Tra cứu ảnh theo nội dung sử dụng tập Pareto và mô hình học thống kê CART

Content-based Image Retrieval using Pareto Fronts Set and CART

Vũ Văn Hiệu, Nguyễn Trường Thắng, Nguyễn Hữu Quỳnh, Ngô Quốc Tạo

Abstract: *Image retrieval systems adopt a combination of multiple features and then total distance measures of particular features for ranking the results. Therefore, the top-ranked images with smallest total distance measures are returned to the users. However, images with smallest partial distance measures which are suitable for users' purpose may not be included in these results. Therefore, partial distance measure should be considered. In this paper, we propose to adopt the Pareto set in the distance measure space. This set assures that the returned results contain not only points with smallest total distance obtained by linear combinations, but also other points have smallest partial distance measures which cannot be found by the linear combination in the distance measure space. Especially, the searching space based on the distance measures is compacted by our algorithm, namely PDFA. This algorithm collects all the Pareto set with different depths, and is efficient for the classification and regression tree (CART). The experimental results on three image collections show the effectiveness of our proposed method.*

Keyword: *Pareto set, classification and regression tree (CART), content-based image retrieval (CBIR), relevance feedback (RF).*

I. GIỚI THIỆU

Từ hai thập kỉ qua, sự xuất hiện của Internet đã thay đổi hoàn toàn cách thức chúng ta tìm kiếm thông tin. Ví dụ, khi làm việc với văn bản, ta chỉ cần đơn giản gõ một vài từ khóa vào máy tìm kiếm Google hay Bing để ngay lập lức có được một danh sách tương đối chính xác các trang web có liên quan. Ta cũng có các hệ thống tương tự với ảnh. Với hệ thống này, bằng cách lấy một ảnh đầu vào từ người sử dụng, hệ thống cố gắng tìm kiếm các ảnh giống nhất trong dữ liệu, rồi trả lại cho người sử dụng. Một cách lý tưởng, sự giống nhau ở đây được định nghĩa dựa trên sự giống nhau giữa các khái niệm được thể hiện trong ảnh. Đây là hệ thống Tra cứu ảnh theo nội dung hay đơn giản là tra cứu ảnh ("content-based image retrieval" viết tắt là CBIR). Lĩnh vực này đã được cộng đồng nhiên cứu quan tâm trong những năm qua, bài báo [6] đã cho thấy điều đó.

Thông thường các hệ thống biểu diễn ảnh trong màu sắc, kết cấu, hình dạng và các đặc trưng bề mặt. Các hàm tìm kiếm được xây dựng để tra cứu theo sự quan tâm. Bài báo này sử dụng kết hợp nhiều biểu diễn đặc trưng được miêu tả như trong [2, 5, 7, 9, 22, 23, 24, 26]. Trong xếp hạng các kết quả trả về cho người dùng thông thường sử dụng khoảng cách toàn cục bằng kết hợp tuyến tính khoảng cách cục bộ theo biểu diễn đặc trưng thành phần. Một ảnh được xếp thứ hạng cao hơn nếu và chỉ nếu độ đo khoảng cách toàn cục là nhỏ hơn.

Ví dụ I.1. *Giả sử chúng ta có hai đặc trưng màu (C) và kết cấu (T). Độ đo khoảng cách của ba đối tượng o_1, o_2, o_3 tương ứng với truy vấn Q là* $D_Q^{(C)}(o_1) = 0.6$, $D_Q^{(T)}(o_1) = 0.3$, $D_Q^{(C)}(o_2) = 0.5$, $D_Q^{(T)}(o_2) = 0.2$, $D_Q^{(C)}(o_3) = 0.45$, $D_Q^{(T)}(o_3) = 0.35$. *Khoảng cách toàn cục áp dụng kết hợp tuyến tính độ đo khoảng cách thành phần của các đặc trưng màu và kết cấu tương ứng là* $D_Q(o_1) = 0.9$, $D_Q(o_2) = 0.7$, $D_Q(o_3) = 0.8$. *Dễ dàng xếp hạng độ đo khoảng cách là o_2, o_3, o_1. Khi không kết hợp tuyến tính độ đo khoảng cách toàn cục, xếp hạng dựa vào độ đo khoảng cách thành phần chúng ta chỉ có thể xếp hạng được o_1 và*

o_2, đối tượng o_3 không thể so sánh được với hai đối tượng còn lại.

Như vậy cách xếp hạng sử dụng tổng toàn bộ độ đo khoảng cách của các thành phần trong kết quả cuối cùng còn nhiều vấn đề cần xem xét và cải tiến.

Trong các nghiên cứu [15, 36] sử dụng kỹ thuật tối ưu đa mục tiêu dựa vào kiến trúc Pareto, định nghĩa độ đo toàn cục như một kết hợp tối ưu tuyến tính của các hàm khoảng cách thành phần. Các nghiên cứu này chỉ sử dụng cách tiếp cận Pareto trong việc lựa chọn kết quả cuối cùng như một bài toán tối ưu đa mục tiêu như trong nghiên cứu [12].

Không giống như cách tiếp cận trên, chúng tôi sử dụng Pareto như một bài toán tiền xử lý dữ liệu (rút gọn tập mẫu). Qua đó, không gian tìm kiếm trên tập độ đo khoảng cách với truy vấn được thu gọn nhất của tập Pareto. Tập thu gọn này được sử dụng như dữ liệu đầu vào giúp cho bộ máy phân lớp hoạt động hiệu quả hơn. Các phương pháp thống kê, như hồi quy thực hiện tốt hơn với tập mẫu nhỏ như số mẫu huấn luyện chỉ có được dựa vào đánh giá của người dùng trong một số lần phản hồi. Do đó chúng tôi kết hợp sử dụng mô hình cây dự báo hồi quy (Classification and Regression Tree - CART) để dự báo phân lớp trên tập mẫu được thu gọn này.

Phần còn lại của bài báo được tổ chức như sau. Phần hai, một số nghiên cứu liên quan sử dụng phương pháp tối ưu Pareto và kỹ thuật máy học. Phần ba là đề xuất phương pháp giảm không gian mẫu của tập độ đo khoảng cách dựa vào tiếp cận tập Pareto và mô hình cây hồi quy phân lớp. Các kết quả thực nghiệm trong phần bốn. Kết luận và hướng nghiên cứu tương lai ở phần năm.

II. NGHIÊN CỨU LIÊN QUAN

II.1. Phương pháp tối ưu Pareto

Để giải bài toán tối ưu nhiều tác giả áp dụng phương pháp thích nghi dựa trên giải thuật di truyền [8, 11, 32]. Các nghiên cứu này đảm bảo không bỏ sót các ảnh có ít nhất một độ đo khoảng cách thành phần với truy vấn là nhỏ nhất. Tuy nhiên, các nghiên cứu này không thay đổi hoặc rút gọn được không gian tìm kiếm. Arevalillo-Herraez và cộng sự [1] sử dụng phương pháp tối ưu Pareto và cách tiếp cận NSGA-II để sắp xếp tập có độ đo khoảng cách không trội (non-dominated). Nghiên cứu này không đưa ra tập rút gọn không gian tìm kiếm. Hsiao và cộng sự [12] sử dụng Pareto độ sâu (dựa trên nghiên cứu của Torlone và cộng sự [31]). Nghiên cứu này sử dụng cách xếp hạng EMR (efficient manifold ranking) theo các mục tiêu như các truy vấn độc lập. Để lựa chọn kết quả cuối cùng, họ sử dụng nhiều điểm rìa Skyline cho xếp hạng các đối tượng theo các rìa. Tối ưu Pareto được sử dụng rộng rãi trong cộng đồng học máy [10]. Các hệ thống CBIR sử dụng bộ máy phân lớp ít sử dụng cách tiếp cận Pareto để giảm tập dữ liệu và đây chính là yếu tố quan trọng giúp cải thiện các bộ máy phân lớp dữ liệu.

II.2. Tra cứu ảnh theo nội dung dựa vào các mô hình học máy

Phản hồi liên quan (Relevance feedback, hay viết tắt là RF) được sử dụng để giảm khoảng cách ngữ nghĩa giữa khái niệm mức cao và đặc trưng mức thấp trong miêu tả ảnh. Thông thường người dùng không dễ dàng dùng trực giác nhận biết ảnh dựa trên đặc trưng mức thấp như màu sắc và hình dạng. Một vấn đề khác liên quan tới nhận thức chủ quan về hình ảnh, người khác nhau có thể có nhận thức trực quan khác nhau về cùng một ảnh. Những ảnh khác nhau có những ý nghĩa khác nhau hoặc có tầm quan trọng khác nhau với mỗi người. Ví dụ, cho một ảnh con chim bay trên bầu trời, trong khi người này có thể quan tâm đến con chim, người khác lại quan tâm đến bầu trời. Do tầm quan trọng của các đặc trưng cụ thể là khó xác định nên sự kết hợp tuyến tính các khoảng cách đặc trưng thành phần có thể dẫn đến bỏ sót các thành phần quan trọng trong kết quả trả về người dùng.

