

MỘT MÔ HÌNH TÌM KIẾM ẢNH VỚI PHƯƠNG PHÁP LAI GIỮA KỸ THUẬT PHÂN CỤM K-MEANS VÀ DBSCAN

Hồ Thị Nữ^{1,2}, Nguyễn Dũng², Nguyễn Thị Uyên Nhi^{3,*}

¹ Trường THPT Nguyễn Huệ, ĐăcĐoa, Gia Lai

² Trường Đại học Khoa học, Đại học Huế

³ Trường Đại học Kinh tế, Đại học Đà Nẵng

* Email: nhintu@due.edu.vn

Ngày nhận bài: 13/12/2022; ngày hoàn thành phần biện: 22/12/2022; ngày duyệt đăng: 26/6/2023

TÓM TẮT

Các phương pháp tìm kiếm ảnh dựa trên nội dung nhằm mục đích truy xuất chính xác tập các ảnh tương tự từ cơ sở dữ liệu lớn dựa trên đặc trưng cấp thấp của hình ảnh. Trong nghiên cứu này, một phương pháp trích xuất vec-to đặc trưng từ màu sắc, hình dạng và kết cấu theo chuẩn MPEG-7 đã được thực hiện. Từ đó, phương pháp lai được đề xuất để phân cụm tốt hơn cho các đặc trưng này dựa trên việc kết hợp thuật toán phân cụm K-means và DBSCAN: (1) Thuật toán DBSCAN được sử dụng để phát hiện các cụm có mật độ dữ liệu tối thiểu và lọc nhiễu ban đầu; (2) sau đó các cụm được chia thành k nhóm phân tán đồng đều bằng thuật toán K-means. Để chứng minh tính hiệu quả của đề xuất, bài báo thực nghiệm trên các tập dữ liệu ảnh COREL-1K và Caltech-256, với độ chính xác lần lượt là 0.863 và 0.803. Kết quả thực nghiệm được so sánh với các phương pháp khác cho thấy hệ thống đề xuất có hiệu suất vượt trội so với các hệ thống còn lại về độ chính xác.

Từ khóa: DBSCAN, K-means, phương pháp lai, tìm kiếm ảnh theo nội dung.

1. MỞ ĐẦU

Trong thời đại của máy tính và Internet, hầu hết mọi lĩnh vực trong cuộc sống của con người, chẳng hạn như thương mại, giáo dục, thư viện, bệnh viện, phòng chống tội phạm, v.v., đều sử dụng hình ảnh cho các dịch vụ [10]. Do đó, tìm kiếm ảnh tương tự nhanh chóng, chính xác trong kho ảnh lớn một cách thức và nhiệm vụ cấp thiết trong lĩnh vực thị giác máy tính. Nhiều hệ thống tìm kiếm hình ảnh hiệu quả đã được phát triển. Các phương pháp tìm kiếm ảnh thường được phân thành hai loại chủ yếu: tìm kiếm ảnh dựa trên văn bản TBIR (Text-Based Image Retrieval) [7] và tìm kiếm ảnh dựa trên nội dung CBIR (Content-Based Image Retrieval) [5,8]. Trong các hệ thống

TBIR, hình ảnh được chú thích thủ công bằng bộ mô tả văn bản để thực hiện truy xuất hình ảnh một cách đơn giản, nhanh chóng và hiệu quả. Tuy nhiên, TBIR tốn nhiều thời gian và nhân lực cho việc chú thích thủ công cũng như thiếu sự chính xác của chú thích do sự chủ quan trong nhận thức của con người.

Để khắc phục những nhược điểm trên hệ thống tìm kiếm ảnh dựa trên nội dung (CBIR) đã được giới thiệu. Trong CBIR, hình ảnh được lập chỉ mục theo nội dung trực quan của chúng, như màu sắc, kết cấu, hình dạng [9]. Mỗi hình ảnh trong cơ sở dữ liệu được trích xuất bằng một vec-tơ đặc trưng. Sau đó, vec-tơ đặc trưng của hình ảnh đầu vào được so sánh với cơ sở dữ liệu đặc trưng dựa trên độ đo khoảng cách để tìm kiếm và truy xuất tập các hình ảnh tương tự. Có nhiều phương pháp trích chọn đặc trưng hình ảnh, việc lựa chọn phương pháp ảnh hưởng đến độ chính xác của hệ thống CBIR: Nhóm nghiên cứu Ahmed K.T. (2018) [1] đã nghiên cứu một hệ CBIR sử dụng đặc trưng kết hợp hình ảnh: màu RGB được sử dụng để trích xuất đặc trưng màu, ma trận đồng xuất hiện mức xám GLCM (Gray-Level Co-Occurrence Matrix) để trích xuất các cạnh và góc của đối tượng trong quá trình hình thành hình dạng. Kết quả của thí nghiệm được thực hiện dựa trên kỹ thuật này cho thấy nó vượt trội hơn hẳn so với kỹ thuật CBIR hiện có. Nazir và cộng sự (2018) [9] đề xuất một hệ CBIR với sự kết hợp màu sắc bằng không gian màu HSV (Hue, Saturation, Value) và kết cấu được trích xuất bằng cách sử dụng biến đổi wavelet rời rạc (DWT) và bộ mô tả biểu đồ cạnh Canny. Alsmadi Mutasem (2020) [2] sử dụng bộ mô tả đặc trưng với màu RGB, phương pháp biểu đồ cạnh Canny để trích xuất các đặc điểm hình dạng, biến đổi wavelet rời rạc và ma trận đồng xuất hiện mức xám GLCM (gray-level co-occurrence matrix) để trích xuất đặc điểm kết cấu. Với bộ đặc trưng kết hợp này, thực nghiệm trên tập ảnh COREL cho độ chính xác cao (0.9015). Từ các nghiên cứu về trích xuất đặc trưng cho thấy, các đặc trưng kết hợp của hình ảnh cho hiệu quả cao trong tìm kiếm ảnh. Tùy thuộc vào yêu cầu của bài toán và bộ dữ liệu ảnh để trích chọn các đặc trưng từ các phương pháp khác nhau.

