



(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)



1-0042069

(51)^{2020.01} G06K 9/00; G06T 11/00

(13) B

(21) 1-2020-05642

(22) 01/10/2020

(45) 25/12/2024 441

(43) 25/04/2022 409

(73) Công ty cổ phần nghiên cứu và ứng dụng trí tuệ nhân tạo Vinai (VN)

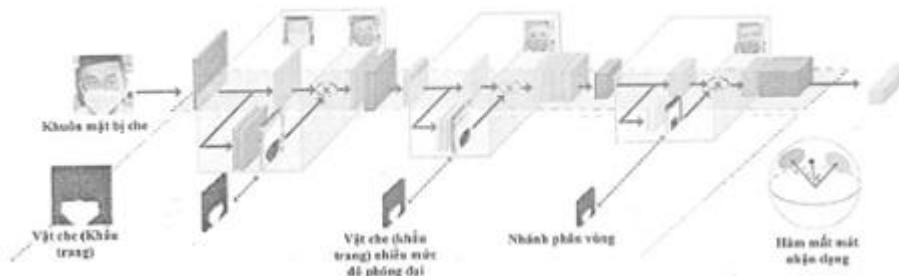
Tòa nhà văn phòng Symphony, đường Chu Huy Mân, KĐT sinh thái Vinhomes
Riverside, phường Phúc Lợi, quận Long Biên, thành phố Hà Nội, Việt Nam

(72) Hung Hai Bui (US); Bùi Đức Toàn (VN); Phạm Hoàng Anh (VN).

(74) Công ty TNHH Tầm nhìn và Liên danh (VISION & ASSOCIATES CO.LTD.)

(54) PHƯƠNG PHÁP NHẬN DẠNG KHUÔN MẶT BỊ CHE

(57) Sáng chế đề cập đến phương pháp nhận dạng khuôn mặt bao gồm bước thu hình ảnh khuôn mặt bị che bao gồm vùng bị che và vùng không bị che; thu đặc trưng hình ảnh từ hình ảnh khuôn mặt bị che; nhập đặc trưng hình ảnh vào mô hình phân vùng được huấn luyện trước để tự động tính toán đặc trưng của vùng bị che; và tinh chỉnh đặc trưng hình ảnh sử dụng đặc trưng được tính toán của vùng bị che, trong đó bước tinh chỉnh bao gồm tập trung vào đặc trưng của vùng không bị che và loại bỏ đặc trưng được tính toán của vùng bị che.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến phương pháp và thiết bị nhận dạng khuôn mặt bị che.

Tình trạng kỹ thuật của sáng chế

Mặc dù đã có những nghiên cứu chuyên sâu về các thuật toán so khớp khuôn mặt dùng máy điện toán, nhưng vẫn rất khó khăn trong việc phát triển các thuật toán nhận dạng khuôn mặt để triển khai mạnh mẽ trong các cài đặt không ràng buộc, đặc biệt là khi có sự che khuất khuôn mặt.

Việc nhận dạng khuôn mặt đóng vai trò quan trọng trong nhiều ứng dụng trên thực tế như các hệ thống camera giám sát, xác minh đối tượng hoặc nhận dạng tội phạm. Nhờ sự thành công của các mạng nơron tích chập sâu (deep convolutional neural network - DCNN), việc nhận dạng khuôn mặt đã đạt được hiệu quả đáng kể trong thực tế.

Tuy nhiên, hầu hết các phương pháp nhận dạng khuôn mặt trên cơ sở DCNN hiện tại xử lý tất cả các điểm ảnh đầu vào, tức là, cả vùng mặt và vùng bị che, như nhau. Do đó, độ chính xác của việc thực hiện giảm đáng kể khi xử lý các khuôn mặt bị che. Điều này là do các khuôn mặt bị che có các phần nhìn thấy không hoàn toàn và các tín hiệu thông tin nhận dạng không đầy đủ. Do đó, rất khó để các hệ thống nhận dạng khuôn mặt đạt được độ chính xác cao.

Bản chất kỹ thuật của sáng chế

Sáng chế đề xuất phương pháp và thiết bị nhận dạng khuôn mặt bị che có cách tiếp cận đơn giản để đáp ứng một số mục tiêu nhận dạng đối tượng mạnh mẽ trong các tình huống thực tế và giải quyết vấn đề nhận dạng khuôn mặt bị che.

