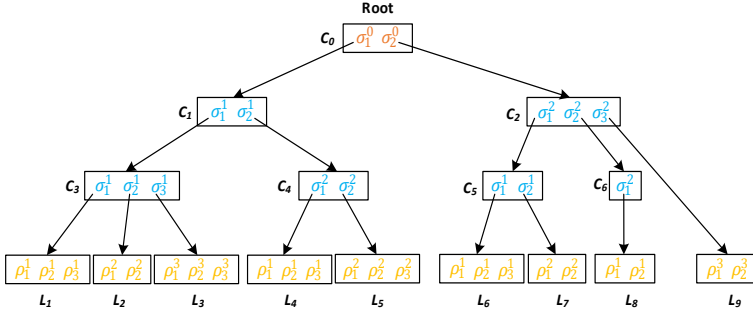
Chương này trình bày các các nghiên cứu liên quan đến cấu trúc cây nhằm lưu trữ và lập chỉ mục cho tập dữ liệu ảnh. Mô tả cấu trúc cây phân cụm GP-Tree và các thao tác thêm, sửa, tách và xóa phần tử trên cây. Một mô hình kiểm ảnh theo ngữ nghĩa trên cây GP-Tree dựa trên ontology được đề xuất và hệ thống tìm kiếm ảnh được thực nghiệm trên các bộ dữ liệu ảnh phổ biến như Wang, MS-COCO và ImageCLEF nhằm đánh giá kết quả của mô hình đề xuất. Nội dung của chương này có liên quan trực tiếp đến hai công trình đã công bố là [CT4] và [CT5]; đồng thời cũng liên quan gián tiếp đến các công trình [CT1], [CT2], [CT3].

2.1. Giới thiệu

Chương này trình bày một số cấu trúc về cây phân cụm và các nghiên cứu liên quan đến cấu trúc cây; kỹ thuật thiết kế cấu trúc cây phân cụm cân bằng GP-Tree để lưu trữ dữ liệu ảnh lớn và các thao tác thêm, sửa, tách và xóa trên cây; đề xuất mô hình, thuật toán và thực nghiệm tìm kiếm ảnh dựa trên ngữ nghĩa trên cây GP-Tree với các tập dữ liệu ảnh đã đề xuất.

2.2. Cấu trúc dữ liệu GP-Tree

Trên cơ sở cấu trúc cây đa nhánh và phương pháp phân cụm K-Means, cấu trúc cây GP-Tree được xây dựng bằng cách tách nút lá thành hai nút nếu như số lượng phần tử tại nút lá đó vượt trên giá trị M cho trước, đồng thời tại mỗi một nút trong các nút con có thể được tạo thành dựa trên một ngưỡng θ , ngưỡng này nhằm đánh giá độ tương tự của dữ liệu. Do đó, cây GP-Tree tăng trưởng theo hướng lá và phát triển thành cây đa nhánh mà mỗi một nút lá là một cụm dữ liệu gồm các phần tử tương tự nhau. Các phần tử dữ liệu là các véc-tơ đặc trưng cho mỗi hình ảnh và lần lượt được lưu trữ trên cây GP-Tree, từ đó các thao tác trên cây được thực hiện. Cấu trúc của cây phân cụm phân cấp GP-Tree được minh họa trong **Hình 2.1**

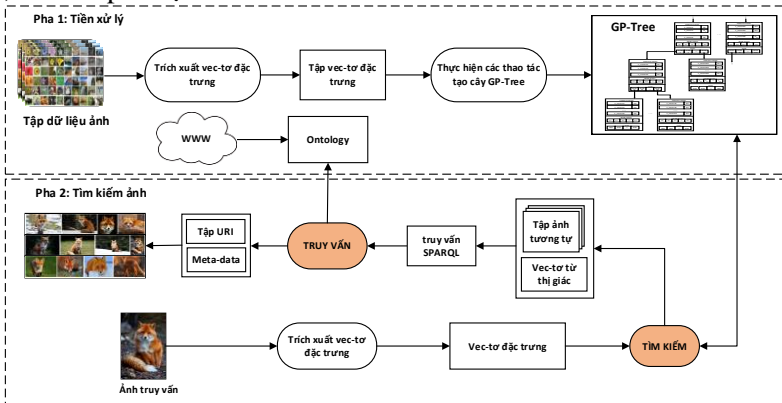


Hình 2.1. Cây phân cụm phân cấp GP-Tree gồm 3 mức

2.3. Tìm kiếm ảnh dựa trên cấu trúc GP-Tree

2.3.1. Hệ tìm kiếm ảnh dựa trên cây GP-Tree

Hệ tìm kiếm ảnh bao gồm 2 pha: (1) pha tiền xử lý thực hiện phân đoạn, trích xuất đặc trưng của ảnh trong tập dữ liệu, lấy các phân lớp khái niệm và tổ chức lưu trữ trên cây GP-Tree; (2) Pha truy vấn có nhiệm vụ tìm kiếm ảnh tương tự theo nội dung trên cây GP-Tree và trích xuất ngữ nghĩa của hình ảnh. Hình 2.4 là mô hình hệ tìm kiếm ảnh theo ngữ nghĩa dựa trên cây GP-Tree, với hai pha cụ thể như sau:



Hình 2.4 Mô hình hệ tìm kiếm ảnh dựa trên GP-Tree (GP-SBIR)

2.3.2. Thực nghiệm và đánh giá hệ tìm kiếm GP-SBIR

Bảng 2.1. Kết quả thực nghiệm cây GP-Tree

Tên tập ảnh	Số lượng ảnh	Tham số thực nghiệm		Thời gian thực nghiệm (giây)	Số cụm nút lá	Số cụm nút trong	Số mẫu lấy P@	Tỉ lệ P@
		M	N					
WANG	10.800	100	40	17.839,47	218	32	2.240	20%
ImageCLEF	20.000	150	50	32.173,95	432	67	6.000	20%
MS-COCO	163.957	300	70	158.765,84	782	314	44.188	20%

Bảng 2.2. Hiệu suất truy vấn ảnh của hệ GP-SBIR trên các tập dữ liệu thử nghiệm.

