

NHẬN DẠNG CÁ NHÂN TRONG HÌNH ẢNH CỦA MỤC TIN DỰA VÀO PHÂN TÍCH VĂN BẢN

Choochart Haruechaliyasak và Chaianun Damrongrat
Trung tâm Quốc gia Công nghệ Điện tử và Máy tính Thái Lan

1. Mở đầu

Để tăng thêm hình ảnh và cung cấp cho người đọc tài liệu minh họa, phần lớn website tin tức thường đưa vào bài mục các nội dung đa phương tiện gắn với hình ảnh và video clip. Một trong những mục tiêu và thách thức trong việc tìm hình ảnh là phải nhận dạng đúng và tìm được hình ảnh của cá nhân cụ thể từ bộ sưu tập hình ảnh lớn.

Những công trình nghiên cứu trước đây về nhận dạng con người trong các hình ảnh thường dựa vào hai cách tiếp cận khác nhau: *nhận mặt* và *tìm chỉ mục*. Nhận mặt dựa vào một số kỹ thuật xử lý hình ảnh để xây dựng và nhận biết các đặc điểm từ hình ảnh bộ mặt [16]. Với một loạt cải tiến được thực hiện trong nhiều năm qua, nhận mặt đã được áp dụng thành công trong nhiều ứng dụng như trắc lượng sinh học và điện tử học tiêu dùng (ví dụ, máy ảnh số). Tuy nhiên, áp dụng nhận dạng mặt trong lĩnh vực tìm hình ảnh người có một số hạn chế vì những lý do sau đây: (1) tách ra và biết được các đặc điểm về hình từ các hình ảnh đòi hỏi rất nhiều thời gian xử lý; (2) để nhận ra từng cá nhân cụ thể đòi hỏi phải có một kích cỡ mẫu hình mặt thích hợp, do đó, cách tiếp cận này không được mở rộng nhiều khi xử lý một số lượng lớn cá nhân và (3) kỹ thuật nhận mặt đơn giản không đủ mạnh để nhận dạng con người từ những góc nhìn khác nhau và trong những hình ảnh độ phân giải thấp.

Để giải quyết quy mô lớn của Web, nhiều máy tìm hình ảnh đã chấp nhận cách tiếp cận tìm theo chỉ mục đang được áp dụng thành công cho việc tìm trang web [8]. Phép tìm theo chỉ mục điển hình lại dựa vào mô thức túi từ để xác định cấu trúc chỉ mục đảo, nghĩa là, mỗi thuật ngữ sẽ chỉ tới một danh sách tư liệu có chứa thuật ngữ đó. Theo mô thức túi từ, một tư liệu có chứa câu hỏi đầu vào ở bất kỳ chỗ nào trong nội dung đều được coi là trúng đích và sẽ được phản hồi cho người sử dụng.

Đối với phép tìm hình ảnh, một chỉ mục có thể được xây dựng từ nhiều mẫu thông tin văn bản như tên file ảnh, các lời thuyết minh gắn liền với hình ảnh và các đoạn văn bản liền kề với hình ảnh [6].

Nói chung, sử dụng phép tìm theo chỉ mục cho các hình ảnh trên Web là hữu hiệu. Tuy nhiên, tìm hình ảnh của các cá nhân cụ thể trong lĩnh vực tin tức đôi khi có thể đưa ra kết quả không mong muốn. Với mô thức túi từ đơn giản, một câu hỏi tìm một người có thể tương ứng với một vài mục tin (nghĩa là người đó được nói tới trong nhiều tin) mà hình ảnh của các mục đó có cả ảnh của những người khác. Do đó, những hình ảnh không tương thích đôi khi cũng được đưa ra cho người sử dụng cùng với hình ảnh đúng của cá nhân mục tiêu. Chúng tôi xin minh họa vấn đề này bằng cách sử dụng một ví dụ từ *Phép tìm hình ảnh tin tức theo Google*¹ dựa trên phương pháp tiếp cận tìm theo chỉ mục. Hình 1 cho

