

MỘT MÔ HÌNH KẾT HỢP KỸ THUẬT PHÂN LỚP KNN VỚI ONTOLOGY CHO BÀI TOÁN TÌM KIẾM ẢNH THEO NGỮ NGHĨA

Phạm Thị Minh Quyên^{1,2}, Nguyễn Văn Trung², Nguyễn Thị Uyên Nhi^{3,*}

¹Trường THPT Quang Trung – An Khê, Gia Lai

²Trường Đại học Khoa học, Đại học Huế

³Trường ĐH Kinh Tế – Đại học Đà Nẵng

* Email: nhintu@due.edu.vn

TÓM TẮT

Tìm kiếm ảnh theo ngữ nghĩa một bài toán được quan tâm và đã có nhiều phương pháp được nghiên cứu và công bố trong thời gian gần đây. Trong nghiên cứu này, một mô hình kết hợp giữa phân lớp KNN (K- nearest neighbor) với Ontology được đề xuất xây dựng trong tìm kiếm ảnh theo ngữ nghĩa. Trước hết, hệ thống trích xuất đặc trưng của ảnh đầu vào và so sánh với các vec-to đặc trưng trong cơ sở dữ liệu, từ đó sử dụng thuật toán phân lớp KNN để xác định lớp của ảnh đầu vào. Sau đó, câu lệnh SPARQL được tự động tạo ra từ các từ vựng phân lớp để truy vấn trên ontology và truy xuất tập ảnh tương tự về ngữ nghĩa. Khung ontology kế thừa trước đó được bổ sung dữ liệu với bộ ảnh CIFAR-10 (60.000 ảnh), Caltech-101 (9.145 ảnh). Kết quả thực nghiệm trên hai bộ ảnh này được so sánh với một số công trình gần đây trên cùng bộ dữ liệu cho thấy tính hiệu quả và đúng đắn của phương pháp đã được đề xuất.

Từ khóa: KNN, ontology, tìm kiếm ảnh theo ngữ nghĩa

1. MỞ ĐẦU

Trong xu thế của cuộc cách mạng công nghiệp 4.0, các thiết bị điện tử như camera, smartphone, tablet, v.v, phát triển mạnh mẽ, khiến cho việc tạo ra ảnh số vô cùng đơn giản và đạt đến một khối lượng khổng lồ [10]. Tuy nhiên, người dùng gặp khó khăn trong việc tổ chức và tìm kiếm số lượng lớn hình ảnh trong cơ sở dữ liệu, vì hệ thống cơ sở dữ liệu thương mại hiện tại được thiết kế cho dữ liệu văn bản và không phù hợp với hình ảnh kỹ thuật số. Vì vậy, thách thức lớn của ngành thị giác máy tính là làm thế nào để tìm kiếm chính xác, khai thác ngữ nghĩa, truy cập và sử dụng những bộ dữ liệu ảnh số lớn một cách hiệu quả và đạt được yêu cầu của người dùng [1].

Trong thập kỷ qua nhiều phương pháp hiệu quả cho tìm kiếm hình ảnh theo ngữ nghĩa đã được đề xuất. Việc tìm kiếm ảnh theo nội dung cấp thấp của hình ảnh (màu sắc, hình dạng, kết cấu, ...) [4] có hạn chế bắt nguồn từ nguyên nhân không thể mô tả được ngữ nghĩa của hình ảnh, nghĩa là, tồn tại một “khoảng cách ngữ nghĩa” giữa đặc trưng cấp thấp của hình ảnh với ngữ nghĩa cấp cao của người dùng [1,10]. Có nhiều phương pháp tìm kiếm ảnh theo ngữ nghĩa được đề xuất để giảm khoảng cách ngữ nghĩa, trong đó tìm kiếm ngữ nghĩa ảnh dựa trên ontology được coi là phương pháp hiệu quả và gần gũi với ngữ nghĩa cấp cao của con người nhất [11], do nó tập trung vào việc nắm bắt nội dung ngữ nghĩa, mối quan hệ phân cấp giữa các từ vựng, điều này có thể giúp hệ thống đáp ứng yêu cầu của người dùng một cách tốt hơn. Trong bài báo này, một phương pháp kết hợp giữa đặc trưng cấp thấp và ngữ nghĩa cấp cao dựa trên ontology được đề xuất: (1) trích xuất đặc trưng cấp thấp của hình ảnh và chuyển đổi các đặc trưng này sang từ vựng ngữ nghĩa dựa vào kỹ thuật phân lớp KNN; (2) câu lệnh SPARQL được tạo ra từ các từ vựng thị giác và truy vấn trên ontology để truy xuất tập ảnh tương tự và ngữ nghĩa cấp cao mô tả hình ảnh. Đây chính là việc ánh xạ các đặc trưng cấp thấp sang ngữ nghĩa cấp cao nhằm giảm “khoảng cách ngữ nghĩa” của hình ảnh. Khung ontology được kế thừa từ nghiên cứu của nhóm tác giả Nhi N.T.U. [10] và bổ sung dữ liệu với các tập ảnh CIFAR-10 (60.000 ảnh), Caltech-101 (9.145 ảnh). Thử nghiệm trên hai tập ảnh này được so sánh với các phương pháp từ những nghiên cứu khác đã được công bố nhằm chứng minh tính hiệu quả và đúng đắn của đề xuất trong bài báo này.