Sáng chế cũng đề xuất phương pháp và thiết bị nhận dạng khuôn mặt bị che cung cấp cơ chế tự động để phân biệt vùng mặt và vùng bị che một cách riêng rẽ không giống các phương pháp hiện có trong đó tất cả các điểm ảnh đầu vào, tức là, cả vùng mặt và vùng bị che, trong khuôn mặt được căn chỉnh của đối tượng đều được xem xét.

Sáng chế cũng đề xuất phương pháp và thiết bị nhận dạng khuôn mặt bị che

mà thực hiện tích chập cổng được dẫn hướng (guided gating convolution) để buộc mạng học các vùng mặt và loại bỏ các vùng bị che.

Sáng chế cũng đề xuất phương pháp và thiết bị nhận dạng khuôn mặt bị che dựa vào kiến trúc nhận dạng khuôn mặt không bị che.

Sáng chế cũng đề xuất phương pháp và thiết bị nhận dạng khuôn mặt bị che dùng mạng học sâu mới.

Sáng chế cũng đề xuất phương pháp và thiết bị nhận dạng khuôn mặt bị che trong đó kiến trúc mạng mới dựa vào tích chập cổng được dẫn hướng hoặc cảm ứng được đề xuất để nhận dạng khuôn mặt bị che.

Các mục tiêu cần đạt được bằng các phương án của sáng chế không bị giới hạn ở đó và sáng chế cũng có thể bao gồm các mục tiêu hoặc các hiệu quả có thể được suy ra từ các giải pháp hoặc các phương án được mô tả dưới đây.

Theo một khía cạnh, sáng chế đề xuất phương pháp nhận dạng khuôn mặt bao gồm các bước thu hình ảnh khuôn mặt bị che bao gồm vùng bị che và vùng không bị che; thu đặc trưng hình ảnh từ hình ảnh khuôn mặt bị che; nhập đặc trưng hình ảnh vào mô hình phân vùng được huấn luyện trước để tự động tính toán đặc trưng của vùng bị che; và tinh chỉnh đặc trưng hình ảnh sử dụng đặc trưng được tính toán của vùng bị che, trong đó bước tinh chỉnh bao gồm bước tập trung vào đặc trưng của vùng không bị che và loại bỏ đặc trưng được tính toán của vùng bị che.

Mô hình phân vùng có thể được huấn luyện bằng cách sử dụng nhãn vật che và đặc trưng khuôn mặt bị che làm các đầu vào và học hệ số của hàm softmax (softmax score) để phân loại một vùng là vùng bị che nếu hệ số của hàm softmax chỉ báo trị số cao và phân loại một vùng là vùng không bị che nếu hệ số của hàm softmax chỉ báo trị số thấp.

Phương pháp nhận dạng khuôn mặt có thể còn bao gồm bước tối thiểu hóa hàm mất mát entropi chéo (cross entropy loss function) giữa đặc trưng được tính toán của vùng bị che và nhãn vật che.

Hình ảnh khuôn mặt bị che có thể được chọn từ nhóm bao gồm hình ảnh khuôn mặt bị che và hình ảnh khuôn mặt không bị che.

Phương pháp nhận dạng khuôn mặt có thể dựa vào mạng nơron tích chập sâu

và phương pháp nhận dạng khuôn mặt có thể được thực hiện trên nhiều lớp của mạng nơron tích chập sâu.

Đặc trưng hình ảnh có thể được lấy mẫu xuống (down-sampled) từ cấp thô nhất đến cấp nét nhất trên nhiều lớp.

Nhãn vật che có thể được lấy mẫu xuống.

Theo một khía cạnh khác, sáng chế đề xuất vật ghi đọc được bằng máy tính được tạo cấu hình để lưu trữ chương trình để thực thi phương pháp nhận dạng khuôn mặt. Phương pháp nhận dạng khuôn mặt bao gồm bước thu hình ảnh khuôn mặt bị che bao gồm vùng bị che và vùng không bị che; thu đặc trưng hình ảnh từ hình ảnh khuôn mặt bị che; nhập đặc trưng hình ảnh vào mô hình phân vùng được huấn luyện trước để tự động tính toán đặc trưng của vùng bị che; và tinh chỉnh đặc trưng hình ảnh sử dụng đặc trưng được tính toán của vùng bị che, trong đó bước tinh chỉnh bao gồm bước tập trung vào đặc trưng của vùng không bị che và loại bỏ đặc trưng được tính toán của vùng bị che.