¹ Google News, <http://news.google.com>

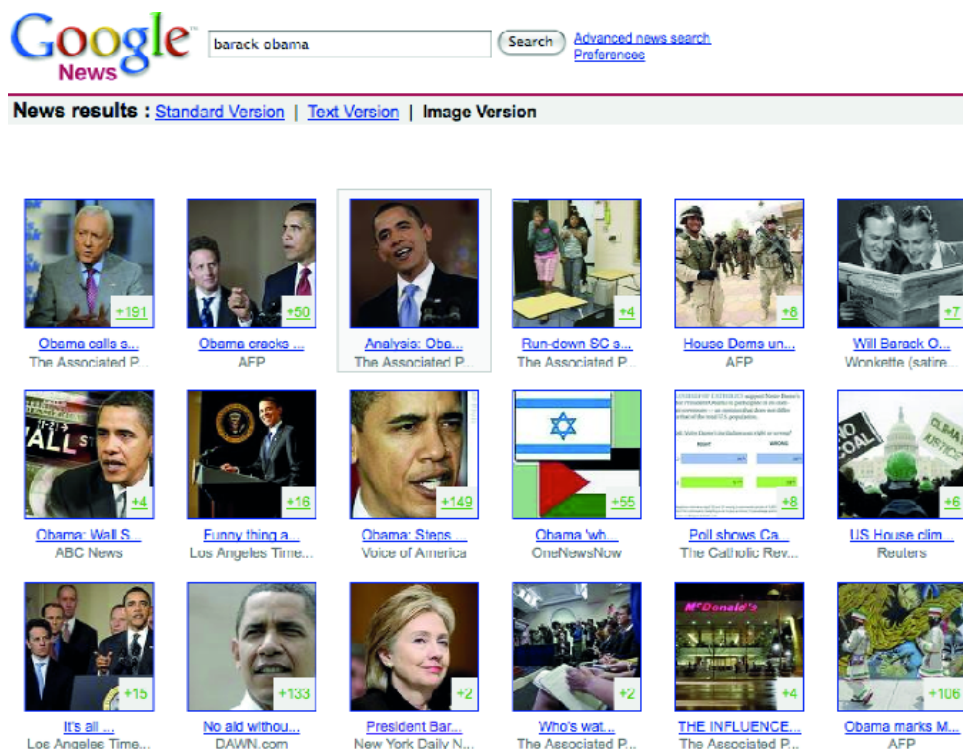
Nhìn ra thế giới

thấy kết quả tìm bằng cách sử dụng “Barack Obama” làm câu hỏi. Từ các kết quả này, hơn một nửa số hình ảnh phản hồi không có hình ảnh của “Barack Obama”.

Theo quan sát ban đầu của chúng tôi, một mục tin thường chứa hình ảnh của cá nhân mục tiêu nếu tên của người đó xuất hiện như chủ ngữ chính trong các câu khác nhau của mục tin, ví dụ trong đầu đề và chú thích hình ảnh. Hình 2 là ví dụ về một mục tin lấy từ website tin tức của CNN² có chứa hình ảnh của một cá nhân với đầu đề và chú thích hình ảnh. Từ hình này, hai thực thể tên người “Hillary Clinton” và “Barack Obama” xuất hiện trong cả đầu đề lẫn chú thích hình ảnh. Khi sử dụng phép tìm theo chỉ mục đơn giản, một câu hỏi tìm “Barack Obama” có thể sẽ đưa

nhằm hình ảnh này của “Hillary Clinton”.

Mục đích chính của bài này là cải tiến hiệu suất của phép tìm theo chỉ mục cơ bản để tìm được các cá nhân trong các hình ảnh tin tức. Cách tiếp cận của chúng tôi là áp dụng phân tích cú pháp về nội dung văn bản (ví dụ, các đầu đề và chú thích hình ảnh) để nhận dạng khẳng định và phủ định cá nhân trong các hình ảnh tin tức. Phép phân tích cú pháp trong phương pháp của chúng tôi bao gồm việc định nhãn thực thể có tên và phân tích cú pháp sơ bộ [1, 3]. Từ Hình 2, bằng cách thực hiện phân tích đầu đề tin, “Clinton” sẽ được phân tích và định nhãn như là một cụm danh từ (NP), sau đó “chấp nhận” như một cụm động từ (VP) và “Obama” như là một cụm danh từ (NP)