Kỹ thuật phản hồi liên quan sử dụng máy học cũng đã được nghiên cứu trong nhiều bài báo những năm gần đây. SVM-AL [30] là một nghiên cứu tiên phong và có đóng góp quan trong trong cộng đồng CBIR. Những giới hạn của nó đã được giải quyết bằng các giải pháp mới. Jiang và cộng sự [14] cải tiến hiệu năng của SVM-AL sử dụng dụng kỹ thuật AdaBoost. Tuy nhiên chỉ đơn thuần sử dụng AdaBoost thì khó cải tiến

được SVM. Các phương pháp phân lớp dựa trên kỹ thuật SVM thường ít hiệu quả khi không có mẫu huấn luyện trước, hay số mẫu được huấn luyện rất ít có được sau một số lần phản hồi của người dùng. AdaBoost được xem như ý nghĩa tăng cường cho thuật toán học yếu. Từ cải tiến AdaBoost gốc, kỹ thuật boosting đã được áp dụng trong các hệ thống CBIR như các nghiên cứu [16, 29, 34]. Tuy nhiên các kỹ thuật dựa trên AdaBoost thường phân lớp chậm, điều này là hạn chế khi áp dụng phân lớp trong các ứng dụng tra cứu ảnh. Một nhược điểm của các phương pháp trên là thường "overfit" khi phân lớp, dẫn đến kết quả không cao.

Trong một số bài báo, kỹ thuật cây quyết định (học giám sát) như C4.5, ID3 được sử dụng trong phản hồi liên quan để phân lớp các ảnh trong cơ sở dữ liệu ảnh vào hai lớp (liên quan/không liên quan) phụ thuộc vào tương tự với ảnh truy vấn như nghiên cứu của MACARTHUR và cộng sự [18]. Kỹ thuật CART do Breiman và cộng sự [4] xây dựng một cấu trúc cây bằng cách phân hoạch đệ quy không gian thuộc tính đầu vào. Một tập các luật quyết định có thể thu được theo các đường dẫn từ gốc tới các lá của cây. So sánh với các phương pháp học khác, cây quyết định học khái niệm đơn giản, mạnh với các đối tượng không đầy đủ và nhiễu các đặc trưng đầu vào.

III. KỸ THUẬT ĐỀ XUẤT

III.1. Giảm không gian tìm kiếm dựa vào tập Pareto

Tập Pareto hoặc rìa Pareto là một tập con của tập các điểm thoả hiệp của các lời giải trong đó chứa tất cả các điểm mà có ít nhất một mục tiêu tối ưu trong khi giữ nguyên mọi mục tiêu khác. Các điểm đó được gọi là các điểm tối ưu Pareto[1].

Bài toán tối ưu trên miền không gian độ đo khoảng cách của truy vấn với các mẫu trong cơ sở dữ ảnh phát biểu như sau:

$$\begin{cases} \min\ D_Q^t(I), & t \in \{1,...,T\} \\ s.t.\ I \in \left\{E_i^F\right\}, & i \in \{1,...,M\} \end{cases}, \qquad (1)$$

trong đó truy vấn Q biểu diễn bởi một tập T đặc trưng và các phần tử ảnh I của tập dữ liệu $\left\{E^F\right\}$ gồm M ảnh bao gồm các đặc trưng tương ứng như truy vấn. $D_Q^t(I) = D(Q_t, I_t)$ là độ đo khoảng cách giữa đặc trưng thứ t biểu diễn bởi các thành phần Q_t và I_t. Ký hiệu $D_Q(I) = \{D_Q^t(I)\} = \{D^t(Q^t, I^t)\}_{1 \le t \le T}$ là tập T độ đo khoảng cách của ảnh I và truy vấn Q.

Để tìm tập các đối tượng tối ưu trên miền không gian độ đo khoảng cách, dựa trên quan hệ trội tìm tập tối ưu Pareto theo định nghĩa 3.1.

Định nghĩa 3.1. *(Trội Pareto trên độ đo khoảng cách) Cho truy vấn Q, xác định một quan hệ trội (ký hiệu là f) trên tập độ đo khoảng cách của hai ảnh* I_1 *và* I_2 *như sau:*

- *Quan hệ trội yếu, ký hiệu là* $D_Q(I_1) \succcurlyeq D_Q(I_2)$ *khi và chỉ khi:*

$$\begin{cases} \forall t, 1 \le t \le T, D_Q^t(I_1) \le D_Q^t(I_2), \\ \exists t_0, 1 \le t_0 \le T, D_Q^{t_0}(I_1) < D_Q^{t_0}(I_2), \end{cases} \qquad (2a)$$

- *Quan hệ trội mạnh, ký hiệu là* $D_Q(I_1) \succ D_Q(I_2)$ *khi và chỉ khi:*

$$\forall t, 1 \le t \le T, D_Q^t(I_1) < D_Q^t(I_2), \qquad (2b)$$

Ví dụ III.1: *Xét ví dụ I.1 ta có,* $D_Q(o_2) \succ D_Q(o_1)$.

Định nghĩa 3.2. ***(****Rìa Pareto****)*** *Cho* $\forall I \in \{E^F, D_Q(I)\}$ *nếu* $\nexists I_0 \in \{E^F, D_Q(I_0)\}$ *mà* $D_Q(I_0) \succ D_Q(I)$ *thì* $D_Q(I)$ *được gọi là điểm tối ưu Pareto. Tập các điểm tối ưu Pareto (không trội) của* $\left\{E^F, D_Q(I)\right\}$ *được gọi là rìa Pareto đầu tiên, ký hiệu là* PF^1.

Tập Pareto chứa tất cả các điểm không trội với các điểm khác trong $\left\{E^F, D_Q(I)\right\}$. Tập này chứa tất cả các phần tử tối thiểu hoá bằng cách kết hợp tuyến tính, nhưng cũng chứa các phần tử khác mà không tìm thấy nếu kết hợp tuyến tính.

[1] http://en.wikipedia.org/wiki/Pareto_efficiency

Mệnh đề 3.1. $\forall I \in \left\{E^F, D_Q(\mathrm{I})\right\}$, *nếu:*

$\exists t_0, 1 \le t_0 \le T, D_Q^{t_0}(I) = \underset{I' \in \{E^F\}}{Min} D_Q^{t_0}(\mathrm{I}'),$ *thì* $I \in PF^1$.

Chứng minh: Giả sử

$I \notin PF^1 \Rightarrow \exists I' \in E^F$, $\forall 1 \le t \le T$, $D_Q^t(I') < D_Q^t(I) \Rightarrow$

$D_Q^{t_0}(I') < D_Q^{t_0}(I)$, vô lý vì $D_Q^{t_0}(I) = \underset{I' \in \{E^F\}}{Min} D_Q^{t_0}(\mathrm{I}')$ ■

Định nghĩa 3.3 *(Mức rìa Pareto) Rìa Pareto thứ i được xây dựng:*

$$PF^i = PF^1\left(\left\{\left\{E^F, \{D_Q^t(I)\}_{1 \le t \le T}\right\}\right\} \setminus \left(\cup_{j=1}^{i-1} PF^j\right)\right) \quad (3)$$

Ví dụ III.2. *Xét quan hệ trội trên ví dụ I.1:* $D_Q(o_2) \succ D_Q(o_1)$, *thì ta có* $PF^1 = \{o_1, o_3\}$, $PF^2 = \{o_2\}$.

Tập các điểm Pareto nhiều mức rìa (mức rìa tăng dần) được gọi là Pareto depth.

Mệnh đề 3.2.

$(i) \forall I_1, I_2 \in PF^1 (l \ge 1) \Rightarrow I_1 \not\succ I_2, I_2 \not\succ I_1,$

$(ii) \forall I \in PF^{1+1} (l \ge 1) \Rightarrow \exists J \in PF^1, \mathrm{D}_Q(J) \succ D_Q(I)$

Chứng minh: *(i) được suy từ định nghĩa* PF^1.