Tập dữ liệu ảnh	Chỉ số đánh giá			
	Độ chính xác	Độ phủ	Độ dung hòa	Thời gian tìm kiếm trung bình (ms)
WANG	0.6780	0.684	0.6810	98.75
ImageCLEF	0.6802	0.775	0.7245	132.09
MS-COCO	0.7170	0.724	0.7205	217.65

Semantic-based Image Retrieval

0.0837037037037037, 0.0, 0.1566666666666667, 0.0, 0.0, 0.0162962962962963, 0.00271604938271605, 0.1311111111111111, 0.0, 0.0465432098765432, 0.444320987654321, 0.0, 0.0037037037037037, 0.0583950617283951, 0.0565432098765432, 0.0, 0.0, 0.489490196078431, 0.518588235294118, 0.75243137254902, 0.751803921568627, 0.240862745098039, 0.262509803921569, 0.725764705882353, 0.541254901960784, 0.0658923529411785, 0.708765432098765, 0.471047641162254, 0.453159952010527, 0.553951035701544, 0.594472693032015, 0.0, 0.240862745098039, 0.262509803921569, 0.541254901960784, 0.489490196078431, 0.518588235294118, 0.75243137254902, 0.751803921568627, 0.0, 0.725764705882353, 0.0658923529411785, 0.0, 0.100493827160494, 0.902839506172839, 0.5625,

elephant mountain bus beach meal

elephant

COREL Wang Stanford Dogs ImageCLEF

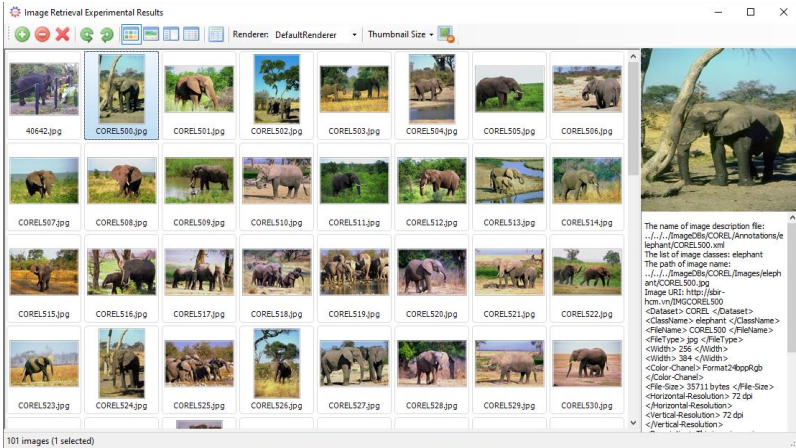
Load Image Classification on GP-Tree Classification on Graph GP-Tree AND Query objects UNION Query objects Refresh Close

AND Query images UNION Query images **Ontology Retrieval**

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
 PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
 PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
 PREFIX owl: <http://www.w3.org/2002/07/owl#>
 PREFIX xml: <http://www.w3.org/XML/1998/namespace>
 PREFIX sbrir: <http://sbrir-hcm.vn/>
 SELECT DISTINCT ?imgName
 WHERE { (?IMG sbrir:imgName ?imgName. ?IMG sbrir:opELEPHANT
 sbrir:inELEPHANT sbrir:inELEPHANT rdf:type owl:NamedIndividual .
 sbrir:inELEPHANT rdf:type sbrir:ELEPHANT.) }

40642.jpg COREL500.jpg COREL501.jpg COREL502.jpg COREL503.jpg COREL504.jpg
 COREL505.jpg COREL506.jpg COREL507.jpg COREL508.jpg COREL509.jpg COREL510.jpg
 COREL511.jpg COREL512.jpg COREL513.jpg COREL514.jpg COREL515.jpg COREL516.jpg
 COREL517.jpg COREL518.jpg COREL519.jpg COREL520.jpg COREL521.jpg COREL522.jpg
 COREL523.jpg COREL524.jpg COREL525.jpg COREL526.jpg COREL527.jpg COREL528.jpg
 COREL529.jpg COREL530.jpg COREL531.jpg COREL532.jpg COREL533.jpg COREL534.jpg
 COREL535.jpg COREL536.jpg COREL537.jpg COREL538.jpg COREL539.jpg COREL540.jpg
 COREL541.jpg COREL542.jpg COREL543.jpg COREL544.jpg COREL545.jpg COREL546.jpg
 COREL547.jpg COREL548.jpg COREL549.jpg COREL550.jpg COREL551.jpg COREL552.jpg
 COREL553.jpg COREL554.jpg COREL555.jpg COREL556.jpg COREL557.jpg COREL558.jpg
 COREL559.jpg COREL560.jpg COREL561.jpg COREL562.jpg COREL563.jpg COREL564.jpg
 COREL565.jpg COREL566.jpg COREL567.jpg COREL568.jpg COREL569.jpg COREL570.jpg

Hình 2.18. Hệ tìm kiếm ảnh GP-SBIR



Hình 2.19. Kết quả tập ảnh tương tự của ảnh truy vấn trên hệ GP-SBIR

Hệ tìm kiếm ảnh dựa trên cấu trúc GP-Tree (GP-SBIR) được minh họa bởi **Hình 2.18**. Với mỗi ảnh đầu vào I_q (Load Image) được chọn theo từng bộ ảnh (Wang, ImageCLEF, MS-COCO), trích xuất véc-tơ đặc trưng được tìm kiếm trên GP-Tree để trích xuất tập ảnh tương tự theo nội dung. Kết quả tìm kiếm tập ảnh tương tự theo nội dung SI của ảnh đầu vào I_q dựa trên GP-Tree là quá trình trích xuất tập từ thị giác. Tập từ thị giác được xây dựng dựa trên những ảnh cùng phân lớp thuộc tập có tần số xuất hiện nhiều nhất. Từ đó, câu truy vấn SPARQL được xây dựng (Create SPARQL) để tìm tập ảnh tương tự theo ngữ nghĩa dựa trên ontology đã xây dựng (Load Ontology). Kết quả tập ảnh tương tự theo ngữ nghĩa với ảnh đầu vào được minh họa bởi **Hình 2.19**.