Mô tả vắn tắt các hình vẽ

Các đối tượng, các dấu hiệu và các lợi thế ở trên và các đối tượng, các dấu hiệu và các lợi thế khác của sáng chế sẽ trở nên rõ ràng hơn đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này bằng cách mô tả chi tiết các phương án làm ví dụ dựa vào các hình vẽ kèm theo, trong đó:

Fig.1 là lưu đồ minh họa phương pháp nhận dạng khuôn mặt theo một phương án làm ví dụ của sáng chế;

Fig.2 là sơ đồ khối của thiết bị nhận dạng khuôn mặt theo một phương án làm ví dụ của sáng chế;

Fig.3 là sơ đồ thể hiện sự so sánh giữa tích chập vanilla (vanilla convolution), tích chập cổng (gate convolution) và tích chập cổng được dẫn hướng (guided gating convolution);

Fig.4 là sơ đồ thể hiện kiến trúc để nhận dạng khuôn mặt bị che;

Fig.5 thể hiện tập dữ liệu khuôn mặt bị che tổng hợp;

Fig.6a đến Fig.6c thể hiện hình ảnh đã cho, vùng bị che và vùng không bị che

được đóng góp như nhau trong tích chập vanilla, và các phần đóng góp được điều chỉnh thích ứng trong tích chập công được dẫn hướng, một cách tương ứng; và

Fig.7a đến Fig.7c lần lượt thể hiện sự phân bố các đặc trưng được học trong tích chập vanilla, tích chập công và tích chập công được dẫn hướng (các hình ảnh không bị che ở trên cùng và các hình ảnh bị che ở dưới cùng).

Mô tả chi tiết sáng chế

Mặc dù có thể thực hiện nhiều biến thể và các phương án khác nhau theo sáng chế, các phương án làm ví dụ sẽ được thể hiện trên các hình vẽ kèm theo và được mô tả. Tuy nhiên, cần hiểu rằng sáng chế không bị giới hạn ở các phương án làm ví dụ và bao gồm tất cả các thay đổi, dạng tương đương hoặc thay thế nằm trong phạm vi kỹ thuật của sáng chế.

Các thuật ngữ bao gồm các số thứ tự, như thứ hai và thứ nhất, có thể được sử dụng để mô tả nhiều thành phần, nhưng các thành phần không bị giới hạn ở các thuật ngữ này. Các thuật ngữ chỉ được sử dụng để phân biệt một thành phần này với một thành phần khác. Ví dụ, không vượt ra khỏi phạm vi của sáng chế, thành phần thứ hai có thể được gọi là thành phần thứ nhất, và tương tự, thành phần thứ nhất có thể được gọi là thành phần thứ hai. Thuật ngữ “và/hoặc” bao gồm kết hợp bất kỳ của nhiều mục được liệt kê liên quan hoặc bất kỳ một trong số nhiều mục được liệt kê liên quan.

Khi mô tả là một thành phần “được nối” hoặc “được liên kết” với một thành phần khác, cần hiểu rằng thành phần này có thể được nối hoặc được liên kết trực tiếp với thành phần kia nhưng một thành phần khác có thể ở giữa chúng. Mặt khác, khi mô tả là một thành phần “được nối trực tiếp” hoặc “được liên kết trực tiếp” với một thành phần khác, cần hiểu rằng không có thành phần nào khác ở giữa chúng.

Các thuật ngữ được sử dụng ở đây chỉ để mô tả các phương án làm ví dụ và không nhằm mục đích giới hạn sáng chế. Các dạng số ít bao gồm cả các dạng số nhiều, trừ khi ngữ cảnh chỉ khác đi một cách rõ ràng. Xuyên suốt bản mô tả, cần hiểu rằng các thuật ngữ “bao gồm”, “có”, v.v. được sử dụng ở đây để chỉ ra sự có mặt của các dấu hiệu, các số, các bước, các hoạt động, các thành phần, các phần đã nêu hoặc các kết hợp của chúng nhưng không loại trừ sự có mặt hoặc bổ sung một hoặc nhiều dấu hiệu, các số, các bước, các hoạt động, các thành phần, các phần khác hoặc các kết hợp của chúng.