(ii) Giả sử $I \in PF^{l+1} \Rightarrow$

$$\begin{cases} I \notin PF^l \Rightarrow \exists J \in E^F \setminus \underset{i=1}{\overset{l-1}{\cup}} PF^i, \mathrm{D}_Q(J) \succ D_Q(I) \\ \forall I' \notin \underset{i=1}{\overset{l}{\cup}} PF^i, \mathrm{D}_Q(I) \not\succ \mathrm{D}_Q(I'), \mathrm{D}_Q(I') \not\succ \mathrm{D}_Q(I) \end{cases}$$

$\Rightarrow D_Q(J) \succ D_Q(I), J \in PF^l$ ■

Thuật toán *PDFA* tìm tập rìa Pareto nhiều mức sâu hay tập Pareto sử dụng mệnh đề 3.1 và 3.2.

Thuật toán *PDFA* (rìa Pareto nhiều mức sâu)

Đầu vào: Tuple= $\{\mathrm{D}_Q^t(\mathrm{I}_i)\}_{t=1}^T$, $1 \le i \le N$, $1 \le t \le T$

/*Danh sách sắp thứ tự Tuple có T danh sách N ảnh, mỗi ảnh có T giá độ đo khoảng cách theo từng đặc trưng với truy vấn Q */

k /* Số lượng mẫu trong tập rìa Pareto */

Đầu ra: *ListResult* /*Tập rìa Pareto */

/* Biến trung gian */

Result=0; *PF*=*PF_Next*= $\varnothing$; *aTupleMax* =0; *aMax*=0;

/* Khởi tạo */

1. *TopTuple* = 0;
2. While (*Result* <*k*)
3. While $\nexists$ $I_i \in PF$ mà ($\mathrm{D}_Q(\mathrm{I}_i)$ f *aTupleMax*) $\wedge$ (*Result* <*k*)
4. For *t*=1 to *T*
5. Lấy ra ảnh I_i chưa được lấy trong danh sách đã sắp thứ tự $Tuple_t$ cùng với *T* độ đo khoảng cách $D_Q(\mathrm{I}_i)$;
6. IF aMax< $\mathrm{D}_Q(\mathrm{I}_i)$ aMax = $\mathrm{D}_Q^t(\mathrm{I}_i)$;
7. isDominated=false;
8. While (not(isDominate)) $\wedge$ $\exists$ $I_j \in PF$ chưa được so sánh với I_i)
9. IF($\mathrm{D}_Q(\mathrm{I}_i) \succ \mathrm{D}_Q(\mathrm{I}_j)$) chuyển I_j từ *PF* vào *PF_Next;*
10. End IF;
11. IF($I_j \succ I_i$)
12. isDominated = true;
13. Chèn I_i vào *PF_Next*;
14. End IF
15. End While
16. IF not(isDominated) chèn I_i vào *PF*;
17. $aTupleMax_t$ =*aMax*; /* Đặt lại ngưỡng ở *t* */
18. Đưa các ảnh $I_i \in PF$ mà *aTupleMax* $\not\succ$ $\mathrm{D}_Q(\mathrm{I}_i)$ vào *ListResult*;
19. Result = Result+1;
20. End For
21. End While
22. IF (Result<k)
23. PF = *PF_Next; PF_Next*= $\varnothing$;
24. For all I_i, I_j $\in$ PF mà $\mathrm{D}_Q(\mathrm{I}_i)$ f $\mathrm{D}_Q(\mathrm{I}_j)$ thì chuyển I_j sang *PF_Next*;
25. Đưa các ảnh $I_i \in PF$ mà *aTupleMax* $\not\succ$ $\mathrm{D}_Q(\mathrm{I}_i)$ vào *ListResult*;
26. End IF
27. End While

Sau khi sắp xếp *T* danh sách, thuật toán chỉ thực hiện trên phép so sánh, lần lượt lấy từng ảnh chưa được đánh dấu trong mỗi danh sách so sánh tập độ đo khoảng cách với tập giá trị ngưỡng *aTupleMax*. Tập giá trị ngưỡng *aTupleMax* được thiết lập sao cho mỗi thành phần của nó có giá trị cao nhất trong tất cả các điểm Pareto đã tìm được. Thuật toán PDFA sử dụng định nghĩa 3.3 kết hợp với tập giá trị *aTupleMax* để so sánh lấy ra các điểm Pareto theo nhiều mức, quá trình tiếp tục đến khi số điểm cần lấy đạt được *k* điểm, được gọi là tập rìa Pareto nhiều mức sâu. Quá trình tăng dần mức rìa (độ sâu) đến khi tìm đủ số điểm theo độ sâu

hoặc hết cơ sở dữ liệu. Thuật toán có độ phức tạp là $O(n)$, trong đó các phép toán được sử dụng chỉ toàn các phép so sánh nên thời gian thực hiện nhanh.

Theo mệnh đề 3.1, tập rìa Pareto nhiều mức sâu chứa các điểm có độ đo khoảng cách tối thiểu theo thành phần và tối thiểu theo cách kết hợp tuyến tính. Theo mệnh đề 3.2, các điểm trong cùng một mức sâu thì không thể so sánh với nhau, các điểm ở mức trong sâu hơn thì bị làm trội ở mức ngoài. Như vậy tập Pareto depth bao được các điểm liên quan về độ đo khoảng cách mức thấp. Theo trực giác đây là tập khả năng liên quan cao nhất. Tuỳ thuộc số mức rìa, tập này có số mẫu nhỏ hơn toàn bộ cơ sở dữ liệu.

Phản hồi liên quan là cầu nối giúp giảm khoảng trống giữa đặc trưng mức thấp biểu diễn với khái niệm mức cao của người dùng. Trong quá trình phản hồi, người dùng chọn các ảnh như "liên quan", "không liên quan". Kỹ thuật đề xuất sử dụng các ảnh liên quan như một truy vấn độc lập, mỗi truy vấn này lại thu được một tập rìa Pareto nhiều mức sâu.

Định nghĩa 3.4 phát biểu hợp của các rìa Pareto nhiều mức sâu. Kết quả phép hợp rìa Pareto nhiều mức sâu sẽ được sử dụng trong thuật toán PCART ở phần sau.

Định nghĩa 3.4 *(Hợp Pareto) Tập kết hợp của các rìa Pareto được gọi là hợp Pareto, ký hiệu là* PF^+*, thoả mãn:*

$$PF^+ = PF^{l+1} \stackrel{\text{def}}{=} \left\{ \forall I \in \left\{ E^F, D_Q(I) \right\} \setminus \bigcup_{1 \le k \le l} PF^k \right.$$

$$\left. / \forall J \in \left\{ E^F, D_Q(\mathrm{J}) \right\} \setminus \bigcup_{1 \le k \le l} PF^k, \neg D_Q(I) \succ D_Q(\mathrm{J}) \right\}$$

III.2. Cây dự báo hồi quy (CART)

Giả sử mỗi ảnh tương ứng là một mẫu trong không gian độ đo khoảng cách với truy vấn *Q* và tập tất các mẫu $\{\mathrm{D}_Q(\mathrm{I}_i)\}$ có kích thước *M*. Từ kết quả tập hợp rìa Pareto nhiều mức sâu (Thuật toán PDFA) gọi là tập PF_l, ký hiệu *l* là mức sâu của rìa Pareto, thông thường chúng tôi lựa chọn $1 \le l \le L$, và $L \le 20$, $\#\mathrm{PF}_l \ll \#\{\mathrm{D}_Q(\mathrm{I}_i)\}$. Theo mệnh đề 3.2, tập PF_l chứa các đối tượng tối thiểu trên một số bộ nên gồm nhiều các đối tượng liên quan, *k* đối tượng tốt nhất theo các rìa Pareto (gọi là tập *NB* và $NB \subseteq PF_l$) được hiển thị. Người dùng chọn đối tượng liên quan được gán nhãn là "+1" và đưa vào tập NB^+, các mẫu không liên quan được gán nhãn "-1" và đưa vào tập NB^-. Quá trình tiếp tục như vậy ở lần phản hồi sau.

Lời giải của bài toán học máy nằm trong dữ liệu huấn luyện xác định (truy vấn và các ảnh được đánh giá), suy luận một khái niệm từ dữ liệu này, và đưa ra các trường hợp khác từ một cơ sở dữ liệu sao cho phù hợp với khái niệm này (trả về một tập các ảnh). Bài toán học máy có thể được xem như một bài toán phân hai lớp được đề xuất ban đầu trong [27]. Kỹ thuật này áp dụng cho phân lớp ảnh như sau: cho một tập dữ liệu huấn luyện được trả về từ các ảnh tra cứu, tập này đưa tới cho người dùng gán nhãn, sau đó được đưa vào một mô hình học. Một hạn chế của bài toán CBIR là dữ liệu huấn luyện không có trước, dữ liệu huấn luyện chỉ có sau khi người dùng gán nhãn trong các lần lặp phản hồi đối với từng truy vấn. Cách tiếp cận cây quyết định rất hiệu quả trong bài toán phân lớp này. CART đưa ra điều kiện phân bố của y cho x, trong đó x biểu diễn một véc tơ của các dự báo $[\mathrm{x}_1,\mathrm{x}_2,...,\mathrm{x}_n]$.

