



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0053168

(51)^{2022.01} G06F 16/90

(13) B

(21) 1-2023-08082

(22) 15/11/2023

(45) 25/11/2025 452

(43) 25/04/2024 433

(73) TẬP ĐOÀN CÔNG NGHIỆP - VIỆN THÔNG QUÂN ĐỘI (VN)

Lô D26 Khu đô thị mới Cầu Giấy, phường Yên Hoà, quận Cầu Giấy, thành phố Hà Nội

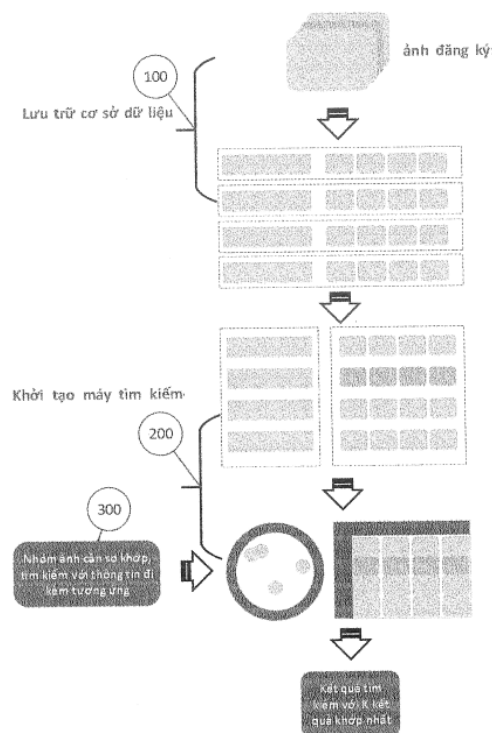
(72) VŨ ĐỨC CHÍNH (VN); LÊ VĂN BẰNG (VN); NGUYỄN HOÀNG NAM (VN);
TRƯƠNG HUY HOÀNG (VN).

(74) Công ty TNHH NACILAW (NACILAW)

(54) PHƯƠNG PHÁP SO KHỚP TÌM KIẾM TỐC ĐỘ CAO ĐỐI VỚI CƠ SỞ DỮ LIỆU
HỖN HỢP QUY MÔ LỚN

(21) 1-2023-08082

(57) Phương pháp so khớp tìm kiếm tốc độ cao với cơ sở dữ liệu hỗn hợp quy mô lớn bao gồm quá trình trích xuất và lưu trữ đặc trưng ảnh đăng ký của người dùng, tạo thành cơ sở dữ liệu, và là tham chiếu khi so khớp, tìm kiếm; Quá trình khởi tạo máy tìm kiếm, trong đó khởi tạo hai đối tượng trong máy tìm kiếm bao gồm đối tượng so khớp, tính toán khoảng cách và đối tượng bảng dữ liệu để xác định đối tượng tìm kiếm mục tiêu cho từng mẫu ảnh được đưa vào; Quá trình so khớp, tìm kiếm đối tượng, bao gồm quá trình tính toán khoảng cách giữa đặc trưng và xác định mảng mục tiêu, cuối cùng đối chiếu và đưa ra kết quả cuối cùng. Phương pháp đề xuất đã thể hiện được khả năng tính toán chính xác cũng như hiệu năng rất cao, đặc biệt trong các trường hợp xử lý dữ liệu quy mô lớn.



Hình 1

Lĩnh vực kỹ thuật được đề cập.

Sáng chế đề cập đến phương pháp so khớp tìm kiếm tốc độ cao với cơ sở dữ liệu hỗn hợp quy mô lớn. Cụ thể phương pháp so khớp tìm kiếm tốc độ cao với cơ sở dữ liệu hỗn hợp quy mô lớn được ứng dụng trong lĩnh vực trí tuệ nhân tạo và thị giác máy tính, giúp xử lý và tìm kiếm đối tượng trong các bài toán nhận diện đối tượng.

Tình trạng kỹ thuật của sáng chế

Đối với những hệ thống nhỏ, ví dụ như dùng để nhận diện khuôn mặt cho nhân sự ra vào một tòa nhà nhất định, cơ sở dữ liệu có tính đồng nhất cao, đó là những dữ liệu chỉ thuộc một đơn vị với một nhà cung cấp dịch vụ duy nhất. Thông thường, đối với trường hợp này, phương pháp so khớp vẫn là tính toán những khoảng cách giữa cơ sở dữ liệu có sẵn và đặc trưng khuôn mặt đang xuất hiện trước máy quay, những mẫu được xác định là khớp là những mẫu có khoảng cách đặc trưng thấp.

Những hệ thống nhỏ lẻ, quy mô dữ liệu nhỏ và tính đồng nhất cao thì vai trò của máy tìm kiếm không cao, và thời gian vận hành các toán tử tìm kiếm cũng nhỏ, không đáng kể, do chỉ thực hiện một phép tính duy nhất cho một (hoặc một nhóm đặc trưng). Ngoài ra, cũng do quy mô dữ liệu nhỏ, có nhiều cách tiếp cận khi tính toán khoảng cách, thậm chí đối với những thiết bị xử lý tại biên còn có thể sử dụng phương án tính toán tuần hoàn (sử dụng ngôn ngữ lập trình Java hoặc C++) mà vẫn bảo đảm thời gian tính toán trong phạm vi cho phép.

Tuy nhiên, với trường hợp ứng dụng vận hành dịch vụ nhận diện đối tượng sử dụng điện toán đám mây cho quy mô nhà cung cấp dịch vụ rộng rãi (bên thứ ba), người dùng lớn, vị trí lắp đặt linh hoạt và cơ sở dữ liệu lớn, thì các mảng dữ liệu tìm kiếm đối với từng mẫu ảnh cần nhận diện là khác nhau, ví dụ như một người dùng sở hữu hai tòa nhà, mỗi tòa sẽ có cơ sở dữ liệu khác nhau, khi có đối tượng đứng trước camera thì chỉ cần tìm kiếm mảng cơ sở dữ liệu của tòa nhà đó, mà không lấy kết quả so khớp đối với phần còn lại của cơ sở dữ liệu.

Bản chất kỹ thuật của sáng chế

Mục đích của sáng chế là khắc phục những vấn đề kỹ thuật còn tồn tại, nâng cao hiệu năng trong quá trình so khớp tìm kiếm và linh hoạt đối với nhiều nhóm dữ liệu được tìm kiếm đồng thời, trong bài toán nhận diện đối tượng trong ảnh theo phương pháp đối chiếu, tính toán khoảng cách, bằng việc đề xuất phương pháp so khớp tìm kiếm tốc độ cao với cơ sở dữ liệu hỗn hợp quy mô lớn. Phương pháp trong sáng chế cũng giúp cho việc xử lý linh hoạt trên các thiết bị có cấu hình khác nhau.

Trong sáng chế này, phương pháp so khớp tìm kiếm tốc độ cao với cơ sở dữ liệu hỗn hợp quy mô lớn được thực hiện dựa trên các quá trình như sau:

- **Quá trình xử lý ảnh đăng ký:** sử dụng để trích xuất đặc trưng của đối tượng (có thể là khuôn mặt, hoặc hình dáng người, phương tiện) và cập nhật, lưu trữ vào trong cơ sở dữ liệu theo cấu trúc phân mảnh theo hàng và cột.

- **Quá trình khởi tạo máy tìm kiếm:** thực hiện quá trình đọc cơ sở dữ liệu, bao gồm việc tải dữ liệu đặc trưng và thông tin kèm theo, và đóng gói vào hai đối tượng, thứ nhất là đối tượng so khớp ma trận và thứ hai là đối tượng bảng dữ liệu. Hai đối tượng này cấu thành máy tìm kiếm, và thực hiện các tác vụ xác định mảng dữ liệu mục tiêu cũng như xác định khoảng cách giữa đặc trưng ảnh cần tìm kiếm với ma trận đặc trưng của cơ sở dữ liệu có sẵn.