Hệ thống CBIR hiệu quả không chỉ phụ thuộc vào việc trích chọn đặc trưng của hình ảnh mà còn phụ thuộc vào phương pháp lưu trữ và tổ chức dữ liệu ảnh [3,4]. Với mục đích này, nhiều kỹ thuật học máy khác nhau đã được phát triển để cải thiện hiệu suất của hệ thống CBIR. Trong đó, phân cụm [10] là một trong những phương pháp học máy không giám sát được ứng dụng rộng rãi. Nhiều kỹ thuật phân cụm dữ liệu được ứng dụng trong bài toán tìm kiếm ảnh theo nội dung trong những năm gần đây. Ouhda Mohamed và cộng sự (2018) [11] xây dựng một mô hình tìm kiếm ảnh dựa trên kỹ thuật phân cụm k-Means, sử dụng khoảng cách Euclide để đối sánh hình ảnh và đo độ tương tự để truy xuất hình ảnh. Bộ ảnh COREL được sử dụng để thực nghiệm cho mô hình cho độ chính xác 0.831. Malviya Neha và cộng sự (2017) [8] sử dụng phương pháp kết hợp giữa phân cụm K-means và phân cụm phân cấp cho bài toán tìm kiếm ảnh y khoa theo nội dung. Do đặc điểm của ảnh y khoa chỉ có màu xám, nhóm nghiên

cứu chỉ trích xuất đặc trưng kết cấu với ma trận đồng xuất hiện mức xám GLCM. Với số lượng ảnh lớn, các đặc trưng ảnh được gom nhóm theo K-means sau phân cụm phân cấp để tạo các cụm có các đối tượng tương tự với tỷ lệ cao. Tuy nhiên do tính phức tạp của ảnh y khoa, độ chính xác chỉ đạt 0.666. Nhóm nghiên cứu Khalid Muhammad Junaid (2020) [5] đề xuất hệ thống CBIR với kỹ thuật trích xuất đặc trưng dựa trên RGB, DWT và sử dụng phương pháp phân cụm dữ liệu dựa trên mật độ DBSCAN (Density-Based Spatial Clustering of Applications with Noise). Thử nghiệm được thực thi trên COREL-1K, COREL5K cho độ chính xác lần lượt là 0.91 và 0.87. Các kỹ thuật phân cụm cho thấy tính hiệu quả trong tìm kiếm hình ảnh, độ phức tạp không quá cao, giảm thời gian huấn luyện và tìm kiếm ảnh. Tuy nhiên, mỗi phương pháp phân cụm đều có những ưu nhược điểm riêng, phù hợp với mỗi kiểu dữ liệu ảnh khác nhau. Do đó, nghiên cứu các kỹ thuật phân cụm cho các tập ảnh khác nhau là một bài toán được nhiều nhà khoa học quan tâm.

Từ những phân tích các công trình nghiên cứu trên đã tạo động lực cho nghiên cứu của chúng tôi, các đề xuất của chúng tôi bao gồm: (1) trích chọn đặc trưng hình ảnh bao gồm màu sắc, hình dạng, kết cấu của bộ mô tả MPEG-7; (2) mô hình tìm kiếm ảnh theo nội dung dựa trên các kỹ thuật phân cụm: K-means, DBSCAN và phương pháp lai K-means và DBSCAN. Phần còn lại của bài báo gồm: Mục 2 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; Mục 3 trình bày phương pháp xây dựng mô hình kết hợp giữa phân lớp KNN với Ontology; Mục 4 là phần thực nghiệm nhằm đánh giá và so sánh các cách tiếp cận; Mục 5 kết luận và hướng phát triển tương lai.

2. MÔ HÌNH TÌM KIẾM ẢNH DỰA TRÊN KỸ THUẬT PHÂN CỤM

Trong bài báo, một mô hình tìm kiếm ảnh theo nội dung dựa trên các kỹ thuật phân cụm dữ liệu như thuật toán K-means, DBSCAN và phép lai giữa hai kỹ thuật này được đề xuất. Mô hình được gọi là CBIR_CLUS.