Cho một tập độ đo khoảng cách mỗi ảnh với truy vấn $D = \{\mathrm{D}_Q(\mathrm{I}_1), ..., \mathrm{D}_Q(\mathrm{I}_n)\}$, trong đó:

$D_Q(I_k) = \{D_Q^1(I_k), ..., D_Q^T(I_k)\}$ bao gồm *T* các bộ độ đo khoảng cách như là các thuộc tính.

Một phương pháp tốt nhất cho lựa chọn các phân hoạch nhiều cách dựa vào thống kê tầm quan trọng [3]. Việc tách được thực hiện quanh việc xác định điểm tách tốt nhất. Ở mỗi bước tìm kiếm toàn bộ được thực hiện để xác định phép tách tốt nhất. Điều đó thực hiện như sau:

$$f\left(\frac{s}{t}\right) = 2P_L P_R \sum_{j=1}^{m} |\,\mathrm{P}(\mathrm{C}_j \mid \mathrm{t}_L) - \mathrm{P}(\mathrm{C}_j \mid \mathrm{t}_R)\,|, \qquad (4)$$

trong đó *t* là nút hiện tại, *s* là các thuộc tính, *L* và *R* là cây con bên trái và phải của nút hiện tại. P_L, P_R là xác suất mà bộ trong tập huấn luyện sẽ ở bên trái hay bên phải của một cây:

$$P_L \text{ hoặc } P_R = \left| \frac{D_Q^t(I) \text{ của các cây con}}{D_Q^t(I) \text{ trong tập huấn luyện}} \right|$$

$$\mathrm{P}(\mathrm{C}_j \mid \mathrm{t}_L) \text{ hoặc } \mathrm{P}(\mathrm{C}_j \mid \mathrm{t}_R) = \left| \frac{D_Q^t(I) \text{ của } C_j \text{ trong các cây con}}{D_Q^t(I) \text{ trong nút mục tiêu}} \right|$$

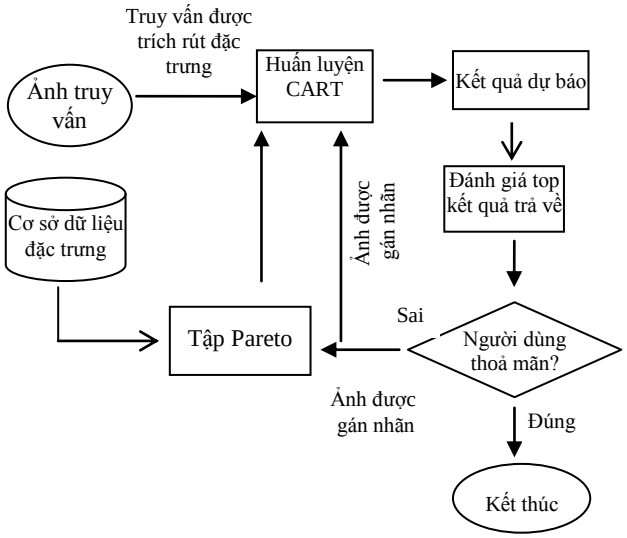
Trong đó $\mathrm{P}(\mathrm{C}_j \mid \mathrm{t}_L)$ hoặc $\mathrm{P}(\mathrm{C}_j \mid \mathrm{t}_R)$ là xác xuất mà một bộ ở trong lớp C_j ở bên trái hoặc bên phải của các cây con. Trong mỗi bước, chỉ một tiêu chuẩn được lựa chọn tốt nhất trong tất cả các tiêu chuẩn có thể có.

Dưới đây là thuật toán PCART thực hiện dự báo phân lớp theo mô hình CART sử dụng tập Pareto. Để tăng cường số mẫu trên tập Pareto và tránh được vấn đề gặp phải số ảnh liên quan nằm rải rác trong không gian vật lý (là tập các véc tơ nhiều chiều của khoảng cách mỗi ảnh với truy vấn), hợp các rìa Pareto nhiều mức sâu được sử dụng trong thuật toán này.

Thuật toán PCART

Input: $\{\mathrm{D}_Q(\mathrm{I}_i)\}$, $1 \le i \le N$, /* Tập độ đo khoảng cách của mỗi ảnh trong cơ sở dữ liệu với truy vấn */
k; /* Số lượng mẫu trong tập phủ Pareto */
Output: Ảnh thoả mãn nhu cầu tìm kiếm
1. Khởi tạo:
$listNB^+ \leftarrow Q$; /* Truy vấn ban đầu được nhãn dương */
$listNB^- \leftarrow \varnothing$; /* Tập mẫu được gán nhãn âm ban đầu * /
$PF \leftarrow \varnothing$; /* Tập Pareto ban đầu */
2. While người dùng chưa thoả mãn
2.1 For each Q_j in $listNB^+$
Tìm tập các điểm rìa Pareto nhiều mức (xem thuật toán 1 và định nghĩa 3.3)
$PF \leftarrow PF \cup \mathrm{Pareto}(\{\mathrm{D}^t_{Q_j}(\mathrm{I}_i)\}_{t=1}^{T}, \mathrm{k})$;
2.2. Chuẩn bị dữ liệu huấn luyện cho CART (X_i, y_i),
$X_i \in listNB^+ \cup listNB^-, y_i = \begin{cases} +1, \text{if } X_i \in listNB^+ \\ -1, \text{if } X_i \in listNB^- \end{cases}$
Xây dựng hàm dự báo phân lớp sử dụng phương trình (3.4) thu được $f\left(\frac{s}{t}\right)$
2.3. $aPredictRF(I_i) = f\left(\frac{s}{t}\right)$; / * aPredictRF($\mathrm{I}_i$) là giá trị dự báo phân lớp cho ảnh I_i trong tập Pareto */
Sắp xếp các ảnh trong *PF* theo giá trị dự báo *aPredictRF*;
2.4. $S_k \leftarrow$ k ảnh đầu tiên trong *PF*;
2.5. Người dùng đánh giá các ảnh theo nhận thức về sự liên quan và không liên quan
$NB^+ \leftarrow S_k$;
$NB^- \leftarrow S_k$;
$listNB^+ \leftarrow listNB^+ \cup NB^+$;
$listNB^- \leftarrow listNB^- \cup NB^-$;
3 End While

Trong thuật toán PCART, *aPredictRF* là một danh sách lưu các giá trị dự báo sử dụng phương trình (4). Thuật toán PCART sử dụng các ảnh liên quan như truy vấn độc lập để mở rộng truy vấn và mở rộng tập rìa Pareto theo nhiều mức sâu bằng cách sử dụng định nghĩa 3.3 và thuật toán *PDFA*. Thuật toán có độ phức tạp là $O(n^2)$. Mô hình đề xuất được mô tả như Hình 1.



Hình 1. Sơ đồ hệ thống đề xuất

IV. THỰC NGHIỆM

Để đánh giá hiệu năng của phương pháp đề xuất, một số thực nghiệm đã được thiết kế và cài đặt. Đề xuất của chúng tôi được so sánh với phương pháp tra cứu ảnh có sử dụng kỹ thuật phân lớp như SVM chuẩn, học tăng cường i.Boost [29] (AdaBoost), và phương pháp phản hồi liên quan tiên tiến MARS. Đây là các phương pháp tiên tiến thường được sử dụng để phân lớp dữ liệu, tuy nhiên với dữ liệu gặp nhiều

nhiễu như "khoảng trống ngữ nghĩa" trong CBIR và số mẫu huấn luyện không có trước nên các phương pháp này gặp nhiều khó khăn. Kỹ thuật phân lớp CART hiệu quả với dữ liệu huấn luyện nhỏ như số các mẫu có được trong một số lần phản hồi.

IV.1. Các miêu tả ảnh

Chúng tôi lựa chọn bộ đặc trưng kết hợp gồm sáu đặc trưng mức thấp và hàm khoảng cách sử dụng tương ứng được miêu tả trong Bảng 1. Các biểu diễn gồm ba kiểu đặc trưng màu sắc, kết cấu và hình dạng, đây là những đặc trưng được sử dụng rất nhiều trong các nghiên cứu tra cứu ảnh hoặc nhận dạng.