Phần còn lại của bài báo gồm: Phần II khảo sát một số công trình liên quan đồng thời phân tích các ưu nhược điểm để chứng minh tính hiệu quả của tìm kiếm ảnh dựa trên KNN và Ontology; Phần III trình bày phương pháp xây dựng mô hình kết hợp giữa phân lớp KNN với Ontology; Phần IV là thực nghiệm, đánh giá so sánh; Phần V kết luận và hướng phát triển tương lai.

2. CÁC CÔNG TRÌNH NGHIÊN CỨU LIÊN QUAN

Bài toán tìm kiếm ảnh dựa trên kỹ thuật phân lớp KNN được quan tâm nghiên cứu trong những năm gần đây. Nhóm nghiên cứu Kumar Munish (2018) [7] đề xuất một hệ thống tìm kiếm ảnh dựa trên nội dung hiệu quả sử dụng hai bộ phân loại BayesNet và KNN. Để chứng minh tính hiệu quả của đề xuất, thực nghiệm được triển khai trên tập ảnh COREL (1000 ảnh), cho độ chính xác cao là 0.88. Tuy nhiên, phương pháp này chỉ thực hiện trích xuất đặc trưng cục bộ mà chưa tính đến đặc trưng toàn cục của hình ảnh. Mardi Kevin Song và cộng sự (2019) [9] sử dụng ba phương pháp trích xuất đặc trưng khác nhau được kết hợp với Phân loại KNN. Kết quả từ hệ thống cho thấy đặc trưng Color Moments có độ chính xác là 45,8%, đặc trưng Vùng kết nối có độ chính xác là 34,1% và DWT có độ chính xác là 44,5%. Trong khi đó, việc kết hợp các phương pháp trích xuất đặc trưng mà không có Phân loại KNN có độ chính xác là 37,1%, còn khi kết hợp với Phân loại KNN, hệ thống có độ chính xác 68,7%. Thử nghiệm cho thấy rằng hệ

thống cần Phân loại KNN để tăng khả năng sử dụng và độ chính xác của việc truy xuất. Tuy nhiên, tập dữ liệu ảnh nhỏ với 1000 hình ảnh động vật, cho kết quả truy vấn chưa cao, do KNN không thực sự hoạt động tốt trong tập ảnh nhỏ. [12] Sathiamoorthy S. và cộng sự (2020) [12] đề xuất hệ thống tìm kiếm ảnh dựa trên KNN. Bộ ảnh Wang (10.800 ảnh) được sử dụng để đánh giá tính hiệu quả của mô hình, cho độ chính xác 80.4%. Kết quả này cho thấy với bộ dữ liệu ảnh lớn, KNN thực hiện nhiệm vụ phân loại và tìm kiếm hình ảnh hiệu quả. Tuy nhiên, do đặc trưng hình ảnh bị loại bỏ đi đặc trưng hình dạng, nên việc xác định vùng đối tượng hay loại bỏ nhiều trong ảnh chưa được quan tâm.

Như vậy, kỹ thuật phân loại KNN hiệu quả trong bài toán tìm kiếm hình ảnh trong nội dung. Tuy nhiên, hệ thống tìm kiếm ảnh theo nội dung chỉ tìm kiếm các hình ảnh tương tự nhau theo đặc trưng ảnh nên thiếu đi sự tương đồng về ngữ nghĩa. Vì vậy bài toán tìm kiếm ảnh trên ontology để trích xuất ngữ nghĩa của hình ảnh được nhiều nhà nghiên cứu quan tâm. Nhóm nghiên cứu Allani Olfa (2017) [3] đề xuất hệ thống truy xuất hình ảnh dựa trên đồ thị mẫu kết hợp các đặc trưng cấp thấp và ngữ nghĩa dựa trên ontology. Để đánh giá cách tiếp cận, nhóm nghiên cứu đã sử dụng tập ảnh ImageCLEF (SIAPR-TC 12) bao gồm 20.000 hình ảnh được chú thích. Mặc dù đây là cách tiếp cận đầy tiềm năng khi ontology được xây dựng trên tập dữ liệu lớn với 276 lớp, tuy nhiên việc xây dựng tự động ontology không thể xác định được tính đúng đắn và không có sự kiểm nghiệm của chuyên gia. Do đó độ chính xác của tìm kiếm ảnh dựa trên ontology này là không cao, chỉ với 35.13%. Nhi N.T.U và cộng sự (2022) [11] đề xuất một mô hình tìm kiếm ảnh theo ngữ nghĩa dựa vào đặc trưng cấp thấp trên GraphC-Tree và ontology. Ontology được xây dựng bán tự động dựa trên tập ảnh ImageCLEF bao gồm 20.000 hình ảnh và 276 phân lớp. Ảnh được phân lớp và bổ sung tự động vào các phân cấp lớp đã xây dựng. Các tập ảnh thực nghiệm là ImageCLEF, COREL, WANG và Stanford Dogs với độ chính xác lần lượt là 83.98%, 88.85%, 76.65% và 82.64%. Kết quả này vượt trội so với các phương pháp khác trên cùng tập dữ liệu cho thấy tính hiệu quả của ontology. Đồng thời, dữ liệu ảnh có thể bổ sung để làm giàu thêm cho khung ontology.