Trừ khi được định nghĩa khác, các thuật ngữ được sử dụng ở đây bao gồm các thuật ngữ khoa học hoặc kỹ thuật có thể có nghĩa giống như nghĩa thường được hiểu bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật này. Các thuật ngữ như các thuật ngữ được định nghĩa trong các từ điển thường được sử dụng cần được hiểu là có nghĩa thống nhất với nghĩa của chúng trong ngữ cảnh của lĩnh vực kỹ thuật liên quan và không nên được hiểu theo nghĩa lý tưởng hóa hoặc quá hình thức, trừ khi được định nghĩa rõ ràng ở đây.

Sau đây, các phương án của sáng chế sẽ được mô tả chi tiết dựa vào các hình vẽ kèm theo. Trên toàn bộ các hình vẽ, các số chỉ dẫn giống nhau sẽ được dùng cho các thành phần giống hoặc tương tự nhau và phần mô tả lặp lại của chúng sẽ được lược bỏ.

Fig.1 là lưu đồ minh họa phương pháp nhận dạng khuôn mặt theo một phương án làm ví dụ của sáng chế và Fig.2 là sơ đồ khối của thiết bị nhận dạng khuôn mặt theo một phương án làm ví dụ của sáng chế.

Theo Fig.1, phương pháp nhận dạng khuôn mặt theo phương án làm ví dụ của sáng chế bao gồm hoạt động S1100 để thu hình ảnh khuôn mặt bị che, hoạt động S1200 để thu đặc trưng hình ảnh từ hình ảnh khuôn mặt bị che, hoạt động S1300 để nhập đặc trưng hình ảnh vào mô hình phân vùng được huấn luyện trước để tự động tính toán đặc trưng của vùng bị che, hoạt động S1400 để tinh chỉnh đặc trưng hình ảnh sử dụng đặc trưng được tính toán của vùng bị che bằng cách tập trung vào đặc trưng của vùng không bị che và loại bỏ đặc trưng được tính toán của vùng bị che.

Hơn nữa, theo một phương án, mô hình phân vùng được huấn luyện bằng cách sử dụng nhãn vật che và đặc trưng khuôn mặt bị che làm các đầu vào và học hệ số của hàm softmax để phân loại một vùng là vùng bị che nếu hệ số của hàm softmax chỉ báo trị số cao và phân loại một vùng là vùng không bị che nếu hệ số của hàm softmax chỉ báo trị số thấp.

Theo phương án, phương pháp nhận dạng khuôn mặt có thể còn bao gồm bước tối thiểu hóa hàm mất mát entropi chéo giữa đặc trưng được tính toán của vùng bị che và nhãn vật che.

Hơn nữa, hình ảnh khuôn mặt bị che có thể được chọn từ nhóm bao gồm hình ảnh khuôn mặt bị che và hình ảnh khuôn mặt không bị che.

Phương pháp nhận dạng khuôn mặt theo phương án làm ví dụ của sáng chế có thể dựa vào mạng nơron tích chập sâu và phương pháp nhận dạng khuôn mặt có thể được thực hiện trên nhiều lớp của mạng nơron tích chập sâu.

Đặc trưng hình ảnh có thể được lấy mẫu xuống từ cấp thô nhất đến cấp nét nhất trên nhiều lớp.

Nhãn vật che cũng có thể được lấy mẫu xuống.

Theo phương án khác, vật ghi đọc được bằng máy tính được đề xuất có thể được tạo cấu hình để lưu trữ chương trình để thực thi phương pháp nhận dạng khuôn mặt như được mô tả ở trên.

Theo Fig.2, thiết bị nhận dạng khuôn mặt 100 theo phương án làm ví dụ của sáng chế có thể bao gồm bộ phận đầu vào 110, bộ thu 120, bộ tính toán 130 và bộ tinh chỉnh 140.