Hình 1. Kết quả tìm “Barack Obama” theo phép tìm hình ảnh tin tức trên Google

² CNN.com International , <http://edition.cnn.com>

Đầu đề: Clinton hậu thuẫn cho Obama, kêu gọi sự thống nhất trong Đảng



Chú thích : Thượng nghị sỹ Hillary Clinton cảm ơn những người ủng hộ bà về bài diễn văn hậu thuẫn cho Thượng nghị sỹ Barack Obama

Hình 2. Ví dụ về hình ảnh một cá nhân với đầu đề và chú thích hình ảnh

Từ câu này, có thể rút ra hai quy tắc: (1) [NP*] [VP] [NP] để nhận dạng khẳng định người trong các hình ảnh : và (2) [NP] [VP] [NP*] để nhận dạng phủ định, trong đó dấu * chỉ vị trí xuất hiện tên người. Quy tắc nhận dạng khẳng định có nghĩa là: nếu tên người đó xuất hiện như một cụm danh từ và là chủ ngữ của một câu thì hình ảnh này chắc chắn có chứa người muốn tìm. Ngược lại, quy tắc nhận dạng phủ định có nghĩa là: nếu tên người này xuất hiện như một cụm danh từ và là tân ngữ của một câu thì hình ảnh này không chắc chắn chứa người muốn tìm.

Sử dụng tập hợp mục tin có định nhãn, chúng ta xây dựng được một bộ quy tắc cú pháp dựa vào những ý tưởng nói trên. Sau đó, bộ quy tắc cú pháp được sử dụng để lựa chọn và sàng lọc các kết quả hình ảnh do phép tìm theo chỉ mục đem lại. Từ ví dụ trên, câu hỏi “Obama” sẽ tìm ra hình ảnh

từ Hình 2, tuy nhiên, hình ảnh này sẽ được lọc ra bằng cách sử dụng quy tắc nhận dạng phủ định [NP] [VP] [NP*] vì “Obama” xuất hiện như là NP* trong câu. Kết quả thí nghiệm về tập hợp hình ảnh tin tức cho thấy, cách tiếp cận phân tích cú pháp có thể cải thiện một cách đáng kể hiệu suất tìm kiếm đối với phép tìm theo chỉ mục cơ bản.

2. Các công trình nghiên cứu liên quan

Các công trình nghiên cứu trước đây về tìm kiếm hình ảnh chủ yếu dựa vào hai kỹ thuật khác nhau : tìm hình ảnh dựa vào văn bản và dựa vào nội dung (CBIR) [4]. Cách tiếp cận dựa vào văn bản căn cứ vào thông tin văn bản kết hợp với hình ảnh để khớp một câu hỏi đầu vào, gồm một hay nhiều từ, vào các hình ảnh. Do khả năng tùy biến kích cỡ, cách tìm hình ảnh dựa vào văn bản đã được áp dụng phổ biến ở nhiều máy tìm

như Google³ và Yahoo⁴. Mặt khác, CBIR áp dụng kỹ thuật xử lý hình ảnh để tách ra và biết được các đặc điểm hình từ các hình ảnh. Trong các hệ thống CBIR, một câu hỏi về hình, mô tả các đặc điểm hình như màu sắc và kết cấu, được dùng để tìm ra hình ảnh tương ứng tốt nhất với các đặc điểm đầu vào [7, 12].