Từ các công trình nghiên cứu liên quan, có thể thấy tính hiệu quả của tìm kiếm ảnh kết hợp giữa đặc trưng cấp thấp và ontology. Tuy nhiên, việc xây dựng ontology cần được tập trung nghiên cứu để đạt được độ chính xác cao. Đây chính là động lực cho nghiên cứu của chúng tôi trong bài báo này.

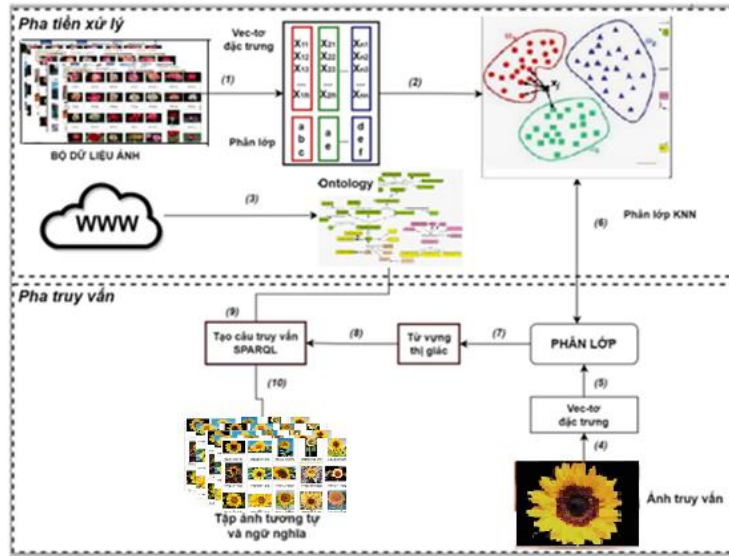
3. TÌM KIẾM ẢNH THEO NGỮ NGHĨA DỰA VÀO KNN VÀ ONTOLOGY

Bài báo này đề xuất một hệ tìm kiếm ảnh theo ngữ nghĩa kết hợp giữa kỹ thuật phân lớp KNN và ontology, gọi là onKNN. Trong phần này, kiến trúc của mô hình onKNN và các thành phần của nó sẽ được trình bày.

3.1. Kiến trúc mô hình tìm kiếm ảnh onKNN

Một mô hình kết hợp kỹ thuật phân lớp KNN với ontology cho bài toán tìm kiếm ảnh theo ngữ nghĩa

Hình 1 mô tả kiến trúc của mô hình tìm kiếm ảnh theo ngữ nghĩa onKNN, bao gồm 2 pha: (a) Pha tiền xử lý và (b) pha truy vấn.



Hình 1. Kiến trúc hệ truy vấn onKNN

(a) Pha tiền xử lý:

- Bước 1. Từ bộ dữ liệu ảnh, thực hiện kỹ thuật trích xuất các vec-tơ đặc trưng cấp thấp (1).
- Bước 2. Thực hiện phân lớp dữ liệu với kỹ thuật phân lớp KNN (2).
- Bước 3. Bổ sung dữ liệu ảnh vào khung ontology được kế thừa (3).

(b) Pha truy vấn:

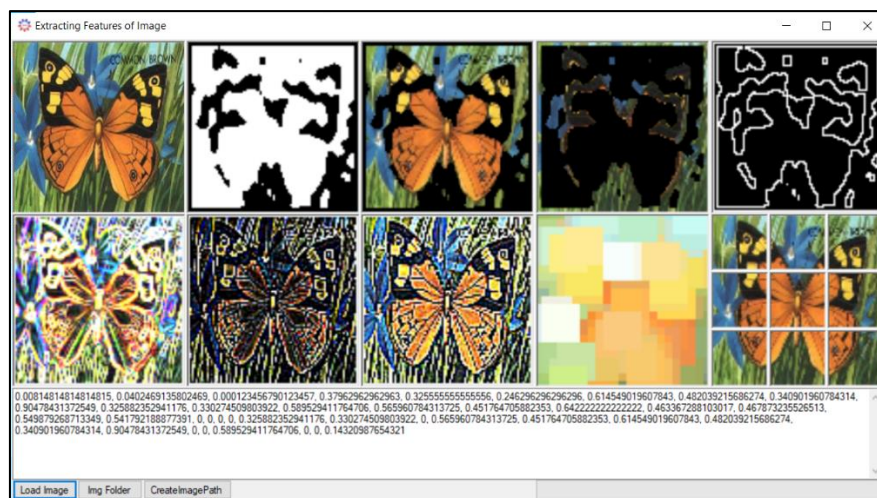
- Bước 1. Từ ảnh truy vấn đầu vào thực hiện trích xuất vec-tơ đặc trưng (4).
- Bước 2. Thực hiện phân lớp (5) trên mô hình KNN đã huấn luyện (6) để tìm ra tập từ vựng thị giác (7).
- Bước 3. Tạo câu truy vấn SPARQL từ từ vựng thị giác (8) và truy vấn trên ontology (9) để tìm tập ảnh tương tự và ngữ nghĩa của nó (10).