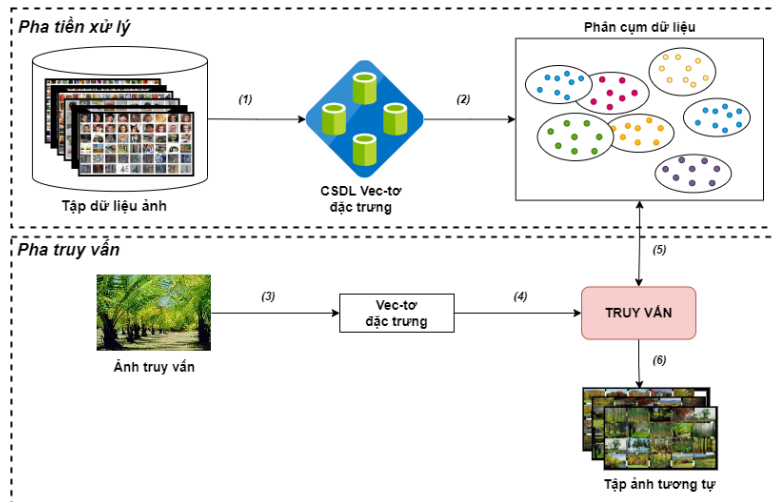
Hình 1 là mô hình tìm kiếm ảnh theo nội dung dựa vào các phương pháp phân cụm dữ liệu, cụ thể như sau: mô tả kiến trúc của mô hình tìm kiếm ảnh theo nội dung bao gồm 2 pha: (a) Pha tiền xử lý và (b) pha truy vấn.

(a) Pha tiền xử lý

- **Bước 1.** Từ bộ dữ liệu ảnh thực hiện trích xuất vec-tơ đặc trưng cấp thấp theo bộ mô tả MPEG-7 (1);
- **Bước 2.** Thực hiện các kỹ thuật học máy khác nhau để phân cụm dữ liệu như K-means, DBSCAN, Phép lai K-means và DBSCAN (2).

(b) Pha truy vấn

- **Bước 1.** Với mỗi hình ảnh đầu vào thực hiện trích xuất vec-tơ đặc trưng cấp thấp (3);
- **Bước 2.** Thực hiện truy vấn (4) trên các phân cụm đã tạo của pha tiền xử lý (5);
- **Bước 3.** Sắp xếp tập ảnh tương tự theo ảnh đầu vào (6) dựa vào độ đo Euclid.



Hình 1. Mô hình tìm kiếm ảnh theo nội dung

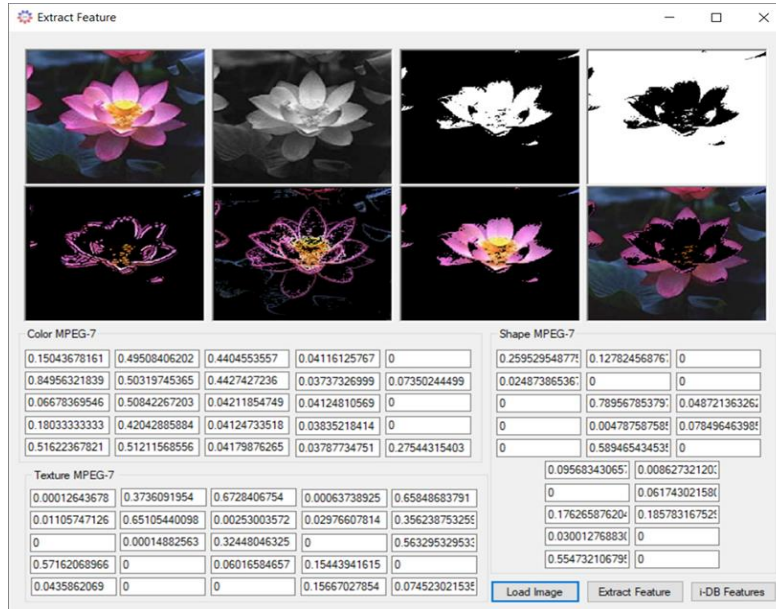
Để mô hình tìm kiếm ảnh hiệu quả cần các thành phần của hệ thống, bao gồm: phương pháp trích xuất đặc trưng, các kỹ thuật phân cụm dữ liệu. Phần sau sẽ trình bày các thành phần của mô hình tìm kiếm ảnh.

3. CÁC THÀNH PHẦN CỦA MÔ HÌNH TÌM KIẾM ẢNH

3.1. Phương pháp trích xuất vec-tơ đặc trưng cấp thấp

Trong bài báo này, một phương pháp trích xuất kết hợp các đặc trưng màu sắc, kết cấu và hình dạng được thực hiện theo bộ mô tả chuẩn MPEG-7 [10]. MPEG-7 là một tiêu chuẩn cung cấp bộ mô tả các đặc trưng của nội dung ảnh phong phú dựa trên danh mục, ngữ nghĩa và cấu trúc. Trong bài báo, bộ mô tả màu chủ đạo DCD (Dominant Color Descriptor) của MPEG-7 được sử dụng để đại diện cho một số lượng nhỏ các thuộc tính màu chủ đạo và cả các đặc điểm thống kê của chúng, bộ mô tả biểu đồ cạnh kết cấu EHD (Edge Histogram Descriptor) được dùng để mô tả sự phân bố không gian của các cạnh và bộ mô tả hình dạng được biểu diễn trên không gian tỷ lệ cong CSS (Curvature Scale Space) của đường bao. Từ các đặc điểm của MPEG-7, có thể thấy ưu điểm của bộ tiêu chuẩn này là: (1) mô tả được nội dung trực quan của hình

ảnh; (2) độ phức tạp thấp và do đó, thời gian tính toán thấp; (3) sử dụng một tập hợp các bộ mô tả dẫn đến một số kết quả tốt hơn so với những kết quả thu được bằng cách sử dụng một bộ mô tả biệt lập. Tập hợp các bộ mô tả khác nhau như màu sắc, kết cấu, hình dạng MPEG-7 được thực hiện để tạo thành một vec-tơ đặc trưng 75 chiều (Hình 2).