Với số lượng lớn hình ảnh trên Web có chứa ảnh của các cá nhân, đề tài nghiên cứu về tìm kiếm hình ảnh phục vụ cho việc tìm hình ảnh người đã ngày càng được chú ý. Trong cách tìm hình ảnh người, nếu đặt câu hỏi tìm một tên người cụ thể, mục đích là phải tìm được chính xác các hình ảnh của người đó (ưu tiên là ảnh cá nhân) từ bộ sưu tập hình ảnh lớn (ví dụ, từ các website tin tức). Kỹ thuật tìm hình ảnh người tương tự như các kỹ thuật được sử dụng để tìm kiếm hình ảnh nói chung. Cách tiếp cận dựa vào nội dung, vào kỹ thuật nhận mặt để xây dựng và biết các đặc điểm hình từ hình ảnh bộ mặt. Rất nhiều công trình nghiên cứu trước đây về tìm kiếm hình ảnh người đều dựa vào kỹ thuật nhận mặt [2,9,10,11]. Chúng đã áp dụng các biến thể của kỹ thuật nhận mặt để liên kết các hình ảnh mặt với tên người được tìm thấy trong các mục tin. Như đã nói ở trên, áp dụng kỹ thuật nhận mặt không có nhiều khả năng nâng cao tùy biến vì có nhiều người trong các tin tức. Do đó, cần có một kích cỡ mẫu lớn của hình ảnh mặt cho mỗi cá nhân. Ngoài ra, việc tách ra và nhận biết các đặc điểm hình từ hình ảnh mặt đòi hỏi nhiều thời gian xử lý.

Để tăng khả năng nâng cao tùy biến, nhiều công trình nghiên cứu đã áp dụng

cách tiếp cận dựa vào văn bản để tìm người trong các mục tin tức. Sihari [13] đã đưa ra hệ thống gọi là *Piction*, dùng các chú thích ảnh để nhận dạng từng cá nhân từ một hình ảnh có chứa nhiều người. Hệ thống này kết hợp thuật toán định vị mặt trong hình ảnh với môđun NLP để phân tích vị trí không gian của cá nhân từ chú thích hình ảnh. Yagnik và Islam [14] đã đưa ra phương pháp dựa vào nhận biết sự đồng xuất hiện của văn bản và hình ảnh trên Web. Cách tiếp cận này gọi là nhận biết nhất quán, thường dùng để chú giải hình ảnh người. Yang và Hauptmann [15] đưa ra phương pháp nhận biết thống kê để nhận dạng mỗi cá nhân xuất hiện trong băng video truyền hình dựa vào văn cảnh tìm thấy trong các bản sao chép video. Edwards và những người khác [5] áp dụng phương pháp phân nhóm văn bản để tập hợp hình ảnh người vào một số nhóm chủ đề. Một số quan sát về cấu trúc cú pháp tên người trong các chú thích hình ảnh đã được bàn đến và sử dụng trong các thí nghiệm.

Phần đóng góp chính của chúng tôi là đưa ra cách tiếp cận có hiệu lực và có khả năng nâng cao tùy biến tự động xây dựng bộ quy tắc cú pháp để nhận dạng cá nhân trong các hình ảnh tin tức. Các quy tắc này được sử dụng để chọn lựa và sàng lọc các kết quả hình ảnh mà phương pháp tìm theo chỉ mục đưa lại.

3. Cách tiếp cận dựa vào phân tích văn bản

Cách tiếp cận của chúng tôi trong tìm kiếm hình ảnh người sẽ kết hợp việc tìm theo chỉ mục với việc phân tích cú pháp các thực thể có tên người (NE) trong nội

³ Google Image Search, <http://images.google.com>

⁴ Yahoo! Search- Image Search <http://images.google.com>

Nhìn ra thế giới

dung văn bản. Nếu cho một câu hỏi tìm tên người, cách tìm kiếm theo chỉ mục thường cung cấp những hình ảnh không đúng ngoài những hình ảnh đúng của cá nhân mục tiêu. Ý tưởng chính là xây dựng bộ quy tắc cú pháp từ thông tin văn bản để lọc ra các hình ảnh không đúng.