- **Quá trình thực hiện tìm kiếm (cụm toán tử vận hành máy tìm kiếm):** bao gồm trích xuất đặc trưng từ ảnh cần so khớp, tìm kiếm, và thực hiện việc truy vấn, đối chiếu với cơ sở dữ liệu có sẵn, tìm ra những mẫu dữ liệu khớp nhất (có khoảng cách nhỏ nhất).

Mô tả vắn tắt các hình vẽ

Hình 1 là hình vẽ mô tả tổng thể các quá trình thực hiện phương pháp so khớp tìm kiếm tốc độ cao với cơ sở dữ liệu hỗn hợp quy mô lớn được đề xuất trong sáng chế.

Hình 2 là hình vẽ mô tả quá trình trích xuất đặc trưng và lưu trữ cơ sở dữ liệu đối với các mẫu ảnh sử dụng để đăng ký để nhận diện đối tượng.

Hình 3 là hình vẽ mô tả ví dụ về mối quan hệ giữa (1) mã định danh đối tác vận hành; (2) mã định danh người dùng; (3) nhóm tìm kiếm, và phương án sử dụng dịch vụ tìm kiếm trên điện toán đám mây.

Hình 4 là hình vẽ mô tả quá trình khởi tạo máy tìm kiếm.

Hình 5 là hình vẽ mô tả ví dụ về cách mã hóa nhóm mã xác định nhóm tìm kiếm từ định dạng chuỗi ký tự sang số nguyên.

Hình 6 là hình vẽ mô tả quá trình so khớp, nhận diện của mẫu ảnh cần nhận diện thông qua máy tìm kiếm.

Hình 7 là hình vẽ mô tả ví dụ về quá trình lọc mảng mục tiêu tìm kiếm của từng mẫu ảnh cần nhận diện.

Mô tả chi tiết các phương án thực hiện sáng chế.

Để giải quyết vấn đề tìm kiếm so khớp phục vụ nhận diện thường có hai cách tiếp cận thông dụng:

+ Một là khởi tạo các mảng dữ liệu riêng lẻ thành các đối tượng tìm kiếm khác nhau.

+ Hai là so khớp chung trên toàn bộ khối dữ liệu, sau đó xác định mảng mục tiêu của từng mẫu ảnh, kết quả khoảng cách được tính toán và so khớp sẽ chỉ sử dụng mảng dữ liệu đó.

Đối với cách thứ nhất, để điều phối từng mẫu ảnh (hoặc một nhóm ảnh) vào từng đối tượng tìm kiếm được khởi tạo riêng rẽ sẽ mất khá nhiều thời gian. Ngoài ra, các tổ hợp mảng dữ liệu sẽ sinh ra sự chồng lấn trong lưu trữ đối tượng, gây lãng phí tài nguyên.

Đối với cách thứ hai, tận dụng năng lực tính toán theo khối của những công cụ có sẵn như Numpy, Cupy... để tăng tốc độ tính toán khoảng cách, đồng thời, dựa vào cách phân tích theo bảng dữ liệu để đánh dấu mảng mục tiêu cho từng mẫu cần tìm kiếm. Đây cũng là căn cứ để thiết kế song song hai đối tượng trong máy tìm kiếm trong sáng chế này.

Đặt yêu cầu của máy tìm kiếm trong trường hợp vận hành với cơ sở dữ liệu phức tạp và quy mô lớn, sáng chế đề xuất ra phương pháp so khớp tìm kiếm tốc độ cao với cơ sở dữ liệu hỗn hợp quy mô lớn, bảo đảm tìm kiếm chính xác và hiệu năng tối ưu nhất.

Tham chiếu hình 1, quá trình lưu trữ và tìm kiếm bao gồm ba giai đoạn xử lý lớn, cụ thể là:

- + quá trình trích xuất và lưu trữ đặc trưng ảnh đăng ký của người dùng, tạo thành cơ sở dữ liệu, và là tham chiếu khi so khớp, tìm kiếm.

- + quá trình khởi tạo máy tìm kiếm, trong đó khởi tạo hai đối tượng trong máy tìm kiếm bao gồm đối tượng so khớp, tính toán khoảng cách và đối tượng bảng dữ liệu để xác định đối tượng tìm kiếm mục tiêu cho từng mẫu ảnh được đưa vào.

- + quá trình so khớp, tìm kiếm đối tượng, bao gồm quá trình tính toán khoảng cách giữa đặc trưng và xác định mảng mục tiêu, cuối cùng đối chiếu và đưa ra kết quả cuối cùng.

Theo đó, quá trình trích xuất và lưu trữ đặc trưng được thực hiện như sau.

Tham chiếu Hình 2, quá trình này bao gồm hai bước:

Ảnh 101 dùng để khởi tạo cơ sở dữ liệu của người dùng được đăng ký qua kênh của đối tác vận hành (sử dụng dịch vụ điện toán đám mây triển khai giải pháp cho người dùng cuối). Khi đăng ký ảnh đối tượng, quá trình trích xuất đặc trưng và cập nhật thông tin kèm theo được mô tả như sau:

- + Đối với ảnh, đối tượng dùng để nhận diện có thể là khuôn mặt, hình dáng người hoặc phương tiện, đối tượng ảnh này được đưa vào mô hình học sâu được huấn luyện từ trước để trích xuất đặc trưng, đặc trưng này là một véc-tơ với định dạng dãy phẩy động 32 bit.

- + Các thông tin kèm theo được xác định khi máy chủ nhận yêu cầu cập nhật cơ sở dữ liệu từ đối tác vận hành, thông tin bắt buộc bao gồm: mã định danh đối tác vận hành; mã định danh người dùng; nhóm tìm kiếm; và nhãn đối tượng.

Các trường thông tin này đều sử dụng định dạng ký tự, trong đó mã đối tác vận hành là chuỗi ký tự xác định đơn vị (hoặc cá nhân) phát hành dịch vụ điện toán đám mây về nhận diện đối tượng. Mã vận hành này do nhà phát triển, cung cấp dịch vụ điện toán đám mây cấp cho đơn vị phát hành dịch vụ.

Mã định danh người dùng là chuỗi ký tự xác định người dùng cuối, là người đang trực tiếp sử dụng dịch vụ nhận diện đối tượng theo phương thức điện toán đám mây được phát hành bởi một đối tác vận hành nhất định. Mã định danh người dùng do đơn vị phát hành dịch vụ cấp cho người dùng, sau khi đăng ký và kiểm duyệt.

Nhóm tìm kiếm đối tượng được mô tả là một chuỗi ký tự thể hiện mảng dữ liệu được đánh dấu khác nhau của cùng một người dùng, sử dụng trong trường hợp người dùng dịch vụ có nhu cầu phân mảnh dữ liệu tại các vị trí lắp đặt camera nhận diện khác nhau, ví dụ như *Căn hộ số 1, Căn hộ số 2...* Nhóm tìm kiếm do người dùng tự định nghĩa, được lựa chọn khi khởi tạo và cập nhật cơ sở dữ liệu.

Nhãn đối tượng là chuỗi ký tự định danh đối tượng, có thể là họ và tên, hoặc một mã ký tự bất kỳ, chỉ cần bảo đảm những mẫu của cùng một đối tượng cùng nhãn là được. Nhãn đối tượng do người dùng tự định nghĩa (hoặc có thể qua kênh mã hóa của đơn vị phát hành dịch vụ), được gửi lên khi đăng ký ảnh mới cho đối tượng cần nhận diện.