Như vậy, từ một hình ảnh truy vấn đầu vào, một vec-tơ đặc trưng cấp thấp được trích xuất, sau đó thuật toán KNN thực hiện phân lớp ảnh và truy xuất một tập từ vựng thị giác. Câu truy vấn SPARQL được tự động tạo ra từ các từ vựng này và truy vấn trên ontology. Kết quả của quá trình này là tập ảnh tương tự và ngữ nghĩa của nó gồm các metadata, URI và các khái niệm ngữ nghĩa, ...

Các thành phần chính trong mô hình tìm kiếm ảnh onKNN bao gồm: (1) kỹ thuật trích xuất vec-tơ đặc trưng của hình ảnh; (2) kỹ thuật phân lớp KNN; (3) bổ sung dữ liệu cho khung ontology. Những thành phần này sẽ được trình bày trong các phần sau của bài báo.

3.2. Trích xuất vec-tơ đặc trưng cấp thấp

Trong bài báo này, phương pháp trích xuất đặc trưng tổng hợp dựa trên các đặc trưng cấp thấp như màu sắc, kết cấu, hình dạng được thực hiện.



Hình 2. Minh họa về trích xuất vec-tơ đặc trưng hình ảnh của mô hình onKNN

Hình 2 là minh họa về trích xuất vec-tơ đặc trưng của mô hình onKNN. Các kỹ thuật trích xuất đặc trưng bao gồm kỹ thuật trích xuất đường biên đối tượng dựa trên phương pháp LoG (Laplacian of Gaussian), kỹ thuật trích xuất đặc trưng dựa trên phương pháp gom cụm màu Newton và phương pháp trích xuất kết bề mặt ảnh dựa trên các phép lọc Sobel, Laplacian. Một vec-tơ đặc trưng cấp thấp cho mỗi hình ảnh gồm 137 chiều được tạo ra nhằm thực hiện cho bài toán tìm kiếm ảnh theo nội dung.

3.3. Kỹ thuật phân lớp KNN

Thuật toán KNN là một thuật toán học có giám sát dùng để phân loại đối tượng mới bằng cách tìm điểm tương đồng giữa đối tượng mới này với dữ liệu đã có. Với KNN, trong bài toán phân lớp, nhãn của một điểm dữ liệu được suy ra trực tiếp từ k điểm dữ liệu gần nhất trong tập huấn luyện. Nhãn của một điểm dữ liệu có thể được quyết định bằng sự trên lệnh số lượng giữa các điểm gần nhất, không quan tâm đến việc có một vài điểm dữ liệu trong những điểm gần nhất này là nhiều. Thuật toán KNN thực hiện một phép tính toán học để đo khoảng cách giữa dữ liệu mới với tất cả các điểm trong tập dữ liệu học D để thực hiện phân lớp. Có nhiều phương pháp để đo khoảng cách. Phép đo khoảng cách Euclid là một trong những phép đo khoảng cách được sử dụng phổ biến nhất, đây cũng chính là khoảng cách được sử dụng trong bài báo này.

Ưu điểm của thuật toán phân lớp KNN là: thuật toán đơn giản, dễ dàng triển khai, độ phức tạp tính toán nhỏ, hoạt động tốt trong trường hợp phân loại với nhiều lớp và xử lý tốt với tập dữ liệu nhiễu. Tuy nhiên, với K nhỏ dễ gặp nhiễu dẫn tới kết quả đưa ra không chính xác. Do đó, thuật toán KNN phù hợp với bài toán đặt ra trong bài báo này, được ứng dụng cho tập dữ liệu có số lượng tương đối lớn và nhiễu.