2.4. Tiểu kết chương

Cây phân cụm GP-Tree có thể lưu trữ được dữ liệu lớn, hiệu quả cho bài toán tìm kiếm ảnh. Cây GP-Tree thực hiện các phương pháp phân cụm phân hoạch và phân cụm phân cấp nên có thể tìm kiếm nhanh theo nhánh có độ đo tương tự nhất để tìm được các phần tử tương tự nhau tại cụm nút lá, do đó GP-Tree

có thời gian tìm kiếm nhanh và độ chính xác khá cao. Tuy nhiên, nhược điểm chính của GP-Tree là mỗi lần tách nút, các phần tử tương tự nhau có thể bị tách thành các nút khác nhau, trong trường hợp xấu nhất, các nút này có thể nằm trên các nhánh riêng biệt. Do đó, quá trình tìm kiếm ảnh trên cây GP-Tree có thể sẽ bỏ sót các phần tử tương tự đã bị chuyển nhánh. Điều này làm ảnh hưởng đến hiệu suất tìm kiếm trên GP-Tree. Vì vậy, các phương pháp cải tiến cây đã được đề xuất trong chương tiếp theo nhằm nâng cao độ chính xác trong tìm kiếm ảnh trên cây GP-Tree. Ngoài ra, một mô hình tìm kiếm ảnh theo tiếp cận ngữ nghĩa dựa trên ontology cũng được đề xuất nhằm tăng hiệu suất tìm kiếm ảnh và trích xuất ngữ nghĩa cho ảnh truy vấn. Vấn đề tìm kiếm ảnh dựa trên ngữ nghĩa thông qua cấu trúc GP-Tree được trình bày trong chương tiếp theo.

CHƯƠNG 3. CẤU TRÚC SGGP-TREE ĐỂ TÌM KIẾM ẢNH THEO NGỮ NGHĨA

Chương này trình bày các phương pháp cải tiến cấu trúc cây phân cụm GP-Tree nhằm nâng cao hiệu quả tìm kiếm ảnh. Cụ thể, các phương pháp như đồ thị cụm Graph-GPTree và mạng kết hợp SgGP-Tree được giới thiệu để cải thiện khả năng lưu trữ và tìm kiếm các phần tử tương tự. Ngoài ra, phương pháp tìm kiếm ảnh theo ngữ nghĩa dựa trên ontology cũng được thảo luận, với cấu trúc SgGP-Tree được sử dụng để phân lớp đối tượng trên ảnh một cách chính xác hơn. Một mô hình tìm kiếm ảnh theo ngữ nghĩa, kết hợp giữa ontology và cấu trúc SgGP-Tree, đã được đề xuất và thử nghiệm trên các bộ dữ liệu phổ biến như WANG, MS-COCO, và ImageCLEF nhằm đánh giá hiệu quả của mô hình. Nội dung cơ bản của chương này được công bố trong các công trình [CT1], [CT2], [CT3], [CT6].

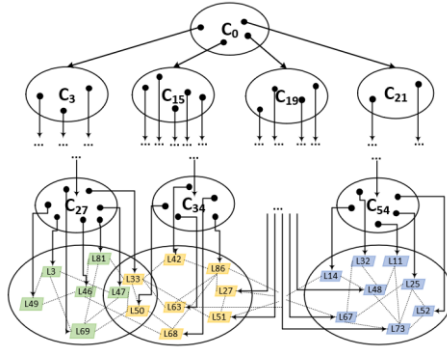
3.1. Giới thiệu

Chương này trình bày các phương pháp cải tiến cây phân cụm GP-Tree nhằm nâng cao hiệu quả tìm kiếm ảnh: Đồ thị cụm Graph-GPTree và mạng kết hợp SgGP-Tree; trình bày phương pháp tìm kiếm ảnh theo ngữ nghĩa dựa trực tiếp trên ontology, cấu trúc SgGP-Tree được sử dụng để phân lớp các đối tượng trên ảnh; phát triển Ontology trên các bộ dữ liệu ảnh đa đối tượng. Một mô hình tìm kiếm ảnh theo tiếp cận ngữ nghĩa dựa trên ontology kết hợp với cấu trúc SgGP-Tree được đề xuất và thực nghiệm để đánh giá kết quả.

3.2. Đồ thị cụm Graph-GPTree

3.2.1. Cấu trúc Graph-GPTree

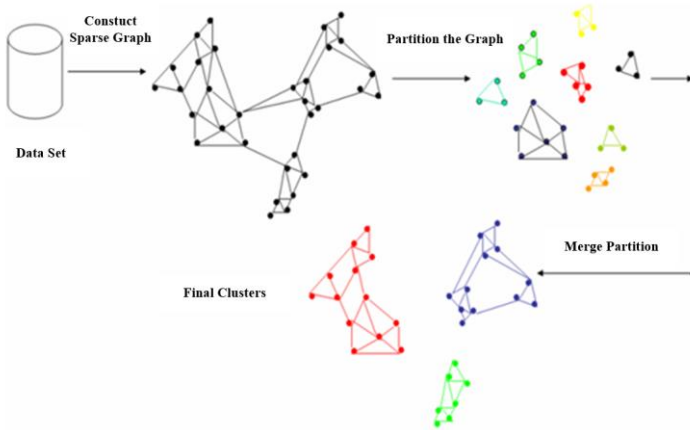
Graph-GPTree được tạo dựa trên các thao tác trên đồ thị thừa tập các nút lá thu được của GP-Tree. Trong đó, các đỉnh biểu thị các nút lá và các cạnh có trọng số biểu thị mức độ tương đồng giữa chúng. Đồ thị thừa được tạo trong quá trình tạo cây GP-Tree khi mỗi lần tách nút lá hệ thống tiến hành đánh dấu các mức lân cận của các nút lá vừa mới tách (**Hình 3.2**)



Hình 3.2. Đồ thị thưa được tạo phải tập nút lá cây GP-Tree

3.2.2. Quá trình tạo Graph-GPTree

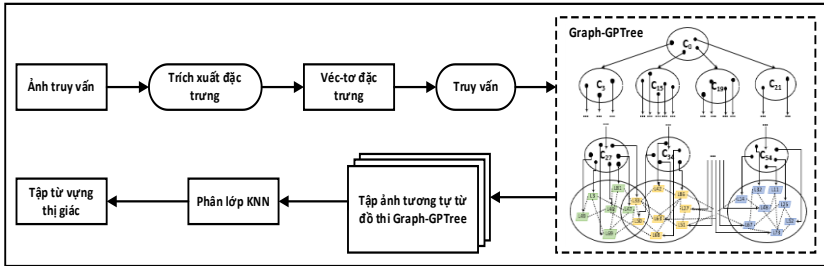
Thuật toán phân cụm đồ thị GraphGP-Tree để tìm các cụm trong một tập dữ liệu được mô tả tổng quát như trong **Hình 3.3**. Thuật toán thực hiện trên một đồ thị thưa trong đó các nút biểu thị các phần tử dữ liệu và các cạnh có trọng số biểu thị sự tương đồng giữa các phần tử dữ liệu. Việc biểu diễn tập dữ liệu bằng đồ thị thưa này cho phép thuật toán phân cụm mở rộng quy mô thành các tập dữ liệu lớn.