Hình 2. Một ví dụ về trích xuất đặc trưng MPEG-7

3.2. Phương pháp lai K-means và DBSCAN

Có nhiều kỹ thuật phân cụm khác nhau, tùy thuộc vào mục đích và tập dữ liệu ảnh để có những hiệu năng khác nhau. Trong bài báo này, chúng tôi nghiên cứu hai kỹ thuật phân cụm phổ biến là K-means và DBSCAN. Mỗi kỹ thuật có các ưu, nhược điểm khác nhau, vì vậy, chúng tôi đề xuất một phương pháp lai K-means và DBSCAN nhằm nâng cao hiệu quả phân cụm dữ liệu. Thuật toán K-Means phân nhóm các đối tượng đã cho vào K cụm (K là số các cụm được xác định trước, K nguyên dương) sao cho tổng bình phương khoảng cách giữa các đối tượng đến tâm cụm là nhỏ nhất. Trong khi đó, phân cụm không gian dựa trên mật độ của các ứng dụng có nhiễu DBSCAN (Density based Spatial Clustering of Application with Noise) là một thuật toán thích hợp với cơ sở dữ liệu có mật độ phân bố dày đặc kể cả có phần tử nhiễu. Điểm chung của những thuật toán phân cụm đều là dựa vào khoảng cách để xác định cụm cho từng quan sát, cập nhật lại cụm dần dần qua các vòng lặp. Ưu nhược điểm của thuật toán K-means và DBSCAN được tổng hợp trong bảng 1.

Từ các so sánh, chúng tôi đề xuất một phương pháp lai giữa K-means và DBSCAN nhằm nâng cao hiệu quả phân cụm dữ liệu. Phương pháp lai với sự kết hợp giữa K-means và DBSCAN được thực hiện nhằm nâng cao hiệu quả phân cụm dữ liệu

ảnh: Thuật toán DBSCAN được sử dụng để phát hiện các cụm có mật độ dữ liệu tối thiểu do người dùng xác định, cũng như để lọc ngoại lệ ban đầu. Các cụm này sau đó được chia thành k nhóm phân tán đồng đều bằng thuật toán K-means. Các trung tâm của các nhóm này sau đó được sử dụng để giảm dữ liệu bằng cách thay thế mỗi nhóm bằng điểm dữ liệu thực tế gần nhất với trung tâm được xác định. Thách thức chung của thuật toán K-means là xác định số cụm đầu vào k . Do đó, DBSCAN cung cấp một ước tính cho mật độ dữ liệu tối thiểu của một cụm thông qua tham số bán kính lân cận của nó, sau đó có thể được sử dụng để đưa ra giá trị hợp lý cho k .

Bảng 1. So sánh K-means và DBSCAN

So sánh	K-Means	DBSCAN
Ưu điểm	<ul style="list-style-type: none"> - Dễ hiểu, dễ triển khai, linh hoạt trong việc chọn độ đo - Hiệu quả cho dữ liệu nhiều chiều, thưa thớt. 	<ul style="list-style-type: none"> - Phù hợp với các bộ dữ liệu nhiễu và dữ liệu ngoại vi - Không yêu cầu chỉ định số lượng cụm
Nhược điểm	<ul style="list-style-type: none"> - Cần xác định trước số cụm và trọng tâm. - Chậm lại khi số lượng dữ liệu hình ảnh nhiều lên - Không hoạt động tốt với dữ liệu nhiễu. 	<ul style="list-style-type: none"> - Cần định nghĩa mật độ của dữ liệu. - Không hoạt động tốt đối với dữ liệu nhiễu, thưa thớt. - Không hiệu quả với những dữ liệu có phân phối đều khắp nơi.

Với đề xuất này của chúng tôi cung cấp một cách tiếp cận mới nhằm giảm dữ liệu linh hoạt dựa trên việc phát hiện các cụm có mật độ khác nhau. Việc thực thi DBSCAN thì các tham số phân nhóm được sửa đổi mỗi lần lặp, để nó phát hiện các vùng có mật độ dữ liệu ngày càng tăng. Khi các điểm được xóa tuần tự qua mỗi lần lặp, bộ nhớ trong sẽ ít hơn được sử dụng để phát hiện và giảm các vùng này. Do đó, thuật toán có khả năng mở rộng hơn đối với các bộ dữ liệu không đồng nhất lớn hơn. Phương pháp lai giữa K-means và DBSCAN được thực hiện như sau:

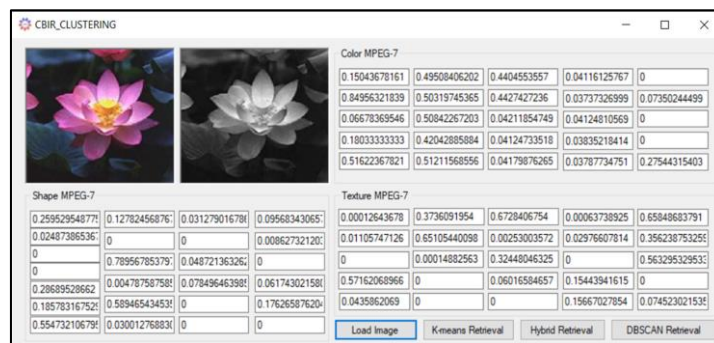
- **Bước 1:** Với tập dữ liệu đầu vào D , áp dụng thuật toán DBSCAN để xác định các vùng R có mật độ dữ liệu cao;
- **Bước 2:** Gắn nhãn tất cả các điểm thuộc $D \setminus R$ là điểm nhiễu và loại bỏ;
- **Bước 3:** Áp dụng thuật toán K-means cho từng cụm R riêng lẻ, tách nó thành k vùng, RC ;
- **Bước 4:** Các điểm dữ liệu $x \in D$ gần nhất đối với các trung tâm cụm K-means được tính toán dựa vào độ đo được gom thành các cụm đầu ra.