Hình 3 minh họa quy trình tổng thể xây dựng bộ quy tắc cú pháp từ một tập hợp mục tin. Nếu mục tin chứa một hình ảnh người, chúng tôi áp dụng việc nhận biết thực thể có tên để định nhãn tất cả các thực thể có tên của con người xã hội trong thông tin văn bản (nghĩa là: đầu đề, phụ đề và chú thích ảnh). Tất cả các câu có chứa các thực thể có tên người được tập hợp và định

nhãn như là các mẫu *khẳng định* hoặc *phủ định*. Mẫu khẳng định là các câu có chứa tên người được hiển thị trong hình ảnh. Mẫu phủ định là các câu có chứa tên người khác mà không xuất hiện trong hình ảnh. Hình 4 là ví dụ về các mẫu khẳng định và phủ định. Đối với mẫu khẳng định, câu hỏi tìm “Tiger Woods” sẽ ánh xạ đúng với hình ảnh đã cho. Đối với mẫu phủ định, câu hỏi tìm “Rafael Nadal” có thể sẽ ánh xạ sai tới hình ảnh của một cầu thủ quần vợt khác “Lleyton Hewitt”. Có thể thấy rằng, đối với mẫu khẳng định, tên người mục tiêu xuất hiện như là chủ ngữ của câu. Mặt khác, đối với mẫu phủ định, tên người mục tiêu xuất hiện như là tân ngữ của câu.



Hình 3. Cách tiếp cận để xây dựng bộ quy tắc sàng lọc

Nhìn ra thế giới

Mẫu khẳng định cho câu hỏi tìm “Tiger Woods”



Đầu đề: Tiger Woods cần bắt đầu ra quân sớm hơn

Mẫu khẳng định cho câu hỏi tìm “Rafael Nadal”



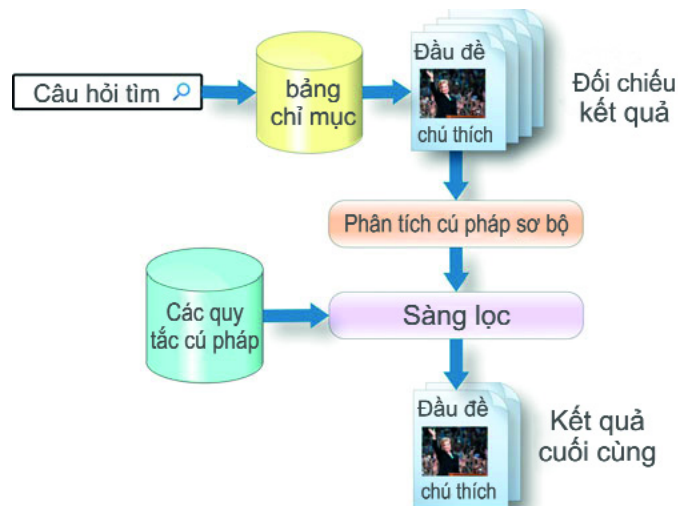
Tiểu đề: Nhà 4 lần vô địch Lieyton Hewitt sẽ cùng Andy Muray và Rafael Nadal tham gia giải vô địch AEGON tháng sau ở Câu lạc bộ Queen

Hình 4. Ví dụ về các mẫu khẳng định và phủ định

Một khi thu thập được các mẫu khẳng định và phủ định từ tập hợp tin tức, thì kỹ thuật phân tích cú pháp sơ bộ được áp dụng để phân đoạn và định nhãn các thực thể có tên người và tất cả các thuật ngữ khác trong câu.

Chẳng hạn, câu “Obama cần 1,5 tỷ đôla cho các chế phẩm chống dịch” sẽ được phân tích cú pháp và định nhãn như sau [NP Obama] {VP cần} [NP1,5 tỷ đôla], [PP dành cho] [NP các chế phẩm chống

dịch]; trong đó NP chỉ cụm danh từ, VP chỉ cụm động từ và PP chỉ cụm giới từ. Vì câu trên có chứa hình ảnh của “Barack Obama”, chúng ta có thể có quy tắc nhận dạng khẳng định như [#] [NP*] [VP] [NP]: *khẳng định*, trong đó # chỉ giới hạn câu, * chỉ vị trí xuất hiện tên người và: *khẳng định* là nhãn loại. Trong bài này, chúng tôi dùng *NLP mở*⁵, một công cụ phân tích văn bản, để thực hiện tất cả các nhiệm vụ xử lý văn bản liên quan.