Hình 3.3. Tổng quan về cách tiếp cận tổng thể của thuật toán tạo cây GP-Tree

3.2.3. Mô hình tìm kiếm ảnh trên đồ thị cụm lân cận Graph-GPTree

Mô hình tìm kiếm ảnh trên đồ thị cụm lân cận Graph-GPTree được mô tả như **Hình 3.4**, gồm các bước sau: Trích xuất vector đặc trưng cho ảnh đầu vào; So sánh vector đặc trưng với cơ sở dữ liệu các đặc trưng trên cây GP-Tree để chọn nhánh tương tự nhất và xác định nút lá phù hợp; Từ nút lá phù hợp được chọn trên cây GP-Tree, dựa trên đồ thị cụm Graph-GP-Tree, tìm ra tập các láng giềng của nút lá và kết quả là tập các ảnh tương tự được sắp xếp theo độ đo tăng dần; Sử dụng thuật toán k-NN để phân lớp cho tập ảnh tương tự và tìm kiếm từ vựng thị giác để giải thích ngữ nghĩa cho hình ảnh truy vấn.



Hình 3.4. Mô hình tìm kiếm ảnh trên đồ thị cụm lân cận Graph-GPTree

3.3. Mạng kết hợp SgGP-Tree

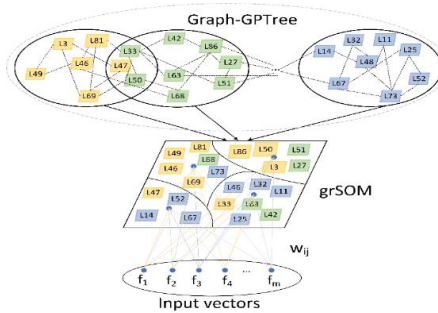
3.3.1. Cấu trúc SgGP-Tree

Cấu trúc SgGP-Tree là sự kết hợp giữa cây GP-Tree, Graph-GPTree và mạng SOM. Trong mạng SOM, sự điều chỉnh các trọng số trong quá trình huấn luyện sẽ làm cho SOM đạt được sự gom cụm tốt nhất [10]. Tuy nhiên, quá trình điều chỉnh trọng số tốn nhiều chi phí đối với tập dữ liệu ảnh đầu vào lớn và cách khởi tạo trọng số ngẫu nhiên có thể sẽ tạo ra các bản đồ hoàn toàn khác nhau. Ngoài ra, việc bổ sung một dữ liệu mới vào bản đồ sau khi huấn luyện SOM sẽ phân loại sai dữ liệu đầu vào do SOM có tính chất tĩnh; lúc này, SOM phải huấn luyện lại từ đầu. Mô hình kết hợp giữa cây GP-Tree, đồ thị láng giềng

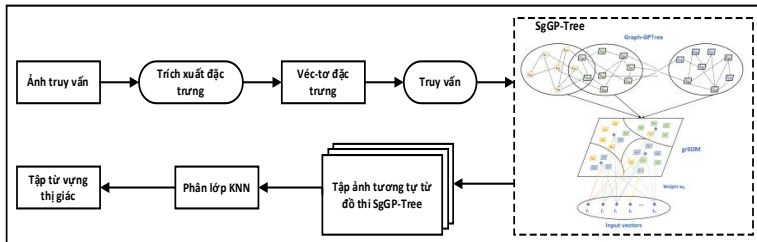
Graph-GPTree và mạng grSOM, gọi là SgGP-Tree, được mô tả như trong **Hình 3.5**.

3.3.2. Mô hình tìm kiếm ảnh trên mạng kết hợp SgGP-Tree

Mô hình tìm kiếm ảnh trên mạng kết hợp SgGP-Tree được gọi là SBIR-SgGP. SgGP-Tree là một cấu trúc được tạo ra bằng cách kết hợp cây GP-Tree, đồ thị cụm lân cận Graph-GPTree và mạng SOM (**Figure 3.5**). Quá trình tiền xử lý bao gồm trích xuất đặc trưng từ tập ảnh và lưu trữ chúng trên cấu trúc SgGP-Tree. Sau đó, trong quá trình tìm kiếm ảnh, SgGP-Tree được sử dụng để tìm kiếm tập ảnh tương tự và tập từ vựng thị giác. Trong mô hình tìm kiếm ảnh này, khối trích xuất vec-tơ đặc trưng được sử dụng lại từ các mô hình trước đó. Một khối cải tiến của mô hình là việc cải tiến cây phân cụm GP-Tree bằng cách kết hợp đồ thị cụm lân cận Graph-GPTree và mạng SOM để tạo ra SgGP-Tree (**Figure 3.6**).



Hình 3.5. Mô hình kết hợp SgGP-Tree

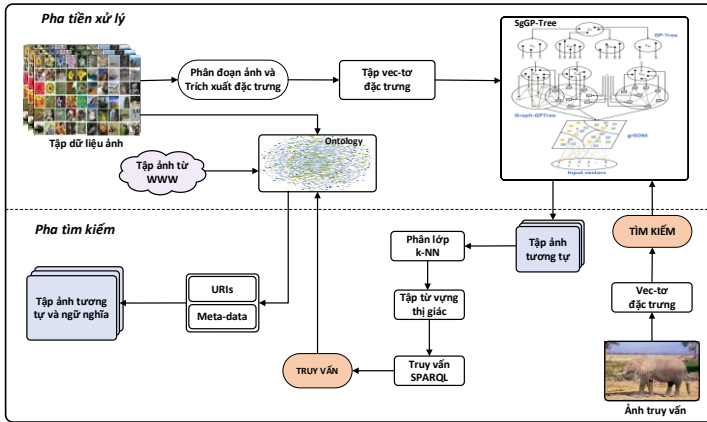


Hình 3.6. Mô trình tìm kiếm ảnh trên SgGP-Tree (SBIR-SgGP)