Bảng 1. Các miêu tả ảnh trong thực nghiệm

Miêu tả	Loại	Kích thước	Hàm khoảng cách
HSV Histogram [28]	Color	32	L1
Color moments [36]	Color	6	L2
Color auto correlogram [13]	Color	64	L1
Lọc Gabor [35]	Texture	48	L2
Wavelet moments [33]	Texture	40	L2
Gist [20]	Shape	512	L2

Chúng tôi sử dụng ba tập ảnh để thực nghiệm. Các ảnh trong mỗi tập được tổ chức theo chủ đề bằng nhận thức chủ quan của con người về tính tương tự ngữ nghĩa. Cụ thể các tập ảnh như sau:

- **Db1**. Đây là tập COREL [17] gồm 1000 ảnh được chia vào 10 chủ đề: biển, Châu Phi, hoa hồng, ngựa, núi, thức ăn, xe buýt, khủng long, toà nhà và voi.
- **Db2**. Tập Oxford Buildings [21] bao gồm 5062 ảnh được lấy ra từ Flickr. Tập này gồm 11 chủ đề địa danh khác nhau gồm 2560 ảnh, mỗi chủ đề sử dụng 5 truy vấn. Tập truy vấn gồm 55 ảnh được sử dụng để đánh giá theo các chủ đề: All Souls Oxford, Ashmolean Oxford, Balliol Oxford, Bodleian Oxford, Christ Church Oxford, Cornmarket Oxford, Hertford Oxford, Keble Oxford, Magdalen Oxford, Pitt Rivers Oxford, Radcliffe Camera Oxford.
- **Db3**. Đây là tập con của tập Caltech 101 [10], gồm 101 chủ đề, mỗi chủ đề có khoảng từ 40 đến 800 ảnh. Chúng tôi sử dụng 10 chủ đề đó là: kiến, cá, gấu, khủng long, súng thần công, bình nước, đàn măng-đô-lin, mỏ lết, ghế, cái ô.

Trên **Db1** và **Db3** 10% số ảnh được lấy ngẫu nhiên ở mỗi chủ đề làm truy vấn và đánh giá chất lượng tra cứu trên các lần lặp với các truy vấn khởi tạo này. Sau khi trích rút đặc trưng, mỗi chiều của đặc trưng được chuẩn hoá vào phạm vi [0,1] sử dụng phương pháp chuẩn Gauss [25].

IV.2. Các hệ thống cơ sở (Baselines)

Hệ thống đề xuất được so sánh với ba phương pháp và được coi như là hệ thống cơ sở và thực nghiệm trên các tập **Db1**, **Db2** và **Db3**. Cả ba phương pháp được thiết lập cùng một môi trường thực nghiệm: các mẫu truy vấn, số lần lặp phản hồi, và cùng một môi trường giả lập người dùng.

- So sánh với học tương tác SVM [30]: Tong và Chang sử dụng SVM để phân lớp các ảnh trong cơ sở dữ liệu ảnh theo sự liên quan và không liên quan.
- So sánh với thuật toán i.Boost [29]: Phân lớp cơ sở dữ liệu ảnh theo truy vấn dựa vào đánh giá của người dùng qua lặp phản hồi liên quan.
- So sánh với kỹ thuật hiệu chỉnh trọng số trong hệ thống MARS [25] của Rui và cộng sự.

IV.3. Độ đo hiệu năng

Hai độ đo Precison với Recall như trong [19] và các ảnh liên quan được tra cứu với số lần lặp (Retrieved relevant - hiệu quả tra cứu) để đánh giá hiệu quả của hệ thống đề xuất. Precision Pr(q) có thể định nghĩa như là tỉ số của số ảnh tra cứu liên quan (Relevant(q), ký hiệu là Rel(q)) với số ảnh tra cứu (N(q)), do đó: $\Pr(q) = \frac{\mathrm{Re}l(q)}{N(q)}$. Recall (Re(q)) được định nghĩa là tỉ số của số ảnh đã tra cứu liên quan với tất cả số ảnh liên quan (C(q)), do đó: $\mathrm{Re}(q) = \frac{\mathrm{Re}l(q)}{\mathrm{C}(q)}$.

Hiệu quả tra cứu được định nghĩa là tỉ số của tổng số ảnh tra cứu liên quan trên tổng số ảnh đã được tra cứu theo lần lặp. Hiệu quả tra cứu được sử dụng cho thấy phần trăm các ảnh tra cứu liên quan cho một lần lặp phản hồi liên quan. Đường cong này cho phép

đánh giá số ảnh liên quan tăng theo các lần lặp. Trung bình Precision với Recall và các ảnh tra cứu liên quan với lần lặp được xem như kết quả cho mọi ảnh truy vấn được sử dụng để so sánh.

IV.4. Các kết quả thực nghiệm

Chúng tôi giả lập ảnh tra cứu được đưa cho người dùng đánh giá. Các ảnh cùng chủ đề với ảnh truy vấn được xem như là liên quan. Bốn phương pháp sử dụng chung các truy vấn trên mỗi tập **Db1**, **Db2**, và **Db3** tương ứng. Với mỗi ảnh truy vấn, ở lần tra cứu khởi tạo các phương pháp đều dùng kết hợp tuyến tính độ đo khoảng cách. Chúng tôi thiết lập 10 lần lặp phản hồi cho mỗi truy vấn.

Các hệ thống CBIR thông thường chọn 20 ảnh tương tự nhất hiển thị cho người dùng đánh giá trong một lần đánh giá. Qua thực nghiệm chúng tôi lựa chọn được các tham số phù hợp cho từng tập dữ liệu như Bảng 2. Như vậy với mức sâu của rìa Pareto chọn hợp lý ta có thể giảm được chi phí tính toán (số mẫu nhỏ hơn), trong khi đó số các ảnh liên quan nhiều nhất.

Bảng 2. Tham số thiết lập rìa Pareto nhiều mức sâu.
Ký hiệu L: mức sâu của rìa Pareto; P: số điểm Pareto.

L			*P*		
Db1	**Db2**	**Db3**	**Db1**	**Db2**	**Db3**
20	30	25	100	500	150

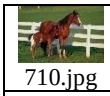
Bảng 3 sử dụng các tham số được thiết lập như Bảng 2, số ảnh tra cứu liên quan với 10 lần lặp là 99 ảnh, giảm được 68.1% không gian số mẫu. Bảng 4, thiết lập tham số tuỳ ý với số điểm Pareto là 300 và độ sâu là 200, số ảnh tra cứu liên quan với 10 lần lặp là 98 ảnh và trung bình giảm 35.8% không gian số mẫu. So sánh Bảng 3 và Bảng 4 cho thấy rõ ràng tính hiệu quả sử dụng tập Pareto.

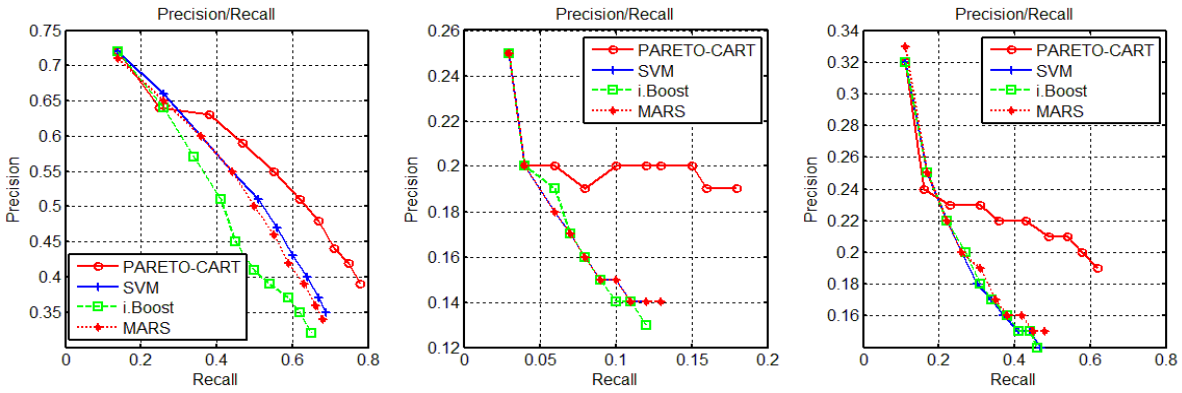
Hình 2(a) là biểu đồ Precision/Recall của cả bốn phương pháp trên tập **Db1**. Trong hai lần lặp đầu tiên, trung bình Precision của phương pháp đề xuất thấp hơn do có rất ít các ảnh được gán nhãn "+" nên nên CART dự báo chưa tốt. Tập dữ liệu này có khoảng trống lớn giữa ngữ nghĩa và đặc trưng mức thấp. Ví dụ: chủ đề xe buýt và hoa hồng về đặc trưng mức thấp rất gần nhau. Ba phương pháp còn lại thực hiện phân lớp ban đầu tốt hơn do tính chất "fitting" của mô hình. Từ lần lặp thứ ba, số ảnh được gán nhãn "+" và "-" tăng lên, CART thực hiện phân lớp hiệu quả rõ rệt trên tập Pareto thu gọn và nhỏ hơn nhiều so với toàn bộ số mẫu. Ngược lại, ba phương pháp còn lại hiệu năng kém hơn từ lần thứ ba vì khi số ảnh được gán nhãn tăng lên, các hệ thống này thường bị "overfitting" và thực hiện phân lớp trên toàn bộ số mẫu rất lớn. Chi tiết số liệu xem trong bảng A.1 ở phụ lục A (Trung bình độ chính xác mô hình đề xuất, SVM, và i.Boost tương ứng là 53.7%, 50.6%, 47.3%, 49.8%).