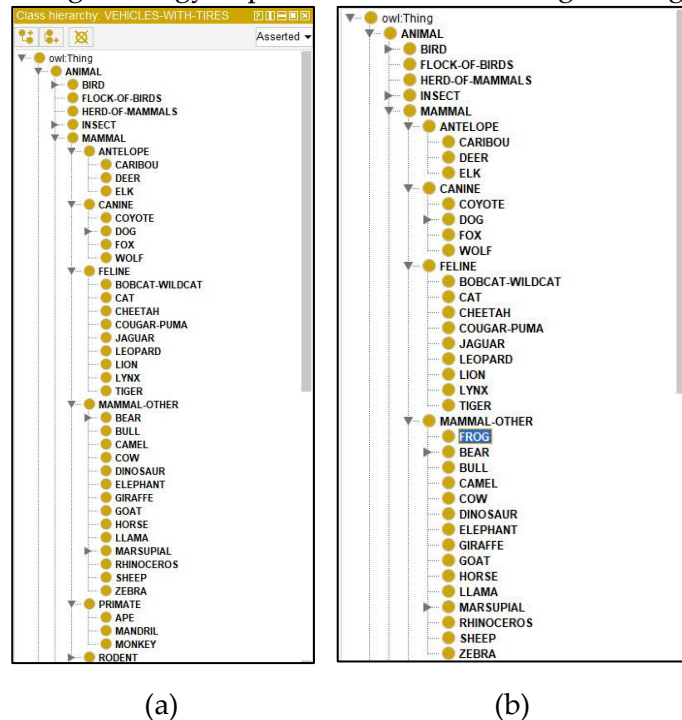
3.4. Bổ sung dữ liệu cho khung ontology

Một mô hình kết hợp kỹ thuật phân lớp KNN với ontology cho bài toán tìm kiếm ảnh theo ngữ nghĩa

Trong bài báo này, chúng tôi kế thừa khung ontology đã được xây dựng trong luận án “Nâng cao hiệu quả tìm kiếm ảnh theo tiếp cận ngữ nghĩa” [3]. Chúng tôi sử dụng khung ontology này làm cơ sở để bổ sung dữ liệu với các tập ảnh mới là CIFAR-10 và Caltech-101. Việc bổ sung dữ liệu vào khung ontology phải đảm bảo tính đúng đắn, nhất quán về cấu trúc và kế thừa các thông tin sẵn có. Vì vậy, việc bổ sung dữ liệu cho khung ontology từ một tập ảnh cho trước được thực hiện theo các quy tắc sau:

- Các lớp chỉ được bổ sung vào Khung ontology nếu lớp đó chưa được định nghĩa trước đó;
- Khi bổ sung lớp, xem xét các phân cấp cha-con để bổ sung vào;
- Bổ sung các định nghĩa cho các lớp mới cũng như các thuộc tính mới cho lớp được thêm vào;
- Các cá thể hình ảnh được phân loại dựa trên KNN để xác định lớp, sau đó thêm vào tự động trong ontology dựa vào phần mềm bổ sung dữ liệu SBIR-Ontology đã được xây dựng [3].

Để trực quan hóa cho quá trình này bổ sung dữ liệu, ontology được hiển thị bằng Protégé. Hình 3 là một minh họa về bổ sung dữ liệu với tập ảnh CIFAR-10 với mười lớp: airplane, car, bird, cat, deer, dog, frog, horse, ship, and truck. Các lớp này được xem xét khi bổ sung vào khung ontology, lớp nào đã tồn tại sẽ không bổ sung.



Hình 3. Một minh họa về bổ sung lớp dữ liệu vào ontology

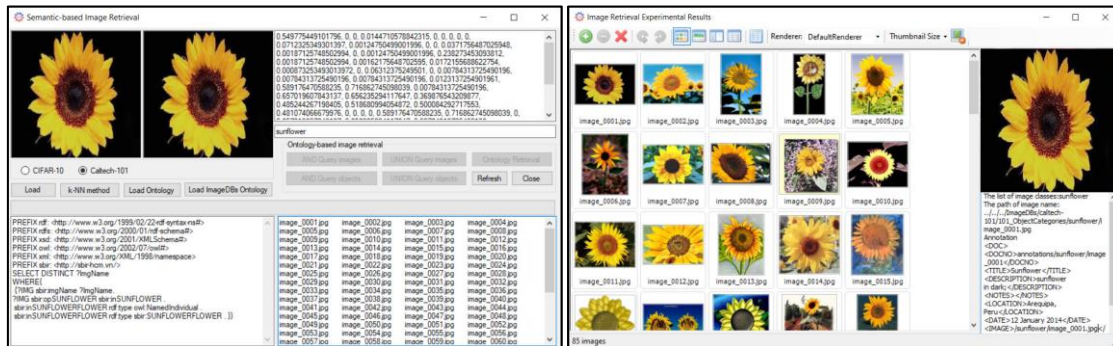
Hình 3 (a) cho thấy ban đầu, khung ontology đã tồn tại các phân lớp bird, cat, deer, dog, horse, do đó không thực hiện bổ sung lớp. Trong khi đó lớp “Frog” chưa tồn

tại, dựa vào từ điển ngữ nghĩa WORDNET, lớp Frog là lớp con của Mammal-other và được bổ sung vào khung ontology (Hình 3b).

4. TỔ CHỨC THỰC NGHIỆM VÀ ĐÁNH GIÁ KẾT QUẢ

4.1. Tổ chức thực nghiệm

Để xác định hiệu quả của các phương pháp đã đề xuất, thực nghiệm được xây dựng trên nền tảng .NET Framework 4.8, ngôn ngữ lập trình C#. Các đồ thị được xây dựng trên Matlab 2019. Cấu hình máy tính của thực nghiệm: Intel(R) CoreTM i7-8750H, CPU 2,70GHz, RAM 8GB và hệ điều hành Windows 10 Professional. Để đánh giá tính đúng đắn và hiệu quả của lý thuyết đề xuất, thực nghiệm của hệ onk-NN được xây dựng trên các tập ảnh CIFAR-10 (60.000 ảnh) và Caltech-101 (9.145 ảnh).