Để chứng minh tính hiệu quả của ba phương pháp phân cụm K-means, DBSCAN và phép lai kết hợp K-means và DBSCAN, chúng tôi thu được một số kết quả thực nghiệm sẽ được nêu trong phần tiếp theo.

4. XÂY DỰNG THỰC NGHIỆM VÀ ĐÁNH GIÁ KẾT QUẢ

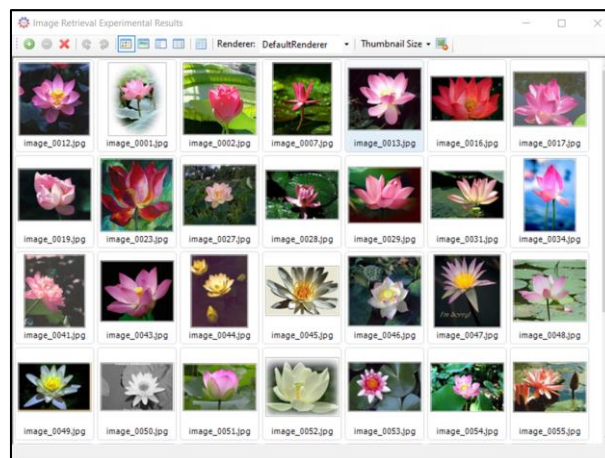
4.1. Tổ chức thực nghiệm

Mô hình tìm kiếm ảnh CBIR_CLUS được xây dựng trên nền tảng dotNET Framework 4.8, ngôn ngữ lập trình C#. Các đồ thị được xây dựng trên Matlab 2015. Cấu hình máy tính của thực nghiệm: Intel(R) Core™ i7-8750H, CPU 2,70GHz, RAM 8GB và hệ điều hành Windows 10 Professional gồm tập dữ liệu ảnh COREL và Caltech-256. Hình 3 là giao diện truy vấn của hệ thống tìm kiếm ảnh theo nội dung dựa trên các phương pháp phân cụm dữ liệu: K-means, DBSCAN, Phương pháp lai (Hybrid Method).



Hình 3. Giao diện tìm kiếm ảnh

Như vậy, với một ảnh đầu vào, đặc trưng cấp thấp của ảnh được trích xuất theo tiêu chuẩn MPEG-7, gồm 75 chiều. Sau khi trích xuất đặc trưng, có thể lựa chọn các kỹ thuật phân cụm khác nhau cho bài toán tìm kiếm ảnh.



Hình 4. Một kết quả tìm kiếm ảnh dựa vào phương pháp lai K-means và DBSCAN

Kết quả của quá trình tìm kiếm là tập các ảnh tương tự được sắp xếp theo độ đo, mô tả như trong **Hình 4**. Đặc trưng cấp thấp theo chuẩn MPEG-7 tập trung vào màu sắc chủ yếu, nên khi tìm kiếm ảnh tương tự, các màu sắc tương tự được ưu tiên. Do đó, với kết quả ở hình 4, ảnh đầu vào là hoa sen do đó tập ảnh tương tự tìm thấy cũng liên quan đến hoa sen, và được sắp xếp theo màu sắc tương tự nhất với ảnh đầu vào.

4.2. Đánh giá kết quả

Thực nghiệm được triển khai trên các tập ảnh COREL và Caltech-256. Trên cơ sở giá trị hiệu suất đã có, các giá trị hiệu suất và thời gian tìm kiếm trung bình của các thư mục ảnh trên tập dữ liệu COREL với các phương pháp phân cụm khác nhau được tổng hợp lần lượt trong **Bảng 2**.

Bảng 2. Hiệu suất của các phương pháp trên tập COREL

Phương pháp	Avg. precision	Avg. recall	Avg. F-measure	Query time (ms)
K-means	0.811275	0.680988	0.739984	80.3492
DBSCAN	0.748855	0.701980	0.723982	68.097
Hybrid	0.862807	0.826651	0.844182	91.734