Hình 5. Quy trình sàng lọc kết quả tìm kiếm hình ảnh

⁵ The OpenNLP Homepage, <http://openlp.sourceforge.net>

Nhìn ra thế giới

Sau khi có bộ quy tắc nhận dạng khẳng định và phủ định, chúng tôi xây dựng bộ quy tắc sàng lọc để lọc ra các kết quả hình ảnh sai được hồi đáp từ phép tìm theo chỉ mục. Cùng một số quy tắc cú pháp có thể

$số\ điểm\ (quy\ tắc\ sàng\ lọc) =$

có những nhân loại khẳng định và phủ định. Do đó, chúng tôi tính tỷ số điểm của từng quy tắc sàng lọc bằng tỷ số giữa số nhân phủ định và số nhân khẳng định.

$số\ lượng\ (quy\ tắc\ phủ\ định)$

$số\ lượng\ (quy\ tắc\ khẳng\ định)$

Trường hợp số nhân khẳng định bằng 0, nghĩa là quy tắc chỉ xuất hiện ở loại phủ định. Những quy tắc này thường được coi là đứng đầu trong số các quy tắc khác. Tỷ số điểm bằng 1, nghĩa là quy tắc xuất hiện ngang nhau trong cả loại khẳng định và phủ định. Chúng tôi xem xét các quy tắc có giá trị tỷ số điểm lớn hơn 1 (nghĩa là, $số\ lượng\ (quy\ tắc\ phủ\ định) > số\ lượng\ (quy\ tắc\ khẳng\ định)$). Hình 5 minh họa quá trình sàng lọc các kết quả tìm kiếm hình ảnh từ cách tiếp cận tìm theo chỉ mục. Nếu đặt câu hỏi tìm theo tên người, hệ thống sẽ tìm thấy tất cả các hình ảnh có chứa tên của cá nhân mục tiêu từ bảng chỉ mục. Kế tiếp, các bản mô tả thành văn các hình ảnh được xử lý với phép phân tích cú pháp sơ bộ, sau đó, các quy tắc sàng lọc được áp dụng để loại ra các hình ảnh chắc chắn không chứa cá nhân mục tiêu.

Các thử nghiệm và trao đổi

Để đánh giá hiệu quả của phương pháp mà chúng tôi đưa ra, chúng tôi thực hiện một số thí nghiệm có sử dụng bộ sưu tập hình ảnh tin tức nhận được từ cách tìm hình ảnh tin tức trên Google. Khi có danh sách các cá nhân thuộc các loại khác nhau (như trình bày trong Bảng 1), chúng tôi tải xuống các nội dung văn bản cùng với hình ảnh và chú thích kèm theo. Số lượng hình ảnh hiện tại trong tập hợp xấp xỉ 500.

Bảng 1. Danh sách tên người trong tập hợp thử nghiệm

Tên người	Loại
Barack Obama	Chính trị
Hillary Clinton	Chính trị
Tiger Woods	Thể thao
David Beckham	Thể thao
Rafael Nadal	Thể thao
Britney Spears	Giải trí
Angelina Jolie	Giải trí

Bảng 2. Các quy tắc sàng lọc hàng đầu

Các quy tắc sàng lọc	Tỷ số điểm
[NP] [VP] [NP*][PP][NP]	7,0
[#] [NP*] [O] [#]	5,0
[NP] [VP] [NP*] [#]	2,67
[#] [NP*] [VP] [ADVP]	2,0
[#] [NP*] [NP] [O]	2,0
[O] [VP] [NP*] [O] [#]	2,0
[O] [VP] [NP*][NP] [O]	2,0
[NP] [VP] [NP*] [VP] [NP]	2,0
[NP] [O] [NP*] [VP] [ADJP]	2,0
[NP] [PP] [NP*][O] [NP]	1,67

Sử dụng tập hợp này, chúng tôi đã xây dựng bộ quy tắc sàng lọc dựa vào quy trình đã nói ở trên. Chúng tôi đặt số lượng tối đa gram cho phép (nghĩa là số lượng mã

Nhìn ra thế giới

thông báo trong một quy tắc) bằng 5 xuyên suốt các thử nghiệm. Bảng 2 cho thấy một số quy tắc hàng đầu dựa vào tính điểm được cho trong phương trình (1), trong đó, [#] chỉ giới hạn câu, * chỉ vị trí xuất hiện tên người, NP chỉ cụm danh từ, VP chỉ cụm động từ, PP chỉ cụm giới từ, ADVP chỉ cụm trạng từ, ADJP chỉ cụm tính từ, và O chỉ một ký tự đặc biệt. Có thể tìm thấy danh sách đầy đủ ký hiệu định nhãn từ Dự án Penn Treebank⁶.