3.4. Hệ tìm kiếm ảnh theo ngữ nghĩa dựa trên ontology

3.4.1. Mô hình tìm kiếm ảnh dựa trên ontology

Dựa trên các phương pháp tìm kiếm ảnh theo ngữ nghĩa trên ontology được đề xuất, một hệ tìm kiếm theo ngữ nghĩa dựa trên ontology được xây dựng bao gồm 2 pha được mô tả trong **Hình 3.7**. Trong mô hình này, phương pháp trích xuất đặc trưng, tìm kiếm ảnh theo nội dung được kế thừa từ các chương trước của luận án. Phương pháp xây dựng ontology, tìm kiếm ảnh trên ontology là đóng góp chính của mô hình này.

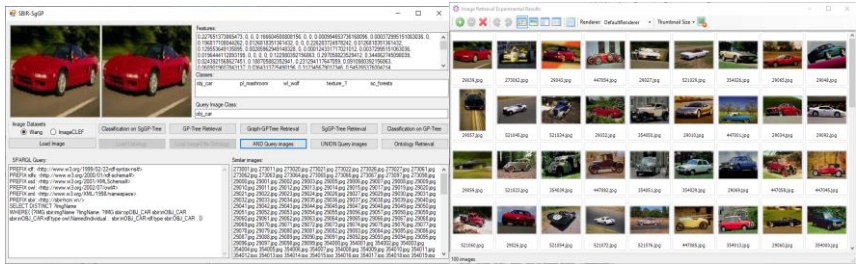


Hình 3.7. Mô hình hệ tìm kiếm SBIR-GP

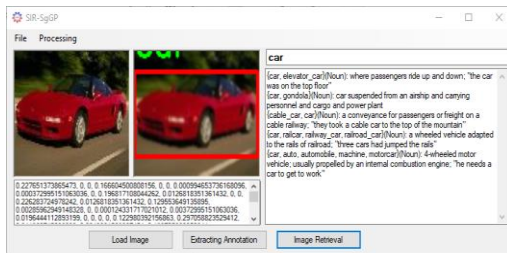
3.4.2. Thục nghiệm và đánh giá hệ tìm kiếm ảnh SBIR-GP

Hệ thống truy xuất hình ảnh SBIR-GP được thiết kế để thực hiện các truy vấn hình ảnh dựa trên ngữ nghĩa bằng cách sử dụng SgGP-Tree và ontology. Khi được cung cấp một hình ảnh đầu vào, hệ thống SBIR-GP trước tiên sẽ trích xuất các vectơ đặc trưng và truy xuất các hình ảnh tương tự dựa trên nội dung của chúng một cách tuần tự bằng cách sử dụng SgGP-Tree. Quá trình này dẫn đến một bộ sưu tập các hình ảnh tương tự. **Hình 3.8** là giao diện của hệ OnSBIR với đầu vào là ảnh. Kết quả là danh sách các ảnh có cùng ngữ nghĩa

với ảnh đầu vào và các mô tả metadata của tập ảnh. Đồng thời, khái niệm ngữ nghĩa của từ vựng phân lớp được trích xuất từ tập từ điển đồng nghĩa của ontology. **Hình 3.9.** là một ví dụ về khái niệm phân lớp của từ điển ontology.



Hình 3.8. Một kết quả của hệ tìm kiếm SBIR-GP từ ảnh đầu vào



Hình 3.9. Khái niệm ngữ nghĩa cho lớp

Các bộ dữ liệu hình ảnh được sử dụng cho các thử nghiệm, bao gồm các bộ dữ liệu WANG, MS-COCO và ImageCLEF. Các giá trị hiệu suất trung bình và thời gian tìm kiếm của bộ dữ liệu thử nghiệm được trình bày trong **Bảng 3.1, Bảng 3.2 và Bảng 3.3**

Bảng 3.1. Hiệu suất tìm kiếm ảnh trên bộ dữ liệu ảnh WANG

Phương pháp	Độ chính xác	Độ phủ	Độ dung hòa	Thời gian tìm kiếm trung bình (ms)
GP-Tree	0.6780	0.6840	0.6810	39.75
Graph-GPTree	0.7665	0.6677	0.7137	202.79
SBIR-GP	0.8004	0.7040	0.7491	696.19

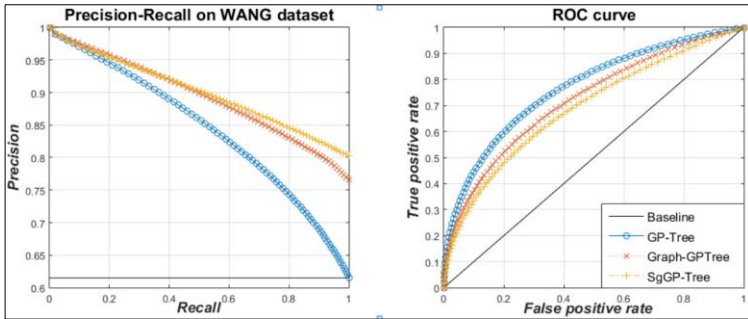
Bảng 3.2. Hiệu suất tìm kiếm ảnh trên bộ dữ liệu ảnh ImageCLEF

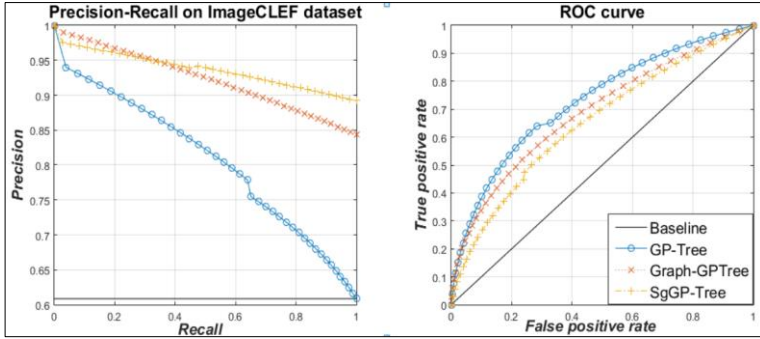
Phương pháp	Độ chính xác	Độ phủ	Độ dung hòa	Thời gian tìm kiếm trung bình (ms)
GP-Tree	0.6802	0.7750	0.7245	44.09
Graph-GPTree	0.8168	0.7637	0.7894	239.29
SBIR-GP	0.8926	0.8764	0.8844	868.51