Hình 4. Một minh họa giao diện truy vấn hình ảnh của hệ onk-NN

Hình 4 là minh họa của giao diện tìm kiếm ảnh: với một ảnh đầu vào, vec-tơ đặc trưng cấp thấp được trích xuất, sau đó thuật toán KNN được sử dụng để phân lớp ảnh. Từ đó, câu lệnh SPARQL được tự động tạo ra từ phân lớp ảnh trong bước trước để truy vấn trên ontology và truy xuất các filename của các hình ảnh tương tự. Kết quả của truy vấn bao gồm: tập ảnh tương tự theo ngữ nghĩa đầu vào, mỗi hình ảnh được giải thích ngữ nghĩa như các metadata, URI, annotation, ... Đồng thời, định nghĩa cho phân lớp ảnh (Hình 4) cũng được truy xuất nhằm bổ sung thêm ngữ nghĩa cho hình ảnh đầu vào.



Hình 4. Một định nghĩa cho phân lớp ảnh

4.2. Đánh giá hiệu suất truy vấn ảnh

Một mô hình kết hợp kỹ thuật phân lớp KNN với ontology cho bài toán tìm kiếm ảnh theo ngữ nghĩa

Để đánh giá hiệu quả tìm kiếm hình ảnh, luận án sử dụng các yếu tố để đánh giá, bao gồm: độ chính xác (precision), độ phủ (recall), độ dung hoà (F-measure), thời gian truy vấn (times). Các giá trị hiệu suất và thời gian tìm kiếm trung bình của các tập dữ liệu Caltech, CIFAR-10 dựa trên những phương pháp khác nhau như tìm kiếm ảnh dựa trên KNN và tìm kiếm ảnh theo ngữ nghĩa dựa trên KNN kết hợp ontology (onKNN) được tổng hợp trong Bảng 1 và Bảng 2.

Bảng 1. Hiệu suất tìm kiếm ảnh của Caltech-101

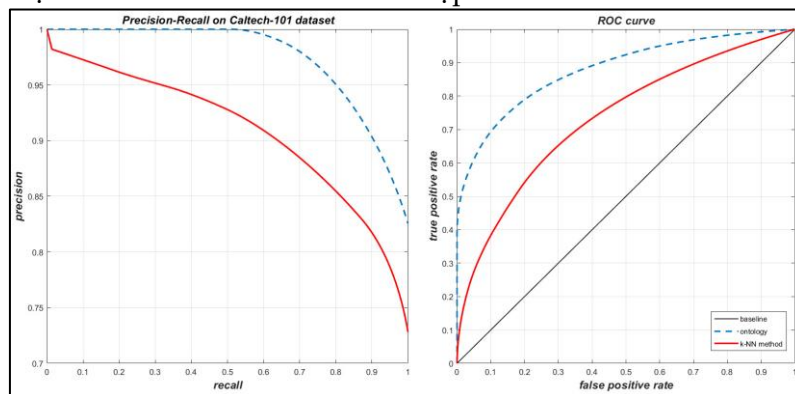
Phương pháp	Avg. precision	Avg. recall	Avg. F-measure	Avg. query time (ms)
KNN	0.743167	0.454464	0.563715	96.5368
onKNN	0.831719	0.61975	0.709236	118.6907

Bảng 2. Hiệu suất tìm kiếm ảnh của CIFAR-10

Phương pháp	Avg. precision	Avg. recall	Avg. F-measure	Avg. query time (ms)
KNN	0.701567	0.422978	0.525208	92.989373
onKNN	0.8290875	0.4885210	0.6123246	116.2424

Bảng 1 cho thấy độ chính xác tìm kiếm ảnh của tập dữ liệu Caltech-101 dựa trên onKNN cao hơn so với phương pháp KNN thuần túy, cụ thể là tăng 5.62%. Tuy nhiên, phương pháp onKNN có thời gian tìm kiếm cao hơn, do việc tìm kiếm phải thông qua thêm một bộ lọc ontology. Tương tự với tập ảnh CIFAR-10, bảng 2 cho thấy độ chính xác của onKNN cao vượt trội 8.33% so với phương pháp KNN. Sự vượt trội này có được do ảnh của bộ CIFAR-10 có độ phân giải thấp, nên sử dụng KNN không có được hiệu quả tốt, kết hợp thêm ontology để tìm kiếm ảnh theo ngữ nghĩa đã nâng cao độ chính xác, nên thời gian tìm kiếm trên onKNN cũng tăng.