Mối quan hệ giữa *điện toán đám mây* $\rightarrow$ *đối tác vận hành* $\rightarrow$ *người dùng* $\rightarrow$ *nhóm tìm kiếm* $\rightarrow$ *đối tượng cần nhận diện* được mô tả như trong Hình 3.

Đặc trưng học sâu và thông tin kèm theo của đối tượng ảnh được đăng ký trong cơ sở dữ liệu được lưu trữ tại 102. Dữ liệu được lưu trữ theo cấu trúc phân mảnh, hay nói cách khác, mỗi một ảnh được đưa vào đăng ký sẽ lưu một tệp nhị phân (binary file) tương ứng. Trong tệp nhị phân đó, bao gồm cả đặc trưng ảnh và thông tin kèm theo dưới dạng đối tượng từ điển (dictionary object). Các tệp này được lưu như các tệp vật lý khác, trong thư mục được xác định trước trong ổ cứng máy tính. Việc lưu trữ bởi tệp nhị phân sẽ giúp hệ thống đọc dữ liệu một cách nhanh nhất, mà vẫn giữ nguyên được định dạng của dữ liệu. Khi khởi tạo máy tìm kiếm, đầu tiên sẽ đọc những tệp nhị phân này, để làm tiền đề cho quá trình tìm kiếm.

Quá trình khởi tạo máy tìm kiếm được thực hiện như sau:

Tham chiếu hình 4, việc khởi tạo máy tìm kiếm bắt đầu từ việc đọc tệp nhị phân từ 201 từ trong cơ sở dữ liệu. Dữ liệu định dạng từ điển được đọc từ tệp nhị phân, với đặc trưng học sâu của ảnh và thông tin đi kèm, lúc này hai dạng thông tin này vẫn đi kèm với nhau trong một đối tượng (từ điển). Sau đó, sử dụng các toán tử đơn giản thao tác, có thể tách phần khóa và giá trị của từ điển để xử lý riêng.

Đối với các đặc trưng học sâu của ảnh, các véc-tơ được ghép lại với nhau theo thứ tự từ trên xuống để thành một mảng tại 202. Với một cơ sở dữ liệu với số chiều của đặc trưng là $1 \times N$, và có M mẫu ảnh đã được khởi tạo, thì mảng (array) được tạo ra có kích thước là $M \times N$. Quá trình này khá đơn giản, đơn thuần là tải đặc trưng vào trong bộ nhớ máy tính theo mảng.

Đối với các thông tin kèm theo, đang được lưu trữ dưới dạng chuỗi ký tự cho từng trường thông tin. Những thông tin này cũng được ghép lại như đối với đặc trưng ảnh tại 203, tuy nhiên khác với định dạng mảng, thông tin kèm theo được đưa vào đối tượng dữ liệu dạng bảng, để tiện trong quá truy đối chiếu và truy vấn.

Tuy nhiên, đối với dữ liệu dạng chuỗi ký tự, tốc độ truy vấn khá chậm, sẽ làm ảnh hưởng đến hiệu năng chung của toàn hệ thống. Do đó, trong sáng chế này, đề xuất phương án chuyển các mã định danh là chuỗi ký tự sang dạng số nguyên, cụ thể được mô tả trong hình 5.

Trong hình 5, thể hiện quá trình mã hóa chuỗi ký tự mã định danh sang định dạng số nguyên. Trong quá trình này, thay vì trực tiếp sử dụng chuỗi ký tự, thì sử dụng vị trí (index) của chuỗi ký tự định danh đó trong toàn bộ mã đối tác vận hành/mã người dùng/nhóm cơ sở dữ liệu, bắt đầu từ 0 và kết thúc là (tổng số chuỗi ký tự - 1), và được chuyển đổi theo thứ tự từ mã đối tác vận hành 2031 $\rightarrow$ mã người dùng 2032 $\rightarrow$ nhóm cơ sở dữ liệu 2033. Ví dụ, khi chuyển đổi mã đối tác vận hành, nếu tổng số đối tác vận hành bao gồm ["si001", "si010", "si101", "si200"] thì đối với mã đối tác vận hành "si010" sẽ được chuyển sang số nguyên bằng 1. Cứ như vậy, các loại mã định danh trên đã được chuyển đổi, bộ số nguyên đầy đủ cho mỗi mẫu dữ liệu có số phần tử bằng số chủng loại mã, bảng 3 2024.

Kết quả cuối cùng của quá trình khởi tạo máy tìm kiếm là 03 đối tượng lưu trữ những thông tin khác nhau, phục vụ cho việc so khớp, tìm kiếm.

Trong đó, tại 204, mảng với kích thước $M \times N$, được tải vào trong các khối xử lý dữ liệu theo các công cụ có sẵn như numpy (để xử lý trên vi xử lý CPU) hoặc cupy (để xử lý trên bộ vi xử lý đồ họa GPU) để thực hiện các phép truy vấn hoặc các toán tử tính toán ma trận một cách nhanh chóng, hiệu quả cao.

Bên cạnh đó theo 205, thông tin kèm theo trước tiên được phân tách và mã hóa sang số nguyên (với 03 chủng loại thông tin định danh), sau đó được tách ra thành hai đối tượng, thứ nhất là đối tượng bảng dữ liệu thông tin kèm theo và thứ hai là đối tượng lưu chuỗi nhãn của cơ sở dữ liệu.

Quá trình thực hiện so khớp, tìm kiếm được thực hiện như sau:

Khi vận hành hệ thống nhận diện đối tượng theo thời gian thực, những mẫu ảnh được chụp lại và tiến hành tìm kiếm theo 300. Đầu tiên, những ảnh cần nhận diện này được phân tích ra hai luồng dữ liệu, đó là từ 301, tính đặc trưng học sâu của ảnh đối tượng (mô hình đồng bộ với 100), và từ 302 xác định nhóm mã xác định tìm kiếm, bao gồm mã đối tác vận hành, mã người dùng và nhóm cơ sở dữ liệu. Những thông tin xác định mã tìm kiếm này sau đó được chuyển sang dạng số nguyên với nguyên lý như tại 203.

Do một số đặc thù của hệ thống tìm kiếm như sau:

+ Những mẫu ảnh cần so khớp là không đồng nhất về mã đối tác vận hành, mã người dùng hoặc nhóm cơ sở dữ liệu cần tìm kiếm. Ngoài ra, các thông tin này hoàn toàn là ngẫu nhiên về thứ tự khi nhận được yêu cầu xử lý từ các camera tại thực địa triển khai.

+ Để tận dụng tối đa khả năng tính toán ma trận của máy tính, toàn bộ cơ sở dữ liệu được đăng ký và khởi tạo thành một mảng khối với nhóm thông tin xác định mảng tìm kiếm khác nhau.

Nên dẫn đến hệ quả đó là, mẫu ảnh cần tìm kiếm không thể sử dụng kết quả so khớp (tính toán khoảng cách giữa các véc-tơ) trên toàn bộ cơ sở dữ liệu như một kết quả tìm kiếm cuối cùng, mà kết quả tính toán có hiệu lực là những khoảng cách véc-tơ trong từng nhóm tìm kiếm khác nhau (cho mỗi lần, mỗi nhóm ảnh cần nhận diện đối tượng). Một mẫu ảnh bất kỳ sẽ chỉ được lựa chọn tìm kiếm nhận diện trên phạm vi về mã đối tác vận hành, mã người dùng và nhóm cơ sở dữ

liệu nhất định. Trong quá trình này, tuy cơ sở dữ liệu chứa toàn bộ thông tin đăng ký dịch vụ nhận diện đối tượng, nhưng sẽ không phát sinh khả năng lỗi nhận diện nhầm, ví dụ như nhầm người ở Căn hộ A sang người đăng ký tại Căn hộ B.