Bảng 3. Số lượng quần thể trong từng lần phản hồi với truy vấn 710.jpg theo10 lần lặp.
Ký hiệu: P – Số điểm rìa Pareto nhiều mức sâu; NB^+ - số ảnh liên quan tồn tại trong tập.

 710.jpg	*Khởi tạo*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
P	102	451	371	352	442	455	291	385	245	96
NB^+	36	98	87	71	51	33	20	14	5	2
Triệu hồi: 99%, trung bình giảm: 68.1% không gian số lượng mẫu.										

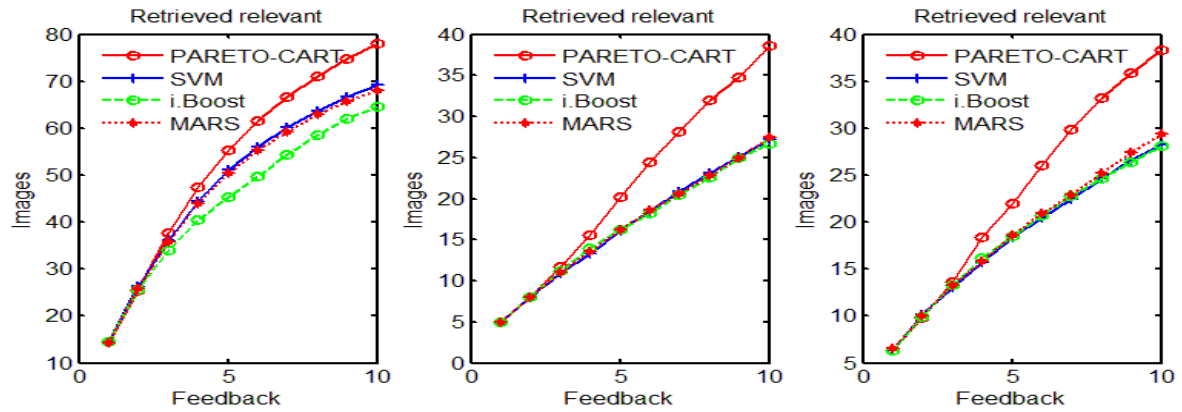
Bảng 4. Số lượng quần thể trong từng vòng phản hồi với truy vấn 710.jpg theo10 lần lặp.
Ký hiệu: P – Số điểm rìa Pareto nhiều mức sâu; NB^+ - số ảnh liên quan tồn tại trong tập.

 710.jpg	*Khởi tạo*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
P	300	833	749	659	742	738	675	691	536	489
NB^+	65	100	88	76	58	43	34	26	10	4
Triệu hồi: 98%; trung bình giảm: 35.8% không gian số lượng mẫu										



(a) (b) (c)

Hình 2. Lược đồ trung bình Precision với Recall cho các mô hình khác nhau (Mô hình đề xuất, SVM, i.Boost, MARS) (a) Db1 (b) Db2 (c) Db3



(a) (b) (c)

Hình 3. Lược đồ hiệu quả tra cứu chp các mô hình khác nhau (Mô hình đề xuất, SVM, i.Boost, MARS) (a) Db1 (b) Db2 (c) Db3

Hình 2(b-c) là biểu đồ Precision/Recall của cả bốn phương pháp trên tập **Db2** và **Db3.** Trên các tập dữ liệu này, hiệu năng của phương pháp đề xuất luôn tốt hơn ba phương pháp còn lại. Hình 3(a-c) cho biết trung bình hiệu quả tra cứu trên ba tập dữ liệu đối với phương pháp đề xuất, SVM, và i.Boost tương ứng sau 10 lần lặp phản hồi liên quan. Trong đó giá trị *Images* là số ảnh tra cứu chính xác và Feedback là lần phản hồi. Kết quả chi tiết trình bày trong bảng A.2, phụ lục A.

Chúng tôi đã phát triển đề xuất thành một ứng dụng cụ thể (Hình A.1 trong phụ lục A), 20 ảnh có thứ hạng đầu tiên được hiển thị trong một lần tra cứu. Trong ứng dụng này, người dùng chọn "-1" và "+1" tương ứng là "không liên quan" và "liên quan". Nếu không chọn, hệ thống không gán nhãn cho đối tượng đó.

V. KẾT LUẬN VÀ HƯỚNG NGHIÊN CỨU

Phương pháp tối ưu Pareto trong tra cứu ảnh theo nội dung ít được sử dụng vì hầu hết các phương pháp khi sử dụng nhiều đặc trưng thường dùng tổng độ đo kết hợp để xếp hạng. Với đề xuất sử dụng tập Pareto để thu hầu hết tập ứng viên với số lượng mẫu nhỏ hơn nhiều so với toàn bộ tập dữ liệu nên cải thiện cho bộ máy phân lớp khi dữ liệu lớn. Mặt khác CART rất phù hợp với số mẫu nhỏ và thường không bị "overfitting" như một số bộ máy phân lớp khác nên sự kết hợp giữa Pareto và CART tạo ra hiệu quả rõ rệt.

Phương pháp đề xuất tránh được tắc nghẽn cục bộ (không tìm được ảnh mong muốn trong khi ảnh đó tồn tại hoặc không tìm thấy ảnh liên quan sau một số lần phản hồi) bằng cách mở rộng truy vấn sử dụng các ảnh

liên quan để thu tập Pareto nhiều mức cho tất cả những ảnh liên quan tránh được những hạn chế có thể gặp phải trong hệ thống MARS.

Để đánh giá hiệu năng của kỹ thuật đề xuất, chúng tôi đã thử nghiệm trên các tập Corel, Oxford Building và Caltech 101. Phương pháp đề xuất được so sánh với các kỹ thuật học tăng cường iBoost, SVM và phân lớp dựa vào hiệu chỉnh trọng số MARS đã chứng tỏ tính hiệu quả của phương pháp đề xuất về: cải thiện hiệu năng bộ máy phân lớp dựa vào giảm số mẫu và tăng chất lượng mẫu bằng hợp các rìa Pareto nhiều mức sâu. Chúng tôi sẽ tiếp tục khai thác thêm một số tích chất của Pareto trong không gian tập độ đo khoảng cách để cải thiện kỹ thuật phân lớp cho học máy trong tra cứu ảnh theo nội dung.

LỜI CẢM ƠN

Chúng tôi xin cám ơn đề tài mã số VAST01.07/15-16 của Viện CNTT, Viện Hàn lâm Khoa học và Công nghệ Việt Nam đã hỗ trợ nghiên cứu này.