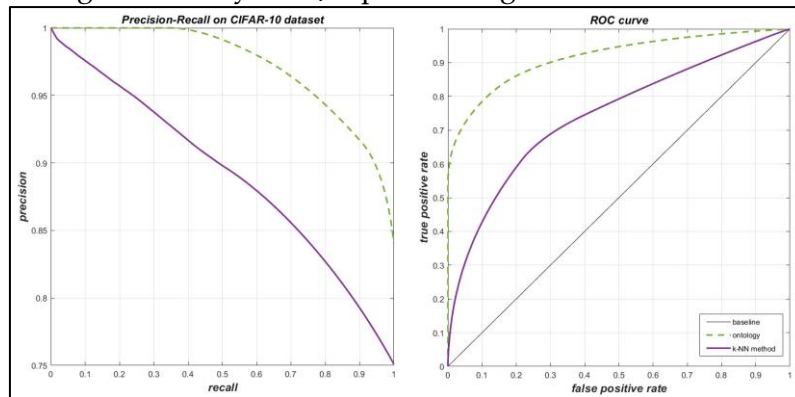
Ngoài ra, để đánh giá các kết quả của hệ truy vấn, một đường cong đặc trưng thu nhận ROC (Receiver Operating Characteristic) được thực hiện. Diện tích phía dưới đường cong AUC (Area Under the Curve), giới hạn trong không gian ROC, là thước đo cho độ chính xác của truy vấn, diện tích càng lớn thì độ chính xác càng cao. Hình 5 và Hình 6 là đồ thị Precision-Recall và ROC của tập ảnh Caltech-101 và CIFAR-10.



Hình 5. Hiệu suất truy vấn hình ảnh dựa trên tập ảnh Caltech-101

Hình 5 và Hình 6 cho thấy, diện tích dưới đường cong AUC của đồ thị Precision-Recall và đồ thị ROC tập ảnh Caltech-101, CIFAR-10 với phương pháp onKNN lớn hơn KNN. Từ các đánh giá thực nghiệm này cho thấy, đề xuất tìm kiếm ảnh dựa vào kỹ thuật

KNN và ontology hiệu quả hơn so với phương pháp sử dụng kỹ thuật KNN thông thường. Mặc dù là tăng thời gian tìm kiếm nhưng không đáng kể, đồng thời, trong luận văn này, chúng tôi tập trung vào phương pháp để tăng độ chính xác ảnh. Vậy, đề xuất của chúng tôi trong bài báo này là hiệu quả và đúng đắn.



Hình 6. Hiệu suất truy vấn hình ảnh dựa trên tập ảnh CIFAR-10

4.3. So sánh kết quả thực nghiệm

Để đánh giá kết quả thực nghiệm, chúng tôi so sánh với các công trình nghiên cứu khác trên cùng tập dữ liệu ảnh, được tổng hợp trong bảng 3 và bảng 4.

Bảng 3. So sánh độ chính xác giữa các phương pháp trên bộ dữ liệu Caltech-101

Phương pháp	Độ chính xác trung bình
Tian X. et al. (2018) [14]	0.814
Ahmed K. T. et al. (2019) [2]	0.698
Sathiamoorthy S. et al. (2020) [12]	0.694
Giveki D. et al. (2021) [6]	0.826
onKNN	0.8317

Bảng 4. So sánh độ chính xác giữa các phương pháp trên bộ dữ liệu CIFAR-10

Phương pháp	Độ chính xác trung bình
Song J. et al. (2018) [13]	0.758
Gao L. et al. (2019) [5]	0.814
Lu H. et al. (2020) [8]	0.599
onKNN	0.829

Bảng 3 và bảng 4 cho thấy phương pháp đề xuất của chúng tôi dựa trên sự kết hợp giữa kỹ thuật KNN và ontology có độ chính xác trung bình cao hơn một số phương pháp khác trên cùng một tập dữ liệu ảnh. Vậy, những đề xuất trong bài báo là chính xác và hiệu quả.

5. KẾT LUẬN

Bài báo này đã đề xuất mô hình tìm kiếm dữ liệu ảnh theo tiếp cận ngữ nghĩa với sự kết hợp giữa KNN và ontology. Mô hình này bổ sung dữ liệu từ các tập ảnh CIFAR-10, Caltech-101 vào khung ontology [3]. Mỗi hình ảnh trong tập dữ liệu ảnh được trích xuất đặc trưng và phân lớp dựa vào thuật toán KNN. Với mỗi ảnh truy vấn, vec-tơ đặc

Một mô hình kết hợp kỹ thuật phân lớp KNN với ontology cho bài toán tìm kiếm ảnh theo ngữ nghĩa

trung được trích xuất và tính toán với KNN để xác định phân lớp. Sau đó, câu lệnh SPARQL được tự động tạo ra từ các từ vựng phân lớp để truy vấn trên ontology và truy xuất tập ảnh tương tự về ngữ nghĩa cũng như khái niệm ngữ nghĩa của phân lớp. Thử nghiệm trên các tập ảnh CIFAR-10, Caltech-101 với độ chính xác lần lượt là 0.839 và 0.8317. Kết quả thử nghiệm được so sánh với các phương pháp khác trên cùng bộ dữ liệu thể hiện ưu điểm của mô hình được đề xuất trong bài báo. Trong định hướng tương lai, chúng tôi tiếp tục nghiên cứu các kỹ thuật phân loại dữ liệu, bổ sung các mối quan hệ trong khung ontology.