Với ý tưởng xử lý như vậy, tại 303, các mẫu cần nhận diện cần xác định được mảng tìm kiếm (trong tổng thể cơ sở dữ liệu). Trong quá trình này, thông tin mã xác định tìm kiếm của nhóm ảnh cần nhận diện được so sánh với dữ liệu bảng dữ liệu thông tin kèm theo của cơ sở dữ liệu để xác định mảng tìm kiếm của từng mẫu ảnh trong nhóm ảnh trên. Cụ thể, mã xác định tìm kiếm của từng mẫu được đối chiếu với mã số nguyên của cơ sở dữ liệu, để tìm những số thứ tự dữ liệu cần so khớp. Gọi M là quy mô của cơ sở dữ liệu, mã số nguyên của cơ sở dữ liệu là ma trận R có kích thước $(M \times 3)$, bên cạnh đó với X mẫu ảnh cần tìm kiếm đồng thời, mã số nguyên tương ứng là T có kích thước $(X \times 3)$, thì ma trận đánh dấu Ψ có kích thước $(M \times X)$, với giá trị chỉ có 0 hoặc 1. Trong đó, tại tọa độ (m, x) có giá trị 0 biểu thị mẫu số m của cơ sở dữ liệu không nằm trong phạm vi tìm kiếm của mẫu thứ x của nhóm tìm kiếm, và ngược lại. Công thức xác định mảng:

$$\Psi(m, x) = \begin{cases} 0 & \text{if } M(m, :) \neq X(x, :) \\ 1 & \text{others} \end{cases}$$

Lúc này, ma trận đánh dấu với các giá trị 0 và 1 sẽ xác định phạm vi tìm kiếm của từng mẫu ảnh cần nhận diện, đại diện cho phạm vi tìm kiếm của mỗi mẫu ảnh là một véc-tơ có độ dài $1 \times M$, vị trí có giá trị bằng 1 là số thứ tự của dữ liệu cần đối chiếu, ví dụ được mô tả như trong hình 7.

Đồng thời, tại 304, nhóm đặc trưng từ 302 được đưa vào đối tượng so khớp tìm kiếm 204 để tiến hành tính toán khoảng cách, ở đây, sử dụng công thức tính *cô-sin* để tính toán khoảng cách giữa hai véc-tơ, giá trị khoảng cách *cô-sin* càng lớn càng nói lên hai véc-tơ đó có độ tương đồng càng cao. Trong quá trình tính toán khoảng cách, sử dụng các toán tử xử lý trên mảng dữ liệu (xử lý theo khối), với tốc độ tính toán cao, đặc biệt khi sử dụng trên GPU. Cụ thể, ma trận khoảng cách $distance(m, x)$ được tính theo công thức:

$$distance(m, x) = dot(f(M, :N), g(X, N))$$

Trong đó, $f(M, :N)$ và $g(X, N)$ lần lượt là đối tượng ma trận đặc trưng của cơ sở dữ liệu và các mẫu ảnh cần nhận diện, với N là số chiều của véc-tơ đặc trưng.

Tiếp theo, để lấy loại bỏ những kết quả so khớp không đúng, sử dụng ma trận đánh dấu để xác định các giá trị khoảng cách hữu dụng đối với từng mẫu cần nhận diện, tại 305, các giá trị ma trận khoảng cách không hữu dụng được đưa giá trị về 0 bởi phép nhân giữa hai ma trận khoảng cách và ma trận đánh dấu, cụ thể:

$$distance'(m, x) = matmul(distance(m, x), \Psi(m, x))$$

Theo từng cột của $distance'(m, x)$, phần tử nào có giá trị càng cao, càng khớp với mẫu cần nhận diện tương ứng, theo số thứ tự. Do đó, để lấy K mẫu dữ liệu khớp nhất (với K được

thiết lập từ trước), thực hiện sắp xếp giá trị từ lớn đến nhỏ, tại **306**, sẽ được số thứ tự $indices(k=1 \rightarrow K)$ và giá trị khoảng cách tương ứng $corr(k=1 \rightarrow K)$ của K mẫu khớp nhất đó:

$$indices(k) = argsort(distance')[: K]$$

$$corr(k) = sort(distance')[: K]$$

Và nhãn của những mẫu dữ liệu khớp này $label(k)$ được xác định bởi:

$$label(k) = L(indices)$$

Với L là đối tượng lưu chuỗi nhãn dữ liệu được xác định tại **205**.

Kết thúc quá trình tìm kiếm, hệ thống trả lại kết quả cuối cùng bao gồm $label(k)$ và $corr(k)$ về địa chỉ máy chủ phát ra yêu cầu cần nhận diện.

Ví dụ thực hiện sáng chế

Phương pháp được đề xuất trong sáng chế này có khả năng áp dụng rộng rãi đối với các bài toán nhận diện đối tượng, ví dụ như nhận diện khuôn mặt, hình dáng người, vân tay, hoặc cho cho vấn đề tái nhận diện phương tiện, cũng như các bài toán nhận diện đối tượng khác.

Đối với bài toán nhận diện khuôn mặt, một số thí nghiệm đã được thực hiện và tiến hành phân tích hiệu quả, cụ thể với môi trường thí nghiệm như sau:

- + Quy mô dữ liệu 10000 mẫu, được khởi tạo như trong quá trình mô tả trong sáng chế, với đặc trưng học sâu và thông tin học sâu được lưu trữ theo tệp nhị phân.

- + Số tổ hợp nhóm tìm kiếm (mã đối tác vận hành + mã định danh người dùng + nhóm tìm kiếm) lên đến 20.

- + Số mẫu tìm kiếm đồng thời là 100, với tổ hợp nhóm tìm kiếm hoàn toàn ngẫu nhiên, và bảo đảm các tổ hợp này đều tồn tại trong cơ sở dữ liệu.

- + Số chiều của véc-tơ đặc trưng là 512, với định dạng dấu phẩy động 32 bit.

- + Vận hành trên GPU NVIDIA 3070Ti.

Đối với mỗi lần tìm kiếm, thực hiện một triệu phép so khớp đồng thời, thời gian trung bình cho mỗi phép so khớp (tính từ sau khi trích xuất đặc trưng đến lúc đưa ra kết quả cuối cùng) là 30 mi-cờ-rô giây với kết quả so khớp hoàn toàn chính xác (100%). Kết quả này có độ chính xác tương đồng nhưng hiệu năng nhanh hơn rất nhiều so với các công cụ có tính năng tương đồng (trên cùng điều kiện thí nghiệm), với mỗi lượt so khớp trung bình mất 15 mi-li giây.

Lợi ích (hiệu quả có thể đạt được bởi sáng chế)

Phương pháp được đề cập trong sáng chế đã được sử dụng trong các giải pháp trí tuệ nhân tạo và thị giác máy tính, phục vụ cho các sản phẩm giám sát an ninh tòa nhà, giám sát an ninh công cộng, tìm kiếm đối tượng diện rộng của Tổng công ty Công nghiệp Công nghệ cao Viettel.

Đặc biệt, phương pháp này là một công nghệ mang tính chất cốt lõi trong sản phẩm dịch vụ điện toán đám mây với các dịch vụ vận hành với dữ liệu trực tiếp từ camera, trả về kết quả tính toán theo thời gian thực. Trong đó, các dịch vụ đã được triển khai như nhận diện khuôn mặt,

nhận diện hình dáng người, nhận diện phương tiện giao thông, và tìm kiếm đối tượng trên diện rộng.