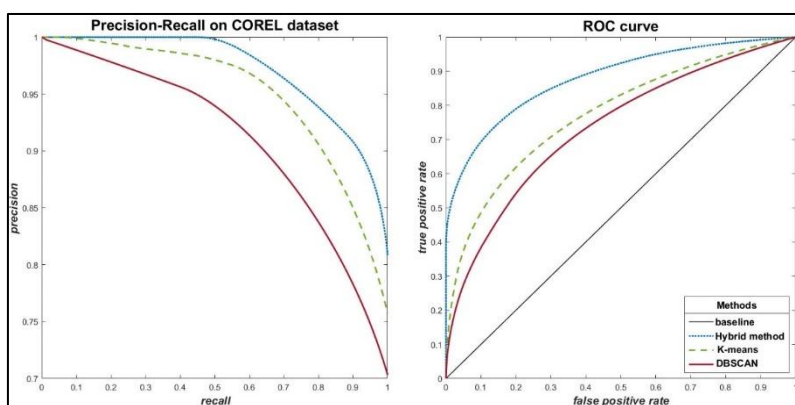
Từ Bảng 2 có thể thấy, độ chính xác tìm kiếm ảnh trên tập ảnh COREL với thuật toán K-means cao hơn so với DBSCAN. Tập ảnh COREL là tập ảnh được phân bố đều, số lượng thưa thớt (1.000 ảnh), ít nhiễu, do đó thuật toán K-means hiệu quả hơn so với DBSCAN. Tuy nhiên DBSCAN là phương pháp phân bố theo mật độ nên thời gian để tìm kiếm sẽ nhanh hơn so với K-means, đồng thời độ phủ cũng tốt hơn. Ngoài ra, độ chính xác tìm kiếm ảnh chưa cao do K-means phụ thuộc vào số lượng cụm đầu vào và cách chọn tâm cụm, các đồ dữ liệu. Vì vậy, kết quả thực nghiệm bộ ảnh COREL với phép lai giữa K-means và DBSCAN đã làm tăng độ chính xác tìm kiếm ảnh (0.862807). Lúc này, phép lai đã loại bỏ những hạn chế của K-means như cần số lượng cụm, tâm cụm đầu vào, cũng như các hạn chế về mặt cấu trúc cân bằng dữ liệu của DBSCAN. Trên cơ sở giá trị hiệu suất đã có, các giá trị hiệu suất và thời gian tìm kiếm trung bình của các thư mục ảnh trên tập dữ liệu Caltech-256 với các phương pháp phân cụm khác nhau được tổng hợp lần lượt trong Bảng 3.

Bảng 3. Hiệu suất của các phương pháp trên tập Caltech-256

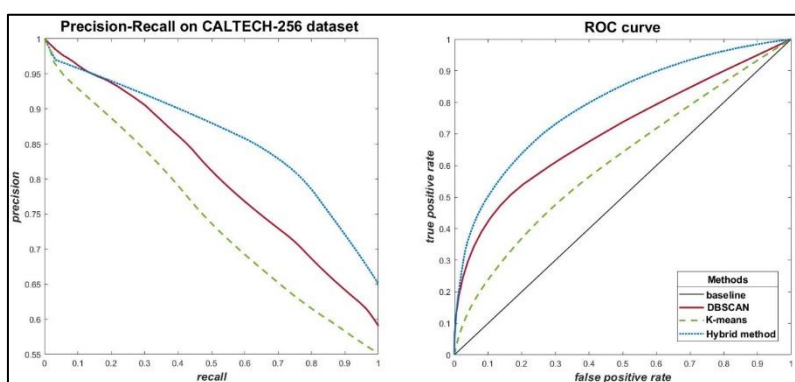
Phương pháp	Avg. precision	Avg. recall	Avg. F-measure	Query time (ms)
K-means	0.612819	0.560605	0.585322	309.991
DBSCAN	0.700851	0.660211	0.679619	204.751
Hybrid	0.803683	0.766553	0.784604	319.204

Với tập ảnh Caltech-256 có số lượng ảnh lớn (30.607) và số lượng chủ đề nhiều (256 chủ đề), sự phân bố hình ảnh cho mỗi chủ đề (thư mục ảnh) không đồng đều, vì vậy K-means không hiệu quả. K-means chỉ tốt với những tập dữ liệu phân bố đồng đều, ít nhiễu. Chính vì vậy, với bộ ảnh Caltech-256, phương pháp DBSCAN hiệu quả hơn vượt trội. DBSCAN xử lý tốt các bộ ảnh phân bố không đồng đều, nhiễu nhiều. Xét về thời gian tìm kiếm ảnh, thuật toán DBSCAN nhanh chóng hơn so với K-means. Tuy nhiên DBSCAN không hiệu quả với tập ảnh lớn nên độ chính xác tìm kiếm ảnh không quá cao. Đồng thời, hiệu suất tìm kiếm ảnh trên tập ảnh Caltech-256 của phương pháp lai cho hiệu quả vượt trội so với hai phương pháp còn lại. Như vậy, phương pháp lai giải quyết được các hạn chế và phát huy được các ưu điểm của K-means và DBSCAN.

Dựa trên các số liệu thực nghiệm, các đồ thị Precision-Recall curve và ROC curve được thực hiện để đánh giá độ chính xác của hệ tìm kiếm ảnh CBIR_CLUS. Ngoài ra, chúng tôi vẽ đồ thị ROC curve để đánh giá mô hình có hiệu quả hay không.



Hình 5. Hiệu suất tìm kiếm ảnh trên tập ảnh COREL



Hình 6. Hiệu suất tìm kiếm ảnh trên tập ảnh Caltech-256

Hình 5 là mô tả hiệu suất tìm kiếm ảnh trên tập ảnh COREL. Hình ảnh này cho thấy, diện tích dưới đường cong AUC trong đồ thị Precision-Recall Curve tương đối lớn, AUC của phương pháp lai (Hybrid method) là lớn nhất, nhỏ nhất là AUC của