Bảng 3.3. Hiệu suất tìm kiếm ảnh trên bộ dữ liệu ảnh MS-COCO

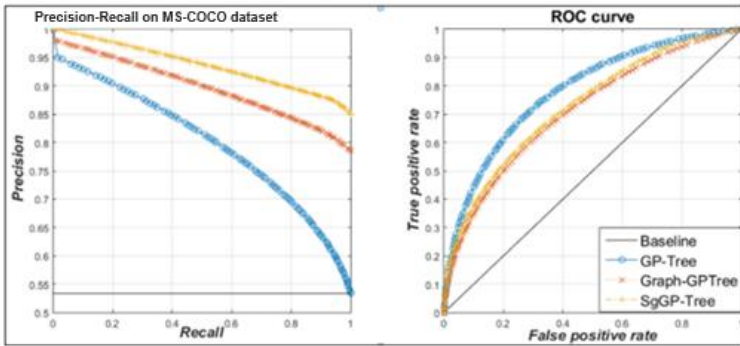
Phương pháp	Độ chính xác	Độ phủ	Độ dung hòa	Thời gian tìm kiếm trung bình (ms)
GP-Tree	0.717	0.724	0.7205	102.32
Graph-GPTree	0.873	0.764	0.815	198.47
SBIR-GP	0.875	0.724	0.783	265.45

Từ các bảng trên, có thể thấy rằng việc cải thiện GP-Tree mang lại hiệu suất tìm kiếm chính xác tốt hơn cho các bộ dữ liệu Wang, ImageCLEF và MS-COCO. Biểu đồ lân cận Graph-GPTree có hiệu suất tốt hơn GP-Tree nhưng thấp hơn SgGP-Tree. Tuy nhiên, thời gian tìm kiếm GP-Tree nhanh hơn Graph-GPTree và SBIR-GP.

**Hình 3.10.** Hiệu suất tìm kiếm ảnh trên GP-Tree, Graph-GPTree và SBIR-GP (SgGP-Tree) trên tập dữ liệu ảnh WANG.



Hình 3.11. Hiệu suất tìm kiếm ảnh trên GP-Tree, Graph-GP-Tree và SBIR-GP (SgGP-Tree) trên tập dữ liệu ảnh ImageCLEF.



Hình 3.12. Hiệu suất tìm kiếm ảnh trên GP-Tree, Graph-GP-Tree và SBIR-GP (SgGP-Tree) trên tập dữ liệu ảnh MS-COCO.

Dựa trên dữ liệu thực nghiệm, các đường cong PR và đường cong ROC được thực hiện để đánh giá độ chính xác của hệ thống tìm kiếm SBIR-GP (**Hình 3.10**, **Hình 3.11** và **Hình 3.12**).

Để đánh giá độ chính xác và hiệu quả của hệ thống tìm kiếm ảnh SBIR-GP, chúng tôi đã so sánh hiệu suất của hệ thống này với các công trình nghiên cứu khác trên cùng một tập dữ liệu ảnh. Các kết quả trong **Bảng 3.4**, **Bảng 3.5** và **Bảng 3.6** so sánh giá trị độ chính xác trung bình của các phương pháp tìm kiếm trên bộ dữ liệu WANG, ImageCLEF và MS-COCO.

Từ các bảng trên cho thấy phương pháp chúng tôi đề xuất đạt được độ chính xác cao hơn so với các phương pháp tra cứu khác trên cùng một tập dữ liệu. Điều này cho thấy phương pháp của chúng tôi có thể trích xuất hiệu quả các đặc trưng để phân biệt các đặc điểm chi tiết của đối tượng trong ảnh. Phương pháp đề xuất của chúng tôi có hiệu quả trong việc giải quyết bài toán truy vấn và phân tích ngữ nghĩa đối với ảnh có một và nhiều đối tượng.

Bảng 3.4. So sánh các phương pháp tìm kiếm ảnh trên bộ dữ liệu ảnh WANG

Phương pháp	Độ chính xác trung bình
K. Kanwal et al. (2020) [11]	0.5067
H. Zeng et al. (2021) [12]	0.6600
O. Sikha and K. Soman (2021) [13]	0.8030
S. Dhingra and P. Bansal (2021) [14]	0.6000
A. Ouni et al. (2022) [15]	0.7800
Graph-GPtree	0.7665
SBIR-GP (SgGP-Tree)	0.8004

Bảng 3.5. So sánh các phương pháp tìm kiếm ảnh trên bộ dữ liệu ảnh ImageCLEF

Phương pháp	Độ chính xác trung bình
A. Yang et al. (2019) [16]	0.8030
Y. Qiang et al. (2020) [17]	0.6670
X. Yue et al. (2021) [18]	0.7140
N. T. U. Nhi et al. (2022) [19]	0.6510
X. Wang et al. (2023) [20]	0.7270
Graph-GPtree	0.8168
SgGP-Tree (SBIR-GP)	0.8926

Bảng 3.6. So sánh các phương pháp tìm kiếm ảnh trên bộ dữ liệu ảnh MS-COCO

Phương pháp	Độ chính xác trung bình
Y. Cao, 2018 [21]	0.8576
Y. Xie, 2020 [22]	0.8628
Wen Gu, 2019 [23]	0.8350
Graph-GPtree	0.8730
SgGP-Tree (SBIR-GP)	0.8753

3.5. Tiêu kết chương

Trong chương này, các phương pháp nâng cao hiệu suất tra cứu ảnh trên GP-Tree được đề xuất. Đầu tiên, một mô hình kết hợp biểu đồ lân cận với GP-Tree, được gọi là Graph-GPTree, được tạo ra để kết nối các phần tử tương tự được phân nhánh trong quá trình phân tách nút trên GP-Tree. Tiếp theo, một mô hình kết hợp GrSOM và Graph-GPTree, được gọi là SgGP-Tree, đã được tạo ra để nâng cao hiệu quả tìm kiếm hình ảnh. Mô hình SgGP-Tree bổ sung tiêu chí để chọn các nút lá chiến thắng, giúp phân cụm tốt hơn và tìm kiếm hình ảnh chính xác hơn. Các thực nghiệm được thực hiện trên bộ dữ liệu WANG, ImageCLEF và MS-COCO. Hệ thống SBIR-GP có độ chính xác vượt trội so với các đề xuất trước đây của chính tác giả luận án. Hiệu suất thử nghiệm được so sánh với các phương pháp khác trên cùng một tập dữ liệu hình ảnh để đánh giá mô hình, phương pháp và thuật toán được đề xuất. Kết quả so sánh cho thấy hệ thống tra cứu SBIR-GP có độ chính xác cao hơn các nghiên cứu khác trên cùng một tập dữ liệu ảnh. Điều này chứng tỏ những đề xuất trong bài luận án là hiệu quả và phù hợp.

KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

1. Đóng góp của luận án

Trong luận án này, các phương pháp tìm kiếm ảnh dựa trên ngữ nghĩa đã được đề xuất và phát triển qua việc phân tích sâu rộng các nghiên cứu liên quan, nhằm xây dựng những mô hình tìm kiếm ảnh có hiệu suất cao. Luận án tập trung vào việc kết hợp các kỹ thuật phân cụm và ontology, với ba đóng góp chính:

- (1) Xây dựng cấu trúc GP-Tree để lưu trữ và lập chỉ mục dữ liệu ảnh: GP-Tree dựa trên phân cụm phân cấp, cho phép lưu trữ hiệu quả dữ liệu ảnh lớn thông qua các véc-tơ đặc trưng cấp thấp. Sự kết nối giữa đặc trưng

ảnh và từ vựng ngữ nghĩa giúp giảm thiểu kích thước dữ liệu, đồng thời tăng tốc độ và độ chính xác của quá trình tìm kiếm.

- (2) Hệ thống tìm kiếm ảnh theo ngữ nghĩa GP-SBIR dựa trên GP-Tree: GP-SBIR kết hợp GP-Tree với ontology, hoạt động qua hai giai đoạn: tiền xử lý và tìm kiếm. Giai đoạn tiền xử lý xây dựng GP-Tree và khung ontology bán tự động dựa trên RDF, trong khi giai đoạn tìm kiếm thực hiện tìm kiếm hình ảnh tương tự và truy vấn ngữ nghĩa thông qua SPARQL. Độ chính xác trên các bộ dữ liệu WANG, ImageCLEF, và MS-COCO đạt lần lượt 0.6780, 0.6802, và 0.7170.
- (3) Phát triển GP-Tree với hai mô hình mới là Graph-GPTree và SgGP-Tree: Graph-GPTree cải thiện độ chính xác tìm kiếm bằng cách liên kết các phần tử tương tự nằm trên các nhánh khác nhau của cây. Trong khi đó, SgGP-Tree kết hợp với SOM để tối ưu hóa phân cụm và chọn các nút lá hiệu quả hơn, từ đó nâng cao hiệu suất tìm kiếm. Thử nghiệm trên các bộ dữ liệu WANG, ImageCLEF, và MS-COCO cho thấy độ chính xác của Graph-GPTree lần lượt đạt 0.7665, 0.8168, và 0.8730, trong khi SgGP-Tree đạt 0.8004, 0.8926, và 0.8753. Những kết quả này khẳng định rằng các mô hình cải tiến GP-Tree mang lại hiệu quả cao trong việc tìm kiếm hình ảnh.

Luận án đã tiến hành thử nghiệm trên các tập dữ liệu nổi tiếng như WANG, ImageCLEF, và MS-COCO, và thu được kết quả vượt trội về độ chính xác so với các phương pháp tìm kiếm hiện đại khác. Những kết quả này không chỉ nâng cao hiệu quả tìm kiếm ảnh dựa trên ngữ nghĩa mà còn mở ra tiềm năng phát triển trong lĩnh vực này.

2. Hướng phát triển

Ngoài những kết quả đạt được vừa nêu, luận văn còn tồn tại một số điểm có thể phát triển mở rộng hơn:

- (1) So sánh với các phương pháp hiện đại: nghiên cứu các phương pháp tìm kiếm ảnh dựa trên mạng nơ-ron sâu (DNN), CNN, R-CNN, GCN... để so sánh với các phương pháp đã được đề xuất trong luận án.
- (2) Ứng dụng trong các lĩnh vực thực tế: phát triển ứng dụng tìm kiếm ảnh ngữ nghĩa trong các lĩnh vực cụ thể như xác định địa điểm du lịch từ ảnh, chẩn đoán bệnh qua ảnh y khoa, phân biệt các loại đá thô nhuộm, và tìm kiếm ảnh từ thông tin trên mạng xã hội.
- (3) Làm giàu ontology ngữ nghĩa: phát triển thêm các mối quan hệ ngữ nghĩa chặt chẽ hơn giữa các đối tượng và hành động trong ảnh, nhằm tăng cường độ chính xác trong việc hiểu và tìm kiếm ảnh.
- (4) Phát triển ontology tiếng Việt: mở rộng việc xây dựng ontology cho tìm kiếm ảnh ngữ nghĩa bằng ngôn ngữ tiếng Việt, mở ra những ứng dụng hữu ích cho người dùng trong nước.

TÀI LIỆU THAM KHẢO

- [1] S. Angra and S. Ahuja, "Machine learning and its applications: A review," in 2017 International Conference on Big Data Analytics and Computational Intelligence (ICBDAC), 2017, pp. 57-60: IEEE.
- [2] D. Karimi, G. Akbarizadeh, K. Rangzan, and M. Kabolizadeh, "Effective supervised multiple-feature learning for fused radar and optical data classification," *IET Radar, Sonar Navigation*, vol. 11, no. 5, pp. 768-777, 2017.
- [3] G. Duan, J. Yang, and Y. Yang, "Content-based image retrieval research," *Physics Procedia*, vol. 22, pp. 471-477, 2011.
- [4] A. Alzu'bi, A. Amira, N. Ramzan, and I. Representation, "Semantic content-based image retrieval: A comprehensive study," *Journal of Visual Communication*, vol. 32, pp. 20-54, 2015.
- [5] J. Filali, H. B. Zghal, and J. Martinet, "Towards visual vocabulary and ontology-based image retrieval system," in *International Conference on Agents and Artificial Intelligence*, 2016, vol. 2, pp. 560-565.
- [6] L. M. Thanh and e. al, "Image retrieval system based on EMD similarity measure and S-tree," (in V), *Intelligent Technologies and Engineering Systems*, Springer, New York, NY, pp. 139-146, 2013.
- [7] L. Piras and G. J. I. F. Giacinto, "Information fusion in content based image retrieval: A comprehensive overview," vol. 37, pp. 50-60, 2017.