Chúng tôi đánh giá hiệu suất của cách tiếp cận phân tích cú pháp đã đưa ra bằng cách so sánh với cách tìm theo chỉ mục.

Đối với cách tìm theo chỉ mục, chúng tôi xây dựng một bảng chỉ mục các thuật ngữ được tách ra và mã hóa thông báo từ thông tin văn bản của mỗi mục tin. Đối với cách tiếp cận phân tích cú pháp, chúng tôi áp dụng bộ quy tắc cú pháp để lọc ra các kết quả hình ảnh mà cách tìm theo chỉ mục đem lại. Một hình ảnh mà trong đó văn bản kết hợp ứng với các quy tắc sàng lọc sẽ được loại khỏi các kết quả. Chúng tôi sử dụng phép trắc lượng tìm chuẩn về độ chính xác, mức độ hồi đáp (mức độ tìm lại được) và số đo F1 để đánh giá. Các kết quả thử nghiệm được tổng kết trong Bảng 3.

Bảng 3. Kết quả đánh giá cách tìm hình ảnh người từ các mục tin tức

Cách tiếp cận	Độ chính xác	Mức độ hồi đáp	Số đo F1
Tìm theo chỉ mục	47,17	100,00	64,10
Phân tích cú pháp	59,08	99,61	74,17

Từ các kết quả đó, cách tiếp cận tìm theo chỉ mục đã cho mức độ hồi đáp bằng 100%. Điều này có được là do câu hỏi tìm người ứng với tất cả các hình ảnh có chứa tên người đó trong thông tin văn bản. Tuy nhiên, hơn một nửa kết quả đưa lại đã không chứa hình ảnh đúng của cá nhân mục tiêu. Kết quả là giá trị về độ chính xác chỉ xấp xỉ 47%. Việc áp dụng quy trình sàng lọc giúp nâng cao độ chính xác lên xấp xỉ 59%, trong khi mức độ hồi đáp hầu như vẫn giữ nguyên. Do đó, cách phân tích cú pháp giúp loại bỏ một số trường hợp khẳng định sai của cách tìm theo chỉ mục. Tóm lại, cách tiếp cận tìm theo chỉ mục đã cho hiệu suất theo số đo F1 bằng 64,10%. Khi dùng cách phân tích cú pháp, F1 đã tăng lên 74,17%. Do đó, các quy tắc cú

pháp có ích cho việc sàng lọc kết quả sai từ các hình ảnh mà cách tìm theo chỉ mục đem lại.

Kết luận và dự định tương lai

Chúng tôi đã đưa ra phương pháp phân tích cú pháp để tìm người trong các hình ảnh của mục tin. Mục đích chính của chúng tôi là nâng cao chất lượng kết quả tìm kiếm hình ảnh thu được khi dùng cách tìm theo chỉ mục dựa vào văn bản. Quá trình định nhãn thực thể có tên và phân tích cú pháp sơ bộ được áp dụng vào nội dung văn bản, ví dụ, đầu đề tin và chú thích hình ảnh để gán nhãn loại từ (POS) thích hợp cho các thực thể tên người và mọi thuật ngữ khác. Một bộ quy tắc cú pháp để nhận dạng người xuất hiện trong các hình ảnh được xây dựng từ tập hợp có định nhãn.

⁶ Penn Treebank Project, <http://www.cis.upenn.edu/treebank>

Nhìn ra thế giới

Chúng tôi sử dụng các quy tắc cú pháp này để sàng lọc các kết quả hình ảnh mà cách tìm theo chỉ mục đem lại. Dựa vào các kết quả đánh giá, phương pháp của chúng tôi đã nâng hiệu suất lên 10% theo số đo F1 so với phép tìm theo chỉ mục đơn giản.