Phương pháp được đề xuất có hiệu năng rất cao, tận dụng được khả năng tính toán song song của hạ tầng triển khai với phương án xử lý theo khối kết hợp với ma trận đánh dấu. Kết quả thí nghiệm và kết quả triển khai thực địa (trong sản phẩm) đã cho thấy sự vượt trội về độ chính xác cũng như thời gian phản hồi kết quả của phương pháp tìm kiếm.

Yêu cầu bảo hộ

1. Phương pháp so khớp tìm kiếm tốc độ cao với cơ sở dữ liệu hỗn hợp quy mô lớn, phương pháp bao gồm ba quá trình được thực hiện lần lượt, gồm:

quá trình xử lý ảnh đăng ký: sử dụng để trích xuất đặc trưng của đối tượng (có thể là khuôn mặt, hoặc hình dáng người, phương tiện) và cập nhật, lưu trữ vào trong cơ sở dữ liệu theo cấu trúc phân mảnh theo hàng và cột; theo đó, ảnh dùng để khởi tạo cơ sở dữ liệu của người dùng được đăng ký qua kênh của đối tác vận hành (sử dụng dịch vụ điện toán đám mây triển khai giải pháp cho người dùng cuối), sử dụng cho bài toán nhận diện đối tượng;

quá trình khởi tạo máy tìm kiếm: thực hiện quá trình đọc cơ sở dữ liệu, bao gồm việc tải dữ liệu đặc trưng và thông tin kèm theo, và đóng gói vào hai đối tượng, thứ nhất là đối tượng so khớp ma trận và thứ hai là đối tượng bảng dữ liệu; hai đối tượng này cấu thành máy tìm kiếm, và thực hiện các tác vụ xác định mảng dữ liệu mục tiêu cũng như xác định khoảng cách giữa đặc trưng ảnh cần tìm kiếm với ma trận đặc trưng của cơ sở dữ liệu có sẵn; quá trình này bao gồm:

đọc và tách thông tin từ tệp nhị phân, tải dữ liệu là đặc trưng học sâu và thông tin đi kèm từ cơ sở dữ liệu được chỉ định từ trước;

tạo đối tượng so khớp, tìm kiếm: đối tượng so khớp, tìm kiếm được tạo ra bởi quá trình ghép xâu chuỗi đặc trưng học sâu của toàn bộ mẫu có trong cơ sở dữ liệu theo chiều từ trên xuống thành một mảng thống nhất; mảng đặc trưng này được tải vào trong đối tượng *numpy* để xử lý song song theo khối trên CPU và trong đối tượng *cupy* để xử lý song song theo khối trên GPU;

tạo đối tượng bảng dữ liệu nhãn đối tượng: được tạo ra sau khi tách thông tin kèm theo với phương thức “khóa: giá trị” của từ điển, nhãn đối tượng được xâu chuỗi thành đối tượng bảng, với số thứ tự đồng bộ với vị trí của mảng đặc trưng học sâu;

tạo đối tượng bảng dữ liệu thông tin kèm theo: được tách biệt và đồng bộ thứ tự với bảng dữ liệu nhãn sau khi tách thông tin kèm theo với phương thức “khóa: giá trị”;

quá trình thực hiện tìm kiếm (cụm toán tử vận hành máy tìm kiếm): bao gồm quá trình trích xuất đặc trưng từ ảnh cần so khớp tìm kiếm và thực hiện việc truy vấn, đối chiếu với cơ sở dữ liệu có sẵn, tìm ra những mẫu dữ liệu khớp nhất (có khoảng cách nhỏ nhất), quá trình này bao gồm việc:

trích xuất thông tin từ những ảnh cần nhận diện: thông tin được phân tích ra hai luồng dữ liệu: tính đặc trưng học sâu của ảnh đối tượng (mô hình đồng bộ với mô hình học sâu khi tạo cơ sở dữ liệu) và xác định nhóm mã xác định tìm kiếm, bao gồm *mã đối tác vận hành*, *mã người dùng* và *nhóm cơ sở dữ liệu*;

xác định mục tiêu tìm kiếm của từng mẫu ảnh trong nhóm cần nhận diện: thông tin mã xác định tìm kiếm của nhóm ảnh cần nhận diện được so sánh với dữ liệu bảng dữ liệu thông tin kèm theo của cơ sở dữ liệu để xác định mảng tìm kiếm của từng mẫu ảnh

trong nhóm ảnh trên; cụ thể, mã xác định tìm kiếm của từng mẫu được đối chiếu với mã số nguyên của cơ sở dữ liệu, để tìm những số thứ tự dữ liệu cần so khớp; gọi M là quy mô của cơ sở dữ liệu, mã số nguyên của cơ sở dữ liệu là ma trận R có kích thước $(M \times 3)$, bên cạnh đó với X mẫu ảnh cần tìm kiếm đồng thời, mã số nguyên tương ứng là T có kích thước $(X \times 3)$, thì ma trận đánh dấu Ψ có kích thước $(M \times X)$, với giá trị chỉ có 0 hoặc 1; trong đó, tại tọa độ (m, x) có giá trị 0 biểu thị mẫu số m của cơ sở dữ liệu không nằm trong phạm vi tìm kiếm của mẫu thứ x của nhóm tìm kiếm, và ngược lại; cách thức xác định mảng: $\Psi(m, x) = \begin{cases} 0 & \text{if } M(m, :) \neq X(x, :) \\ 1 & \text{others} \end{cases}$, lúc này, ma trận đánh dấu với các giá trị 0 và 1 sẽ xác định phạm vi tìm kiếm của từng mẫu ảnh cần nhận diện, đại diện cho phạm vi tìm kiếm của mỗi mẫu ảnh là một véc-tơ có độ dài $1 \times M$, vị trí có giá trị bằng 1 là số thứ tự của dữ liệu cần đối chiếu;

tính toán khoảng cách giữa đặc trưng nhóm ảnh cần nhận diện với đối tượng so khớp, tìm kiếm; xác định nhóm giá trị khoảng cách có hiệu lực đối với từng mẫu ảnh cần nhận diện, cụ thể: nhóm đặc trưng từ nhóm ảnh cần nhận diện được đưa vào đối tượng so khớp tìm kiếm để tiến hành tính toán khoảng cách bằng việc sử dụng công thức tính *cô-sin* để tính toán khoảng cách giữa hai véc-tơ, giá trị khoảng cách *cô-sin* càng lớn càng nói lên hai véc-tơ đó có độ tương đồng càng cao; trong quá trình tính toán khoảng cách, sử dụng các toán tử xử lý trên mảng dữ liệu (xử lý theo khối), với tốc độ tính toán cao, đặc biệt khi sử dụng trên GPU; cụ thể, ma trận khoảng cách $distance(m, x)$ được xác định: $distance(m, x) = dot(f(M, :N), g(X, N))$, trong đó, $f(M, :N)$ và $g(X, N)$ lần lượt là đối tượng ma trận đặc trưng của cơ sở dữ liệu và các mẫu ảnh cần nhận diện, với N là số chiều của véc-tơ đặc trưng; từ kết quả tính toán khoảng cách và ma trận đánh dấu, thực hiện lọc kết quả so khớp và đưa ra kết quả nhận diện cuối cùng;

cuối cùng, căn cứ vào giá trị khoảng cách có hiệu lực, thực hiện sắp xếp, lấy vị trí nhân tương ứng và trả về kết quả cuối cùng, cụ thể:

đầu tiên, loại bỏ những kết quả so khớp không đúng, sử dụng ma trận đánh dấu để xác định các giá trị khoảng cách hữu dụng đối với từng mẫu cần nhận diện, các giá trị ma trận khoảng cách không hữu dụng được đưa giá trị về 0 bởi phép nhân giữa hai ma trận khoảng cách và ma trận đánh dấu, cụ thể: $distance'(m, x) = matmul(distance(m, x), \Psi(m, x))$;

theo từng cột của $distance'(m, x)$, phần tử nào có giá trị càng cao, càng khớp với mẫu cần nhận diện tương ứng, theo số thứ tự; để lấy K mẫu dữ liệu khớp nhất (với K được thiết lập từ trước), thực hiện sắp xếp giá trị từ lớn đến nhỏ, sẽ được số thứ tự $indices(k=1 \rightarrow K)$ và giá trị khoảng cách tương ứng $corr(k=1 \rightarrow K)$ của K mẫu khớp nhất đó:

$$indices(k) = argsort(distance')[: K]$$

$$corr(k) = sort(distance')[: K]$$

và nhãn của những mẫu dữ liệu khớp này $label(k)$ được xác định bởi:

$$label(k) = L(indices);$$

với L là đối tượng lưu chuỗi nhãn dữ liệu được xác định từ lúc khởi tạo máy tìm kiếm; cuối cùng, kết thúc quá trình tìm kiếm, hệ thống trả lại kết quả.

2. Phương pháp so khớp tìm kiếm tốc độ cao với cơ sở dữ liệu hỗn hợp quy mô lớn theo điểm 1 trong đó, những thông tin cần xác định trong quá trình xử lý ảnh đăng ký bao gồm:

đối với ảnh, đối tượng dùng để nhận diện có thể là khuôn mặt, hình dáng người hoặc phương tiện, đối tượng ảnh này được đưa vào mô hình học sâu được huấn luyện từ trước để trích xuất đặc trưng, đặc trưng này là một véc-tơ với định dạng đầy phẩy động 32 bit;

các thông tin kèm theo được xác định khi máy chủ nhận yêu cầu cập nhật cơ sở dữ liệu từ đối tác vận hành, thông tin bắt buộc bao gồm: mã định danh đối tác vận hành; mã định danh người dùng; nhóm tìm kiếm; và nhãn đối tượng;

mã định danh đối tác vận hành là chuỗi ký tự xác định đơn vị (hoặc cá nhân) phát hành dịch vụ điện toán đám mây về nhận diện đối tượng; mã vận hành này do nhà phát triển, cung cấp dịch vụ điện toán đám mây cấp cho đơn vị phát hành dịch vụ;

mã định danh người dùng là chuỗi ký tự xác định người dùng cuối, là người đang trực tiếp sử dụng dịch vụ nhận diện đối tượng theo phương thức điện toán đám mây được phát hành bởi một đối tác vận hành nhất định; mã định danh người dùng do đơn vị phát hành dịch vụ cấp cho người dùng, sau khi đăng ký và kiểm duyệt;

nhóm tìm kiếm đối tượng được mô tả là một chuỗi ký tự thể hiện mảng dữ liệu được đánh dấu khác nhau của cùng một người dùng, sử dụng trong trường hợp người dùng dịch vụ có nhu cầu phân mảnh dữ liệu tại các vị trí lắp đặt camera nhận diện khác nhau; nhóm tìm kiếm do người dùng tự định nghĩa, được lựa chọn khi khởi tạo và cập nhật cơ sở dữ liệu;

nhãn đối tượng là chuỗi ký tự định danh đối tượng, có thể là họ và tên, hoặc một mã ký tự bất kỳ, chỉ cần bảo đảm những mẫu của cùng một đối tượng cùng nhãn là được; nhãn đối tượng do người dùng tự định nghĩa (hoặc có thể qua kênh mã hóa của đơn vị phát hành dịch vụ), được gửi lên khi đăng ký ảnh mới cho đối tượng cần nhận diện.

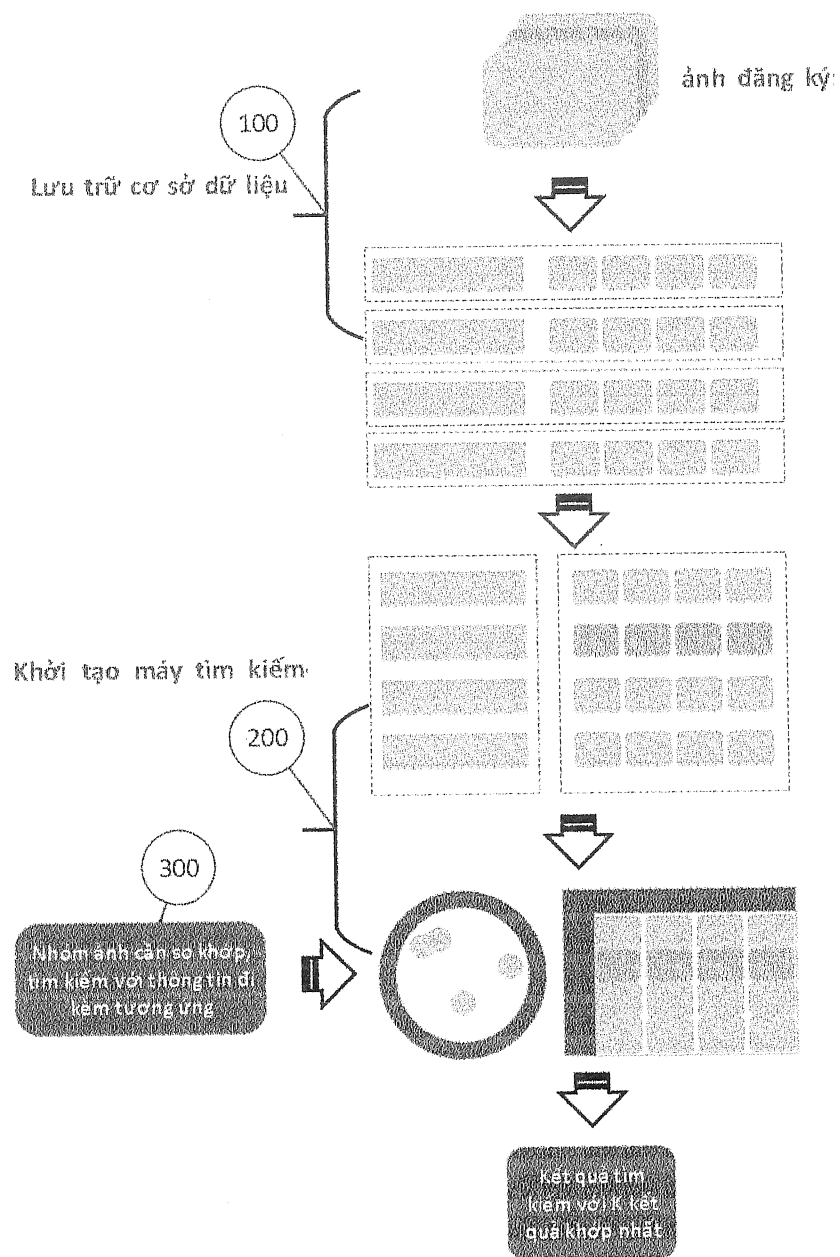
3. Phương pháp so khớp tìm kiếm tốc độ cao với cơ sở dữ liệu hỗn hợp quy mô lớn theo điểm 1 và điểm 2 trong đó, các thông tin trong quá xử lý ảnh đăng ký được lưu trữ như sau:

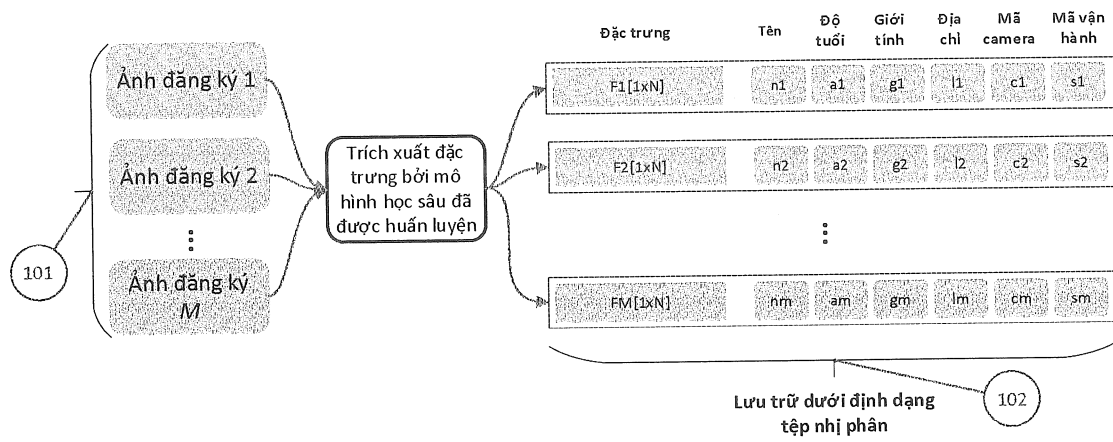
thông tin từ mỗi mẫu ảnh đăng ký được ép kiểu với định dạng là từ điển;

thông tin sau khi trích xuất được lưu trữ dưới dạng tệp nhị phân; các tệp này được lưu như các tệp vật lý khác, trong thư mục được xác định trước trong ổ cứng máy tính.

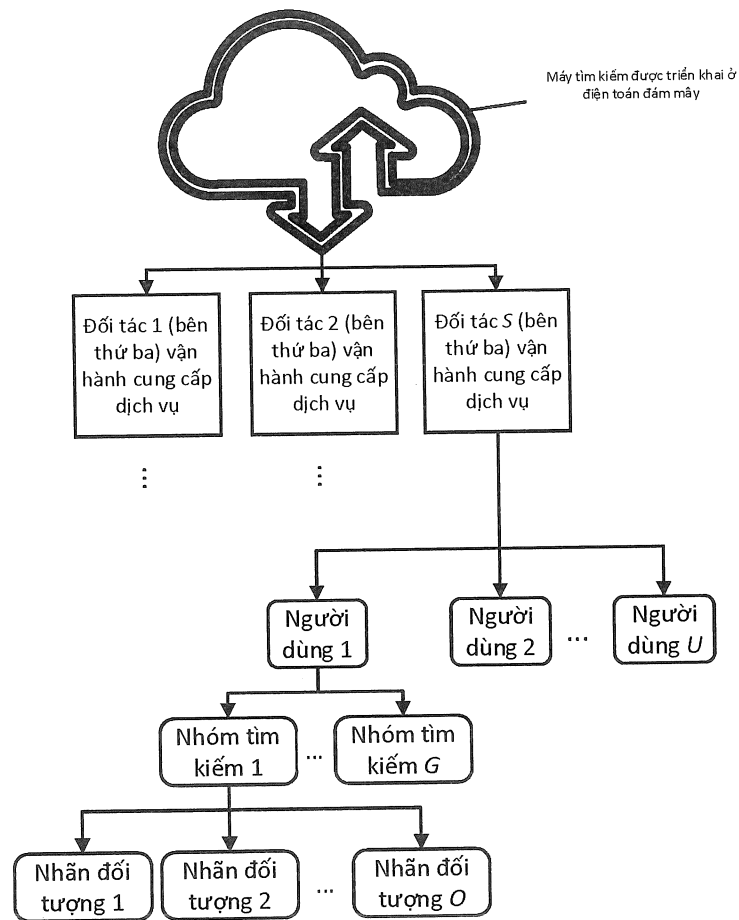
4. Phương pháp so khớp tìm kiếm tốc độ cao với cơ sở dữ liệu hỗn hợp quy mô lớn theo một trong số các điểm từ điểm 1 đến điểm 3, trong quá trình khởi tạo máy tìm kiếm, bảng dữ liệu

thông tin kèm theo được mã hóa sang dạng số nguyên thay vì chuỗi ký tự, trong đó ký tự mã định danh sang định dạng số nguyên, cụ thể: thay vì trực tiếp sử dụng chuỗi ký tự, thì sử dụng vị trí (index) của chuỗi ký tự định danh đó trong toàn bộ mã đối tác vận hành/mã người dùng/nhóm cơ sở dữ liệu, bắt đầu từ 0 và kết thúc là $(\text{tổng số chuỗi ký tự} - 1)$, và được chuyển đổi theo thứ tự từ mã đối tác vận hành $\rightarrow$ mã người dùng $\rightarrow$ nhóm cơ sở dữ liệu.