DBSCAN. Điều này chứng tỏ với tập ảnh COREL, phương pháp lai là hiệu quả nhất. Với đồ thị ROC curve, các đường cong đều nằm trên đường biên baseline, cho thấy các phương pháp gom cụm đều phân loại hình ảnh vào các cụm tốt. Tuy nhiên, phương pháp lai phân cụm tốt hơn so với hai phương pháp còn lại do AUC của nó là lớn nhất và gần với gốc trái (0,1) của đồ thị nhất. Từ các kết luận này cho thấy, với tập ảnh COREL phương pháp lai là hiệu quả nhất, tiếp đến là K-means, và cuối cùng là phương pháp DBSCAN. Hình 6 mô tả hiệu suất tìm kiếm ảnh trên tập ảnh Caltech-256. Hình ảnh này cho thấy, diện tích dưới đường cong AUC của đồ thị Precision-Recall curve là không lớn, do đây là bộ dữ liệu khó và tương đối lộn xộn, độ nhiễu cao. Trong đồ thị Precision-Recall curve phương pháp lai vẫn có AUC lớn nhất, tiếp theo là DBSCAN và thuật toán K-means là kém hiệu quả nhất. Kết quả tương tự với đồ thị ROC curve. Các đường cong trong đồ thị ROC curve đều nằm trên đường chéo baseline, tuy nhiên các đường cong này khá gần baseline, cho thấy kết quả phân loại cho cụm chưa tốt lắm. Với các kết quả này, cho thấy với tập ảnh phức tạp Caltech-256, phương pháp lai là hiệu quả.

Ngoài ra, để đánh giá độ chính xác và hiệu quả của hệ tìm kiếm CBIR_CLUS, chúng tôi so sánh hiệu suất thu được từ thực nghiệm với các phương pháp của các công trình nghiên cứu khác trên cùng tập dữ liệu ảnh. Bảng 4 là kết quả so sánh hiệu suất trên tập ảnh COREL (1000 ảnh) của phương pháp đề xuất với bốn phương pháp hiện đại khác. Mỗi phương pháp có các cách trích xuất đặc trưng cấp thấp, tổ chức dữ liệu và kỹ thuật tìm kiếm khác nhau, do đó cho kết quả tìm kiếm khác nhau.

Bảng 4. So sánh độ chính xác giữa các phương pháp trên tập ảnh COREL

Phương pháp	Độ chính xác trung bình
HSV+Gabor Wavelet+Edge Detection, 2018 [3]	0.6210
Multi-feature with K-means, 2020 [12]	0.7941
Texture features + CFBPNN, 2021 [4]	0.82
Hybrid method	0.8628

Bảng 5. So sánh độ chính xác giữa các phương pháp trên tập ảnh Caltech-256

Phương pháp	Độ chính xác trung bình
Image Features Information Fusion [1], 2019	0.7422
Feature Fusion and linear discriminate analysis [7], 2019	0.78
SCDMH 64 bit [6], 2021	0.7396
Hybrid method	0.8039

Từ Bảng 4 cho thấy, phương pháp đề xuất của chúng tôi là vượt trội so với các phương pháp khác trên cùng một bộ dữ liệu. Điều này chứng tỏ, phương pháp trích xuất đặc trưng, tổ chức dữ liệu theo phương pháp phân cụm kết hợp K-means và DBSCAN là hiệu quả với tập ảnh COREL. Bảng 5 là kết quả so sánh hiệu suất trên tập ảnh Caltech-256 của phương pháp đề xuất với các phương pháp hiện đại khác. Từ Bảng 5 cho thấy, phương pháp đề xuất của chúng tôi là vượt trội so với các phương pháp khác trên cùng một bộ dữ liệu. Điều này chứng tỏ, phương pháp trích xuất đặc trưng, tổ chức dữ liệu theo phương pháp phân cụm kết hợp K-means và DBSCAN là hiệu quả với tập ảnh Caltech-256.

Như vậy, từ thực nghiệm cho thấy, phương pháp đề xuất cho kết quả tốt hơn khi so sánh với các phương pháp từ các công trình nghiên cứu khác. Vì vậy, các phương pháp đã đề xuất của chúng tôi trong bài báo là cho thấy tính hiệu quả.

5. KẾT LUẬN

Trong bài báo này một mô hình tìm kiếm dữ liệu ảnh theo nội dung dựa trên kỹ thuật phân cụm đã được đề xuất. Đầu tiên, kỹ thuật trích xuất đặc trưng gồm màu sắc, hình dạng, kết cấu dựa trên tiêu chuẩn MPEG-7 được thực hiện, để tạo ra một vec-tơ 75 chiều. Các kỹ thuật phân cụm như K-means, DBSCAN được nghiên cứu và phân tích các ưu nhược điểm của các kỹ thuật này. Từ đó, một phương pháp lai dựa trên K-means và DBSCAN được đề xuất. Để đánh giá hiệu quả của các đề xuất, mô hình tìm kiếm ảnh được thực nghiệm trên các tập ảnh COREL-1K, Caltech-256 với kỹ thuật phân cụm khác nhau. Kết quả thực nghiệm với phương pháp lai K-means và DBSCAN cho độ chính xác vượt trội trên tập COREL-1K và Caltech-256, lần lượt là 0.8628 và 0.8039. So sánh kết quả với các phương pháp khác trên cùng tập ảnh đều cho độ chính xác cao hơn. Điều này chứng tỏ đề xuất trong bài báo là đúng đắn và hiệu quả. Trong định hướng tương lai, chúng tôi tiếp tục nghiên cứu các kỹ thuật phân cụm dữ liệu để nâng cao hiệu quả tìm kiếm ảnh.