PHỤ LỤC A

Bảng A.1. Các thống kê trung bình Precsion với Recall cho các mô hình khác nhau (Mô hình đề xuất, SVM, i.Boost, MARS) (a) Db1 (b) Db2 (c) Db3

Lặp	1	2	3	4	5	6	7	8	9	10	Avg
Pr(PARETO-CART)	0.72	0.64	0.63	0.59	0.55	0.51	0.48	0.44	0.42	0.39	0.537
Re(PARETO-CART)	0.14	0.25	0.38	0.47	0.55	0.62	0.67	0.71	0.75	0.78	0.532
Pr(SVM)	0.72	0.66	0.6	0.55	0.51	0.47	0.43	0.4	0.37	0.35	0.506
Re(SVM)	0.14	0.26	0.36	0.44	0.51	0.56	0.6	0.64	0.67	0.69	0.487
Pr(i.Boost)	0.72	0.64	0.57	0.51	0.45	0.41	0.39	0.37	0.35	0.32	0.473
Re(i.Boost)	0.14	0.26	0.34	0.41	0.45	0.5	0.54	0.59	0.62	0.65	0.45
Pr(MARS)	0.71	0.65	0.6	0.55	0.5	0.46	0.42	0.39	0.36	0.34	0.498
Re(MARS)	0.14	0.26	0.36	0.44	0.5	0.55	0.59	0.63	0.66	0.68	0.481

(a)

Lặp	1	2	3	4	5	6	7	8	9	10	Avg
Pr(PARETO-CART)	0.25	0.2	0.2	0.19	0.2	0.2	0.2	0.2	0.19	0.19	0.202
Re(PARETO-CART)	0.03	0.04	0.06	0.08	0.1	0.12	0.13	0.15	0.16	0.18	0.105
Pr(SVM)	0.25	0.2	0.18	0.17	0.16	0.15	0.15	0.14	0.14	0.14	0.168
Re(SVM)	0.03	0.04	0.06	0.07	0.08	0.09	0.1	0.11	0.12	0.12	0.082
Pr(i.Boost)	0.25	0.2	0.19	0.17	0.16	0.15	0.15	0.14	0.14	0.13	0.168
Re(i.Boost)	0.03	0.04	0.06	0.07	0.08	0.09	0.09	0.1	0.11	0.12	0.079
Pr(MARS)	0.25	0.2	0.18	0.17	0.16	0.15	0.15	0.14	0.14	0.14	0.168
Re(MARS)	0.03	0.04	0.06	0.07	0.08	0.09	0.1	0.11	0.12	0.13	0.083

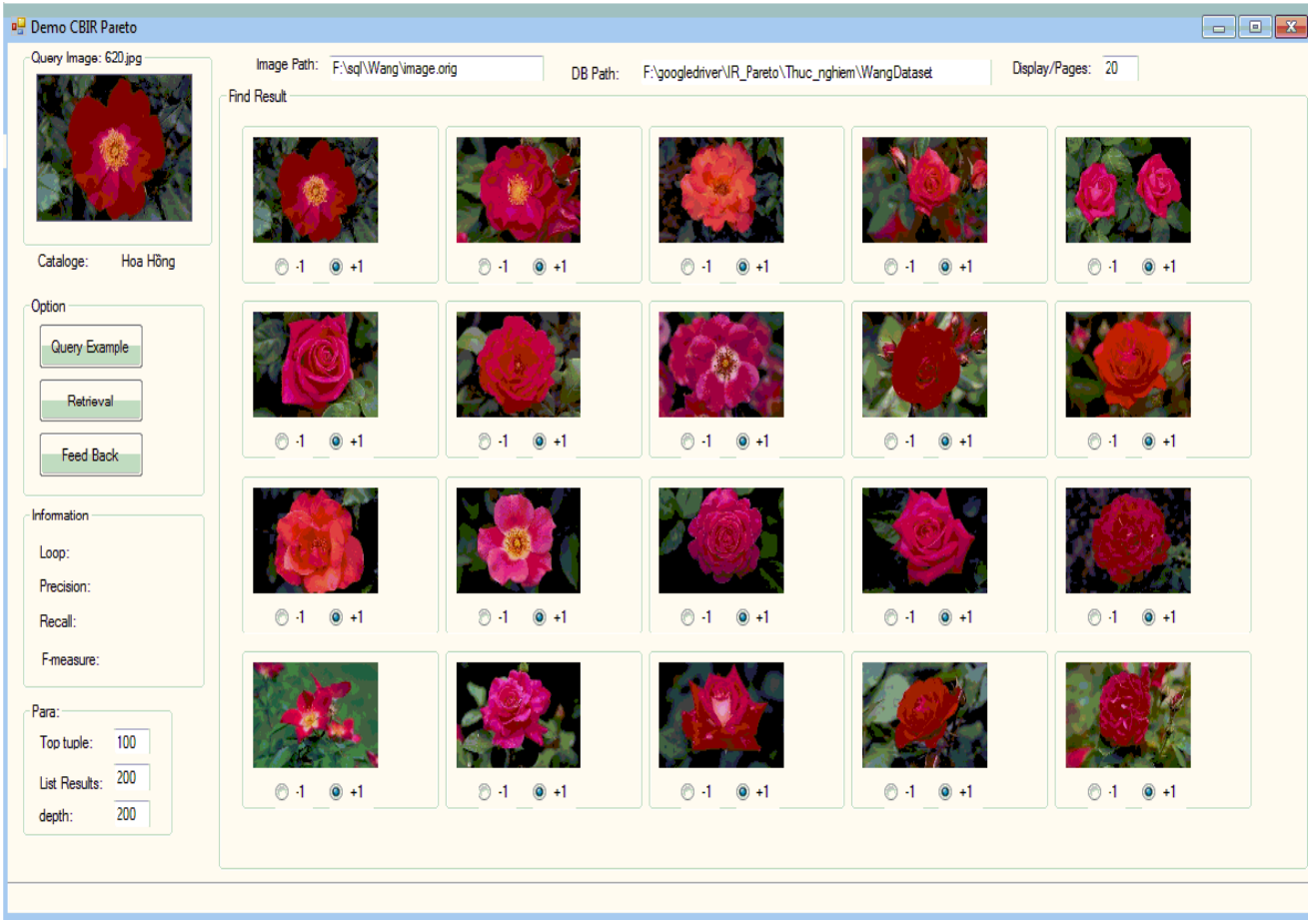
(b)

Lặp	1	2	3	4	5	6	7	8	9	10	Avg
Pr(PARETO-CART)	0.32	0.24	0.23	0.23	0.22	0.22	0.21	0.21	0.2	0.19	0.227
Re(PARETO-CART)	0.11	0.16	0.23	0.31	0.36	0.43	0.49	0.54	0.58	0.62	0.383
Pr(SVM)	0.32	0.25	0.22	0.2	0.18	0.17	0.16	0.15	0.15	0.14	0.194
Re(SVM)	0.11	0.17	0.22	0.26	0.3	0.34	0.37	0.41	0.44	0.47	0.309
Pr(i.Boost)	0.32	0.25	0.22	0.2	0.18	0.17	0.16	0.15	0.15	0.14	0.194
Re(i.Boost)	0.11	0.17	0.22	0.27	0.31	0.34	0.38	0.41	0.44	0.46	0.311
Pr(MARS)	0.33	0.25	0.22	0.2	0.19	0.17	0.16	0.16	0.15	0.15	0.198
Re(MARS)	0.11	0.17	0.22	0.26	0.31	0.35	0.38	0.42	0.45	0.48	0.315

(c)

Bảng A.2. Các thống kê trung bình hiệu quả tra cứu cho các mô hình khác nhau (Mô hình đề xuất, SVM, i.Boost, MARS) (a) Db1 (b) Db2 (c) Db3

	Db1				Db2				Db3			
Lặp	PARETO -CART	SVM	i.Boost	MARS	PARETO -CART	SVM	i.Boost	MARS	PARETO -CART	SVM	i.Boost	MARS
1	14.41	14.41	14.41	14.28	4.93	4.91	4.91	4.93	6.37	6.37	6.37	6.61
2	25.45	26.25	25.54	25.98	8.02	8	8	8.07	9.76	10.13	9.98	10.07
3	37.62	36.14	34.06	35.77	11.7	10.96	11.13	11.06	13.61	13.04	13.26	13.26
4	47.34	44.32	40.58	43.95	15.54	13.3	13.94	13.57	18.39	15.74	16.15	15.87
5	55.2	51.05	45.24	50.41	20.2	16.09	16.28	16.17	22.02	18.37	18.5	18.67
6	61.68	56.1	49.78	55.26	24.48	18.5	18.24	18.57	26.04	20.48	20.67	20.93
7	66.65	60.19	54.29	59.31	28.15	20.8	20.41	20.59	29.83	22.43	22.83	22.96
8	71.06	63.68	58.51	62.88	31.93	23.15	22.52	22.89	33.22	24.7	24.67	25.28
9	74.75	66.71	62.11	65.64	34.81	25.15	24.91	24.94	35.91	26.57	26.41	27.39
10	78.11	69.33	64.54	68.04	38.57	27.26	26.69	27.48	38.33	28.26	28.07	29.35



Hình A.1. Hệ thống tra cứu ảnh dựa vào nội dung

TÀI LIỆU THAM KHẢO

[1] AREVALILLO-HERRÁEZ, MIGUEL, FRANCESC J. FERRI, and SALVADOR MORENO-PICOT, *Improving distance based image retrieval using non-dominated sorting genetic algorithm*, Pattern Recognition Letters 53 (2015): 109-117.