Chúng tôi dự định áp dụng một số kỹ thuật nhận biết bằng máy như các trường

ngẫu nhiên có điều kiện (CRFs) để xây dựng một cách hữu hiệu các quy tắc cú pháp từ các tập hợp có định nhãn. Một tập hợp hình ảnh tin tức lớn hơn cũng sẽ được thu thập và sử dụng để đánh giá toàn diện hơn.

Vũ Văn Sơn (Dịch)

Nguồn: *the Role of Digital Libraries in a Time of Global Change*: ICADL 2010, LNCS 6102, pp. 216-225

Tham chiếu thư mục

1. Abney, S.: Parsing by chunks. In: Berwick, R., Abney, S., Tenny, C. (eds.) *Principle-Based Parsing*. Kluwer Academic Publishers, Dordrecht (1991)
2. Berg, T.L., Berg, A.C., Edwards, J., Maire, M., White, R., Yee-Whye, T., Learned-Miller, E., Forsyth, D.A.: Names and faces in the News. In: Proc. of the 2004 IEEE Conf. on Computer Vision and Pattern Recognition, pp.848-854 (2004)
3. Chinchor, N.: MUC-7 Named Entity Task Definition (Version 3.5) MUC-7, Fairfax, Virginia (1998)
4. Datta, R., Joshi, D., Li, J., Wang, J.Z.: Image Retrieval: Ideas, influences, and trends of the new age. *ACM Computing Surveys (CSUR)* 40(2), 1-60 (2008)
5. Edwards, J., White, R., Forsyth, D.: Words and Pictures in the News. In: Proc. of the HLT-NAACL 2003 workshop on learning word meaning from non-linguistic data, pp. 6-13 (2003)
6. He, X., Cai, D., Wen, J.-R., Ma, W.-Y., Zhang, H.-J.: Clustering and searching WWW images using link and page layout analysis. *ACM Trans. on Multimedia Computing, Communications, and Applications* 3(2) (2007)
7. Horster, E., Lienhart, R., Slaney, M.: Image Retrieval on large-scale image databases. In : Proc. of the 6th ACM int. conf. on image and video retrieval, pp. 17-24 (2007)
8. Kherfi, M.L. Ziou, D., Bernardi, A.: Image Retrieval from the World Wide Web Issues, Techniques, and Systems . *ACM Computing Surveys (CSUR)* 36(1), 35-67 (2004)
9. Kitahara, A, Joutou, T., Yanai, K.: Associating Faces and Names in Japanese Photo News Articles on the Web. In: Proc. of the 22nd Int. conf. on Advanced Information Networking and Applications – Workshops, pp. 1156-1161 (2008)
10. Liu, C., Jiang, S., Huang, Q.: Naming faces in broadcast news video by image google. In: Proc. of the 16th ACM int. conf. on multimedia, pp. 717-720 (2008)
11. Ozkan, D., Duygulu, P.: A Graph Based Approach for Naming Faces on News Photos. In Proc. of the 2006 IEEE Conf. on Computer Vision and Pattern Recognition, pp. 1477-1482 (2006)
12. Smeulders, A.W.M., Worring, M/, Santini, S., Gupta, A., Jain, R.: Content-Based Image Retrieval at the End of the Early Years. *IEEE Transactions on Patterns Analysis and Machine Intelligence* 22(12), 1349-1380 (2000)
13. Shrihari, R.K.: Automatic Indexing and Content-Based Retrieval of Captioned Images. *Computer* 28(9), 49-56 (1995)
14. Yagnik, J., Islam, A.: Learning people annotation from the web via consistency learning. In: Proc. of the int. workshop on multimedia information retrieval, pp. 285-290 (2007)
15. Yang, J., Hauptmann, A.G.: Naming every individual in news video monologues. In : Proc. of the 12th ACM int. conf. on multimedia, pp. 580-587 (2004)
16. Zhao, W., Chellappa, R., Phillips P.J., Rosenfeld, A.: Face recognition: A literature survey: *ACM Computing Surveys (CSUR)* 35(4), 399-458 (2003)