LỜI CẢM ƠN

Nhóm tác giả chân thành cảm ơn Trường THPT Nguyễn Huệ – ĐắcĐoa, Gia Lai, 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] Ahmed, K. T., Ummesafi, S., & Iqbal, A. (2019). Content based image retrieval using image features information fusion. *Information Fusion*, 51, 76-99.
- [2] Alsmadi K.M. (2020). Content-based image retrieval using color, shape and texture descriptors and features. *Arabian Journal for Science and Engineering*, 45(4), 3317-3330.
- [3] Ashraf R., Ahmed M., Jabbar S., Khalid S., Ahmad A., Din S., Jeon G. (2018), Content based image retrieval by using color descriptor and discrete wavelet transform, *Journal of medical systems*, 42(3), pp. 1-12.
- [4] Dhingra S., Bansal P. (2021), 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*, pp. 337-349, Springer, Singapore.
- [5] Khalid, M. J., Irfan, M., Ali, T., Gull, M., Draz, U., Glowacz, A., ... & Hussain, S. (2020). Integration of discrete wavelet transform, DBSCAN, and classifiers for efficient content-based image retrieval. *Electronics*, 9(11), 1886.
- [6] Liang, H., Pu, D., Zhang, L., & Sun, Y. (2021, August). An Image Retrieving Method Based on Strong Constrained Manifold-embedded Hashing. In *Journal of Physics: Conference Series* (Vol. 1995, No. 1, p. 012031). IOP Publishing.
- [7] Liu, H., Wang, W., & Jiao, P. (2019, May). Content Based Image Retrieval via Sparse Representation and Feature Fusion. In *2019 IEEE 8th Data Driven Control and Learning Systems Conference (DDCLS)* (pp. 18-23). IEEE.
- [8] Malviya, N., Choudhary, N., & Jain, K. (2017). Content based medical image retrieval and clustering based segmentation to diagnose lung cancer. *Advances in Computational Sciences and Technology*, 10(6), 1577-1594.
- [9] Nazir, R. Ashraf, T. Hamdani, and N. Ali, "Content based image retrieval system by using HSV color histogram, discrete wavelet transform and edge histogram descriptor," in *Proceedings of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*, pp. 1–6, IEEE, Sukkur, Pakistan, March 2018.
- [10] Nguyễn Thị Uyên Nhi (2022), Nâng cao hiệu quả tìm kiếm dữ liệu ảnh theo tiếp cận ngữ nghĩa, luận án tiến sĩ khoa học, Đại học Khoa học, Đại học Huế.
- [11] Ouhda, M., El Asnaoui, K., Ouanan, M., & Aksasse, B. (2018). A content-based image retrieval method based on K-means clustering technique. *Journal of Electronic Commerce in Organizations (JECO)*, 16(1), 82-96.
- [12] Raja R., Kumar S., Mahmood M. R. (2020), Color object detection based image retrieval using ROI segmentation with multi-feature method, *Wireless Personal Communications*, 112(1), pp. 169-192.

AN IMAGE RETRIEVAL MODEL WITH A HYBRID METHOD BETWEEN K-MEANS AND DBSCAN

Ho Thi Nu^{1,2}, Le Manh Thanh², Nguyen Thi Uyen Nhi^{3,*}

¹Nguyen Hue High School, Dak Doa, Gia Lai

²University of Sciences, Hue University

³University of Economics, Danang University

*Email: nhintu@due.edu.vn

ABSTRACT

Content-based image retrieval methods aim to accurately retrieve a set of similar images from a large database based on low-level image features. In this paper, the method to extract feature vectors from color, shape, and texture according to the MPEG-7 standards is implemented. From that, the proposed hybrid method for better clustering for these features based on the combination of K-means and DBSCAN clustering algorithms: (1) DBSCAN algorithm is used to detect clusters with minimum data density and original noise filtering; (2) then the clusters are divided into k uniformly dispersed groups using the K-means algorithm. To prove the effectiveness of the proposal, the article experimented on the image data sets COREL-1K and Caltech-256, with the accuracy of 0.863 and 0.803, respectively. Experimental results compared with other methods show that the proposed system has better performance compared to the other system in terms of accuracy.

Keywords: Content-based image retrieval, DBSCAN, K-means, hybrid method.



Hồ Thị Nữ sinh năm 1980. Bà nhận bằng Cử nhân chuyên ngành tin học ứng dụng Tại trường Đại Học Duy Tân vào năm 2004. Từ năm 2020 bà học Thạc sĩ tại Khoa CNTT, Trường Đại học Khoa học, Đại học Huế. Hiện nay, bà công tác tại trường THPT Nguyễn Huệ, Huyện ĐăkĐoa Gia Lai.



Nguyễn Dũng sinh ngày 13/06/1988 tại Thừa Thiên Huế. Ông tốt nghiệp cử nhân Tin học tại trường Đại học Khoa học, Đại học Huế năm 2010. Năm 2013, ông tốt nghiệp thạc sĩ chuyên ngành Khoa học máy tính tại trường Đại học Khoa học, Đại học Huế. Hiện nay, ông là giảng viên Khoa Công nghệ thông tin, trường Đại học Khoa học, Đại học Huế.



Nguyễn Thị Uyên Nhi sinh năm 1985. Bà nhận bằng cử nhân và thạc sĩ chuyên ngành Khoa học máy tính và kỹ thuật tính toán, tại trường Đại học tổng hợp kỹ thuật Volgograd, Liên bang Nga, lần lượt vào các năm 2008, 2010. Năm 2022, bà nhận học vị Tiến sĩ chuyên ngành Khoa học Máy tính tại trường Đại học Khoa học, Đại học Huế. Hiện nay, bà công tác tại khoa Thống kê – Tin học, Trường Đại học Kinh Tế (DUE), Đại học Đà Nẵng.