LỜI CẢM ƠN

Nhóm tác giả chân thành cảm ơn Trường Đại học Khoa học, Đại học Huế, Đại học Kinh tế - Đại học Đà Nẵng đã hỗ trợ về chuyên môn và cơ sở vật chất để nhóm tác giả hoàn thành nghiên cứu này.

TÀI LIỆU THAM KHẢO

- [1] Alzubaidi, M. A. (2017). A new strategy for bridging the semantic gap in image retrieval. *International Journal of Computational Science and Engineering*, 14(1), 27-43.
- [2] Ahmed, K. T., Ummesafi, S., & Iqbal, A. (2019). Content based image retrieval using image features information fusion. *Information Fusion*, 51, 76-99.
- [3] Allani, O., Zghal, H. B., Mellouli, N., & Akdag, H. (2017). Pattern graph-based image retrieval system combining semantic and visual features. *Multimedia Tools and Applications*, 76(19), 20287-20316
- [4] Alsmadi, M. K. (2020). Content-based image retrieval using color, shape and texture descriptors and features. *Arabian Journal for Science and Engineering*, 45(4), 3317-3330.
- [5] Gao, L., Zhu, X., Song, J., Zhao, Z., & Shen, H. T. (2019). Beyond product quantization: Deep progressive quantization for image retrieval. *arXiv preprint arXiv:1906.06698*.
- [6] Giveki, D., Shakarami, A., Tarrah, H., & Soltanshahi, M. A. (2022). A new method for image classification and image retrieval using convolutional neural networks. *Concurrency and Computation: Practice and Experience*, 34(1), e6533.
- [7] Kumar, M., Chhabra, P., & Garg, N. K. (2018). An efficient content-based image retrieval system using BayesNet and K-NN. *Multimedia Tools and Applications*, 77(16), 21557-21570.
- [8] Lu, H., Zhang, M., Xu, X., Li, Y., & Shen, H. T. (2020). Deep fuzzy hashing network for efficient image retrieval. *IEEE transactions on fuzzy systems*, 29(1), 166-176.
- [9] Mardi, K. S., Mawardi, V. C., & Perdana, N. J. (2019, October). KNN Classification for CBIR with Color Moments, Connected Regions, Discrete Wavelet Transform. In *2019 International Conference on Electrical, Electronics and Information Engineering (ICEEIE)* (Vol. 6, pp. 238-243). IEEE.
- [10] Nguyễn Thị Uyên Nhi (2022). “Nâng cao hiệu quả tìm kiếm ảnh theo tiếp cận ngữ nghĩa” Luận án Tiến sĩ, Khoa học máy tính, Trường Đại học Khoa học, Đại học Huế.
- [11] Nhi, N. T. U., & Le, T. M. (2022). A Model of Semantic-Based Image Retrieval Using C-Tree and Neighbor Graph. *International Journal on Semantic Web and Information Systems (IJSWIS)*, 18(1), 1-23.
- [12] Sathiamoorthy, S., & Natarajan, M. (2020). An efficient content-based image retrieval using enhanced multi-trend structure descriptor. *SN Applied Sciences*, 2(2), 1-20.

- [13] Song, J., He, T., Gao, L., Xu, X., Hanjalic, A., & Shen, H. T. (2018, April). Binary generative adversarial networks for image retrieval. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 32, No. 1).
- [14] Tian, X., Zheng, Q., & Xing, J. (2018, October). Content-Based Image Retrieval System Via Deep Learning Method. In 2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC) (pp. 1257-1261). IEEE.

A MODEL COMBINED KNN TECHNIQUE AND ONTOLOGY FOR SEMANTIC-BASED IMAGE RETRIEVAL

Pham Thi Minh Quyen^{1,2}, Le Manh Thanh², Nguyen Thi Uyen Nhi^{3,*}

¹Quang Trung High School – An Khe, Gia Lai

²University of Sciences, Hue University

³University of Economics – Danang University

*Email: nhintu@due.edu.vn

ABSTRACT

Semantic-based image retrieval is an interesting problem and many methods have been researched and published recently. In this study, a model combining KNN (K-nearest neighbor) classifier and ontology is proposed to be built in semantic-based image retrieval. First, the system extracts the features of the input image and compares it with the feature vectors in the database, thereby using the KNN classification algorithm to determine the classifier vocabulary. Then, the SPARQL command is automatically generated from this classifier vocabulary to query on the ontology and retrieve the semantically similar image set. The previous legacy ontology framework was supplemented with the image datasets CIFAR-10 (60,000 images), Caltech-101 (9,145 images). The experimental results on these two sets of images are compared with some recent works on the same data set, showing the effectiveness and correctness of the proposed method.

Keywords: KNN, ontology, semantic-based image retrieval