[2] BAI, CONG, KIDIYO KPALMA, and JOSEPH RONSIN, *Color textured image retrieval by combining texture and color features*, Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European. IEEE, 2012.

[3] BIGGS, DAVID, BARRY DE VILLE, and ED SUEN, *A method of choosing multiway partitions for classification and decision trees*, Journal of Applied Statistics18.1 (1991): 49-62.

[4] BREIMAN, LEO, et al, *Classification and regression trees*, CRC press, 1984.

[5] ŞAYKOL, EDIZ, UĞUR GÜDÜKBAY, and ÖZGÜR ULUSOY, *A histogram-based approach for object-based query-by-shape-and-color in image and video databases*, Image and Vision Computing 23.13 (2005): 1170-1180.

[6] DATTA, RITENDRA, et al, *Image retrieval: Ideas, influences, and trends of the new age*, ACM Computing Surveys (CSUR) 40.2 (2008): 5.

[7] DENG, YINING, et al, *An efficient color representation for image retrieval*, Image Processing, IEEE Transactions on 10.1 (2001): 140-147.

[8] DOS SANTOS, J. A., et al, *A relevance feedback method based on genetic programming for classification of remote sensing images*, Information Sciences 181.13 (2011): 2671-2684.

[9] DUBEY, RAJSHREE S., RAJNISH CHOUBEY, and JOY BHATTACHARJEE, *Multi feature content based image retrieval*, International Journal on Computer Science and Engineering 2.6 (2010): 2145-2149.

[10] FEI-FEI, LI, ROB FERGUS, and PIETRO PERONA, *Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories*, Computer Vision and Image Understanding106.1 (2007): 59-70.

[11] FERREIRA, CRISTIANO D., et al, *Relevance feedback based on genetic programming for image retrieval*, Pattern Recognition Letters 32.1 (2011): 27-37.

[12] HSIAO, KO-JEN, JEFF CALDER, and ALFRED O. HERO, *Pareto-Depth for Multiple-Query Image Retrieval*, Image Processing, IEEE Transactions on 24.2 (2015): 583-594.

[13] HUANG, JING, et al. *Image indexing using color correlograms*. Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on. IEEE, 1997.

[14] JIANG, WEI, GUIHUA ER, and QIONGHAI DAI, *Boost SVM active learning for content-based image retrieval*, Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on. Vol. 2. IEEE, 2003.

[15] KNOWLES, JOSHUA D., and David W. Corne, *Approximating the nondominated front using the Pareto archived evolution strategy*, Evolutionary computation8.2 (2000): 149-172.

[16] KORYTKOWSKI, MARCIN, LESZEK RUTKOWSKI, and RAFAŁ SCHERER, *Fast image classification by boosting fuzzy classifiers*, Information Sciences 327 (2016): 175-182.

[17] LI, JIA, and JAMES Z. WANG, *Automatic linguistic indexing of pictures by a statistical modeling approach*, Pattern Analysis and Machine Intelligence, IEEE Transactions on 25.9 (2003): 1075-1088.

[18] MACARTHUR, SEAN D., CARLA E. BRODLEY, and CHI-REN SHYU, *Relevance feedback decision trees in content-based image retrieval*, Content-based Access of Image and Video Libraries, 2000, Proceedings, IEEE Workshop on, IEEE, 2000.

[19] MÜLLER, HENNING, et al, *Performance evaluation in content-based image retrieval: overview and proposals*, Pattern Recognition Letters 22.5 (2001): 593-601.

[20] OLIVA, AUDE, and ANTONIO TORRALBA, *Modeling the shape of the scene: A holistic representation of the spatial envelope*, International journal of computer vision42.3 (2001): 145-175.

[21] PHILBIN, JAMES, et al, *Object retrieval with large vocabularies and fast spatial matching*, Computer Vision and Pattern Recognition, 2007, CVPR'07, IEEE Conference on, IEEE, 2007.

[22] RAHMAN, M. M., BIPIN C. DESAI, and PRABIR BHATTACHARYA, *Multi–modal interactive approach to ImageCLEF 2007 photographic and medical retrieval tasks by CINDI*, Working Notes of CLEF 7 (2007).

[23] RUI, YONG, et al, *Automatic matching tool selection using relevance feedback in MARS*, Proc. of 2nd Int. Conf. on Visual Information Systems, 1997.

[24] RUI, YONG, THOMAS S. HUANG, and SHARAD MEHROTRA, *Content-based image retrieval with relevance feedback in MARS*, Image Processing, 1997 Proceedings., International Conference on, Vol. 2, IEEE, 1997.

[25] RUI, YONG, et al, *Relevance feedback: a power tool for interactive content-based image retrieval*, Circuits and Systems for Video Technology, IEEE Transactions on 8.5 (1998): 644-655.

[26] RUI, YONG, THOMAS S. HUANG, and SHIH-FU CHANG, *Image retrieval: Current techniques, promising directions, and open issues*, Journal of visual communication and image representation 10.1 (1999): 39-62.

[27] SALTON, GERARD, and MICHAEL J. MCGILL, *Introduction to modern information retrieval*, (1986).

[28] Swain, Michael J., and Dana H. Ballard, *Color indexing*, International journal of computer vision 7.1 (1991): 11-32.

[29] TIEU, KINH, and PAUL VIOLA, *Boosting image retrieval*, International Journal of Computer Vision 56.1-2 (2004): 17-36.

[30] TONG, SIMON, and EDWARD CHANG, *Support vector machine active learning for image retrieval*, Proceedings of the ninth ACM international conference on Multimedia, ACM, 2001.

[31] TORLONE, RICCARDO, PAOLO CIACCIA, and U. ROMATRE, *Which are my preferred items*, Workshop on Recommendation and Personalization in E-Commerce, 2002.

[32] TORRES, RICARDO DA S., et al, *A genetic programming framework for content-based image retrieval*, Pattern Recognition 42.2 (2009): 283-292.

[33] Yu, Hui, et al, *Color texture moments for content-based image retrieval*, Image Processing. 2002. Proceedings. 2002 International Conference on. Vol. 3. IEEE, 2002.

[34] YU, JIE, et al, *Integrating relevance feedback in boosting for content-based image retrieval*, Acoustics, Speech and Signal Processing, 2007, ICASSP 2007, IEEE International Conference on. Vol. 1, IEEE, 2007.

[35] ZHANG, DENGSHENG, et al, *Content-based image retrieval using Gabor texture features*, IEEE Pacific-Rim Conference on Multimedia, University of Sydney, Australia. 2000.

[36] ZHANG, QIANNI, and EBROUL IZQUIERDO, *Optimizing metrics combining low-level visual descriptors for image annotation and retrieval*, Acoustics, Speech and Signal Processing, 2006, ICASSP 2006 Proceedings, 2006 IEEE International Conference on, Vol, 2. IEEE, 2006.

Nhận bài ngày: 18/02/2016

SƠ LƯỢC VỀ TÁC GIẢ

VŨ VĂN HIỆU

Sinh năm 1976 tại Kiến Thuỵ, Hải Phòng.
Đang là nghiên cứu sinh năm thứ 4 tại Viện CNTT, Viện Hàn lâm KH&CN Việt Nam, chuyên ngành cơ sở toán cho tin học.
Hiện công tác tại Khoa CNTT, Trường ĐH Hải Phòng.
Email: hieuvv@dhhp.edu.vn

NGUYỄN TRƯỜNG THẮNG

Tốt nghiệp năm 1997 tại Đại học tổng hợp New South Wales , Australia, Tiến sĩ Tin học năm 2005 tại Viện Khoa học và Công nghệ tiên tiến Nhật Bản (JAIST). Hiện công tác tại Viện CNTT, Viện Hàn lâm KH&CN Việt Nam.
Email: ntthang@ioit.ac.vn

NGUYỄN HỮU QUỲNH

Tốt nghiệp ĐH, Cao học và Tiến sĩ tại ĐH Quốc gia Hà Nội vào các năm 1998, 2004 và 2010.
Hiện công tác tại Khoa CNTT, Trường ĐH Điện Lực, Hà Nội.
Email: quynhnh@epu.edu.vn

NGÔ QUỐC TẠO

Nhận bằng Tiến sĩ đảm bảo toán học cho các hệ thống tính toán năm 1997, được phong Phó Giáo sư năm 2002.
Hiện công tác tại Viện CNTT, Viện Hàn lâm KH&CN Việt Nam.
Email: nqtao@ioit