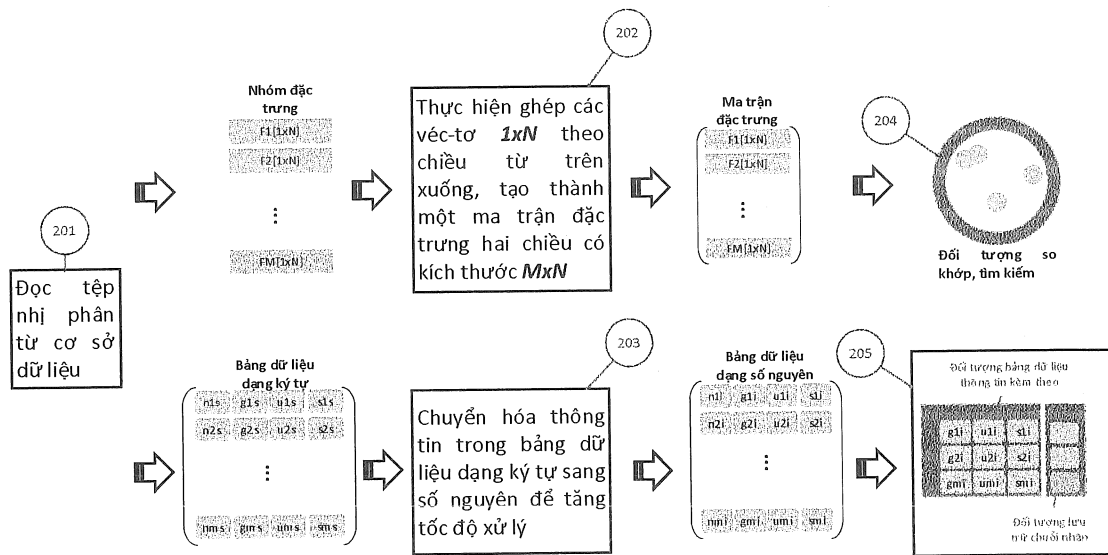
**Hình 1**



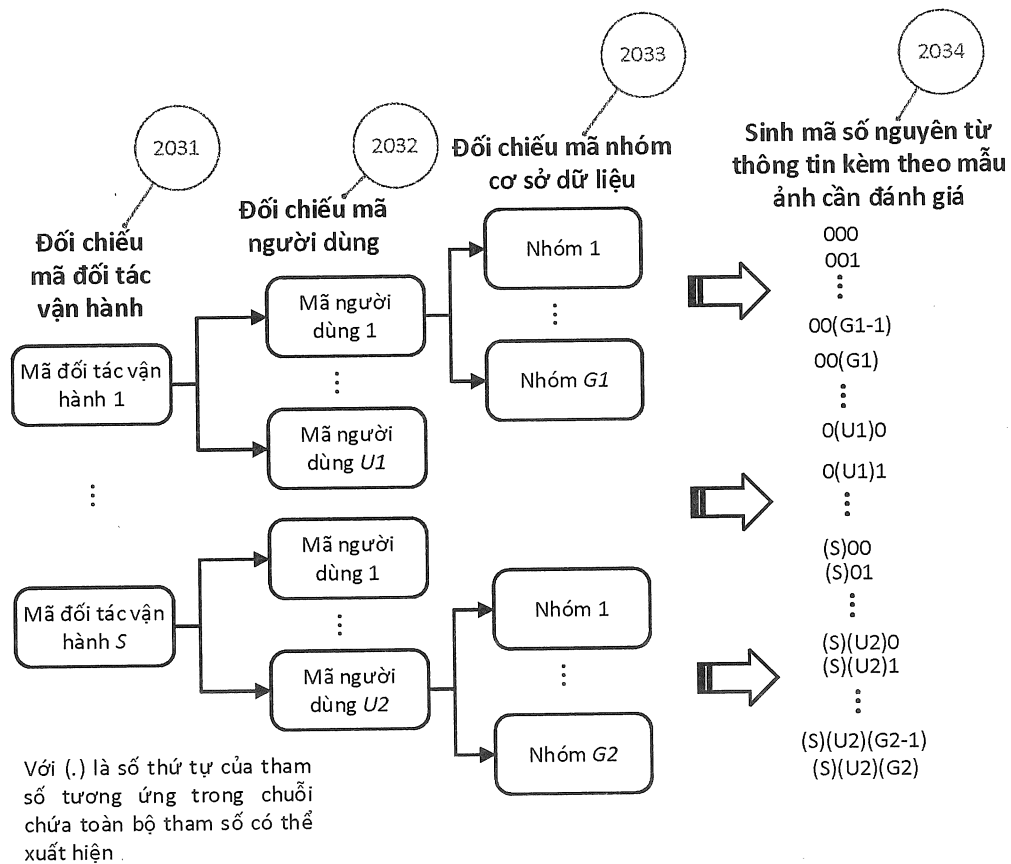
Hình 2



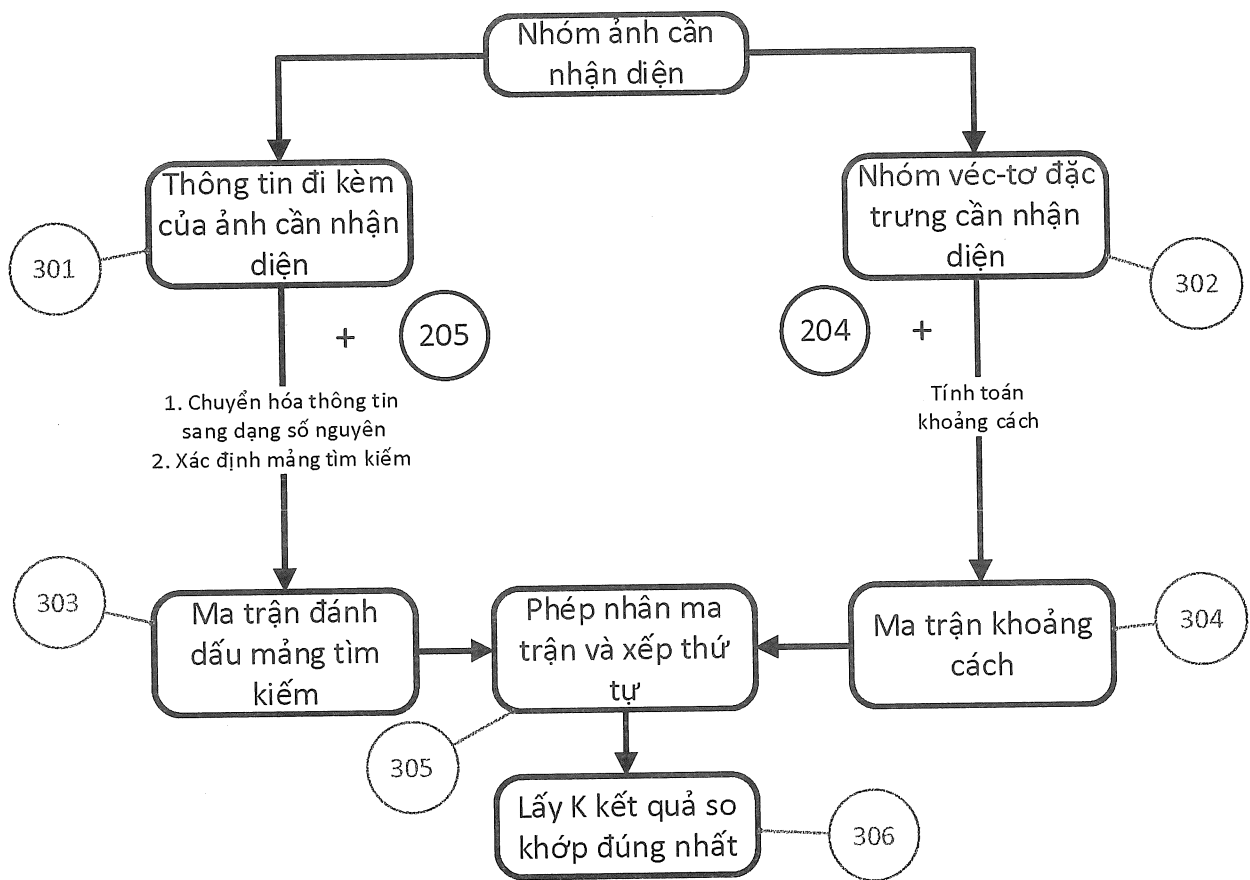
Hình 3



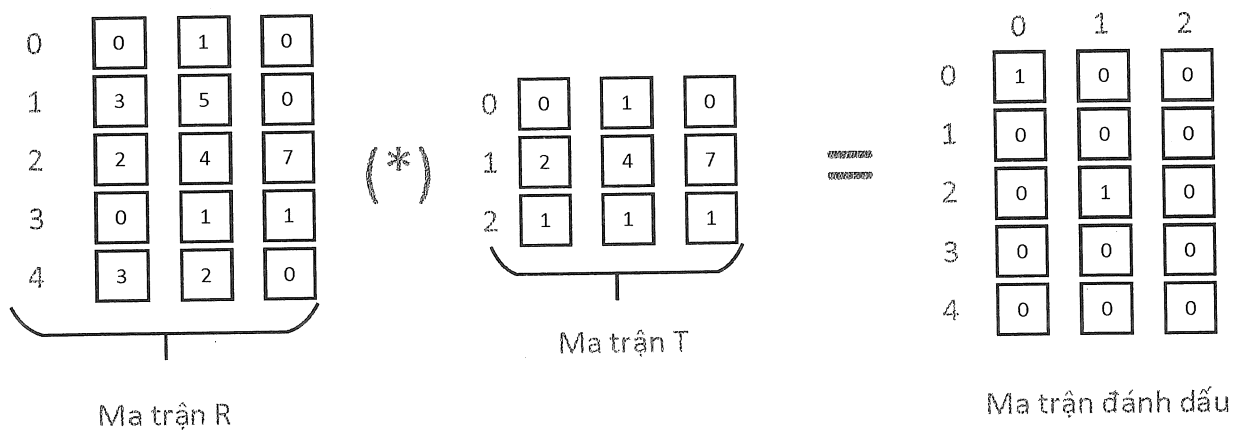
Hình 4



Hình 5



Hình 6



Hình 7