- [8] S. Bruch, *Foundations of Vector Retrieval*. Springer, 2024.
- [9] M. A. Alzubaidi, "A new strategy for bridging the semantic gap in image retrieval," *International Journal of Computational Science Engineering*, vol. 14, no. 1, pp. 27-43, 2017.
- [10] C. S. Wickramasinghe, K. Amarasinghe, and M. Manic, "Parallalizable deep self-organizing maps for image classification," in *2017 IEEE symposium series on computational intelligence (SSCI)*, 2017, pp. 1-7: IEEE.
- [11] K. Kanwal, K. T. Ahmad, R. Khan, A. T. Abbasi, and J. Li, "Deep learning using symmetry, FAST scores, shape-based filtering and spatial mapping integrated with CNN for large scale image retrieval," *Symmetry*, vol. 12, no. 4, p. 612, 2020.
- [12] G.-H. Liu and J.-Y. Yang, "Deep-seated features histogram: a novel image retrieval method," *Pattern Recognition*, vol. 116, p. 107926, 2021.
- [13] O. Sikha and K. Soman, "Dynamic Mode Decomposition based salient edge/region features for content based image retrieval," *Multim. Tools Appl.*, vol. 80, no. 10, pp. 15937-15958, 2021.
- [14] S. Dhingra and P. Bansal, "Relative examination of texture feature extraction techniques in image retrieval systems by employing neural network: an experimental review," in *Proceedings of International Conference on Artificial Intelligence and Applications: ICAIA 2020*, 2021, pp. 337-349: Springer.
- [15] A. Ouni, E. Royer, M. Chevaldonné, and M. Dhome, "Leveraging semantic segmentation for hybrid image retrieval methods," *Neural Computing Applications*, vol. 34, no. 24, pp. 21519-21537, 2022.
- [16] A. Yang, X. Yang, W. Wu, H. Liu, and Y. J. I. a. Zhuansun, "Research on feature extraction of tumor image based on convolutional neural network," vol. 7, pp. 24204-24213, 2019.
- [17] Y. Qiang, C. Sheng, and D. Yin, "Method of tire pattern image retrieval based on wavelet transform and Siamese network," in *Proceedings of the 2020 International Conference on Aviation Safety and Information Technology*, 2020, pp. 587-592.
- [18] X. Yue *et al.*, "Prototypical cross-domain self-supervised learning for few-shot unsupervised domain adaptation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021, pp. 13834-13844.
- [19] N. T. U. Nhi and T. M. Le, "A model of semantic-based image retrieval using C-tree and neighbor graph," *International Journal on Semantic Web Information Systems*, vol. 18, no. 1, pp. 1-23, 2022.
- [20] X. Wang, D. Peng, M. Yan, and P. Hu, "Correspondence-free domain alignment for unsupervised cross-domain image retrieval," *arXiv preprint arXiv:06081*, 2023.
- [21] Y. Cao, M. Long, B. Liu, and J. Wang, "Deep cauchy hashing for hamming space retrieval," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 1229-1237.
- [22] Y. Xie, Y. Liu, Y. Wang, L. Gao, P. Wang, and K. Zhou, "Label-Attended Hashing for Multi-Label Image Retrieval," in *IJCAI*, 2020, pp. 955-962.
- [23] W. Gu, X. Gu, J. Gu, B. Li, Z. Xiong, and W. Wang, "Adversary guided asymmetric hashing for cross-modal retrieval," in *Proceedings of the 2019 on international conference on multimedia retrieval*, 2019, pp. 159-167.

**DANH MỤC CÁC BÀI BÁO ĐÃ XUẤT BẢN
LIÊN QUAN ĐẾN LUẬN ÁN**

1. **N. M. Hai**, T. V. Lang, and V. T. Thanh, "Semantic-Based Image Retrieval Using Hierarchical Clustering and Neighbor Graph," in World Conference on Information Systems and Technologies, 2022, pp. 34-44: Springer, DOI: https://doi.org/10.1007/978-3-031-04829-6_4 (**Scopus, Q4**)
2. **N. M. Hai**, V. T. Thanh, and T. V. Lang, "A method for semantic-based image retrieval using hierarchical clustering tree and graph," *Telkomnika*, vol. 20, no. 5, pp. 1026-1033, 2022, DOI: <http://doi.org/10.12928/telkomnika.v20i5.24086> (**Scopus, Q3**)
3. **N. M. Hai**, T. Van Lang and T. T. Van, "Improving The Efficiency of Semantic Image Retrieval using A Combined Graph and SOM Model," in *IEEE Access*, 2023, doi: <https://doi.org/10.1109/ACCESS.2023.3333678> (**SCIE, Q1**)
4. **N. M. Hai**, V. T. Thanh, and T. V. Lang, "The improvements of semantic-based image retrieval using hierarchical clustering tree," *FAIR'2020*, 2020, pp. 557-570: Natural Science and Technology Publishing House, DOI: <https://doi.org/10.15625/vap.2020.00213>
5. **N. M. Hai**, V. T. Thanh, and T. V. Lang, "A method of semantic-based image retrieval using graph cut," *Computer Science and Cybernetics*, vol. 38, no. 2, pp. 193-212, 2022, DOI: <https://doi.org/10.15625/1813-9663/38/2/16786>
6. **Nguyễn Minh Hải**, Trần Văn Lăng, Văn Thế Thành, "Một tiếp cận tìm kiếm ảnh theo ngữ nghĩa dựa trên mạng nơ-ron tích chập và ontology", *TCKHTrường ĐH Sư phạm TP. HCM*, 2022. tr. 48-59. DOI: [https://doi.org/10.54607/hcmue.js.19.3.3272\(2022\)](https://doi.org/10.54607/hcmue.js.19.3.3272(2022))