Bộ phận đầu vào 110 có thể nhập hình ảnh khuôn mặt bị che. Hình ảnh khuôn mặt bị che có thể bao gồm vùng bị che và vùng không bị che. Sau đó, bộ thu 120 có thể thu đặc trưng hình ảnh từ hình ảnh khuôn mặt bị che. Bộ tính toán 130 có thể tự động tính toán đặc trưng của vùng bị che bằng cách nhập đặc trưng hình ảnh vào mô hình phân vùng được huấn luyện trước. Bộ tính toán 130 thực thi các nội dung của phân vùng (được gọi là nút, mạng, mô hình phân vùng, v.v.).

Bộ tính toán 130 và bộ tinh chỉnh 140 thực thi nội dung của tích chập và kiến trúc cũng như nội dung của nhánh (hoặc nút) phân vùng và nhận dạng huấn luyện đồng thời, sẽ được mô tả dưới đây.

Các hoạt động chi tiết được sử dụng trong phương pháp và thiết bị nhận dạng khuôn mặt theo các phương án làm ví dụ của sáng chế sẽ được mô tả dưới đây.

Fig.3 là sơ đồ thể hiện sự so sánh giữa tích chập vanilla, tích chập công và tích chập công được dẫn hướng. Fig.4 là sơ đồ thể hiện kiến trúc để nhận dạng khuôn mặt bị che. Fig.5 thể hiện tập dữ liệu khuôn mặt bị che tổng hợp. Các hình vẽ từ Fig.6a đến Fig.6c thể hiện hình ảnh đã cho, vùng bị che và vùng không bị che được đóng góp như nhau trong tích chập vanilla, và các phần đóng góp được điều chỉnh thích ứng trong tích chập công được dẫn hướng, một cách tương ứng. Các hình vẽ từ Fig.7a đến Fig.7c thể hiện sự phân bố các đặc trưng được học lần lượt trong tích chập vanilla, tích chập

cộng và tích chập cộng được dẫn hướng (các hình ảnh không bị che ở trên cùng và các hình ảnh bị che ở dưới cùng).

Các phần mô tả sẽ được trình bày dưới đây dựa vào các hình vẽ từ Fig.3 đến Fig.7.

Tích chập vanilla

Như thường được sử dụng trong hầu hết các cấu trúc học sâu để trích xuất đặc trưng, tích chập vanilla có thể được xác định theo phương trình 1.

[Phương trình 1]

$$x_{\ell+1} = \mathcal{F}(x_{\ell}, W_{\ell}) = (x_{\ell} \circledast W_{\ell})_{(i,j)}$$

x_{ℓ} và $x_{\ell+1}$ lần lượt biểu thị đặc trưng đầu vào và đặc trưng đầu ra ở lớp ℓ^{th} . $\mathcal{F}(x_{\ell}, W_{\ell})$ là hàm chuyển đổi phi tuyến với các tham số W_{ℓ} , và $\circledast$ biểu diễn phép toán tích chập.

Phương trình 1 thể hiện rằng các bộ lọc giống nhau được áp dụng cho tất cả các vị trí trong không gian (i, j) để tạo ra các đầu ra từ các lớp tích chập vanilla. Điều này phù hợp cho các tác vụ, như phân loại hình ảnh và phát hiện đối tượng, và tất cả các điểm ảnh của hình ảnh đầu vào đều hợp lệ để trích xuất các đặc trưng cục bộ theo kiểu cửa sổ trượt.

Tuy nhiên, trong phương pháp nhận dạng khuôn mặt bị che, các đầu vào có thể bao gồm cả các vùng với các điểm ảnh hoặc các đặc trưng hợp lệ ngoài các khoảng trống và các điểm ảnh hoặc các đặc trưng không hợp lệ (trong các lớp nông) hoặc các điểm ảnh được tổng hợp (trong các lớp sâu) trong các vùng bị che. Điều này về mặt toán học gây ra sự không rõ ràng và dẫn đến các nhiễu trực quan như sự khác biệt màu sắc, nhòe và các đáp ứng mép rõ ràng khi kiểm thử.

Tích chập cộng

Để đảm bảo tích chập chỉ phụ thuộc vào các điểm ảnh hợp lệ, tích chập cộng có thể tối thiểu hóa ảnh hưởng của các vùng bị che trong quy trình học.

Như được thể hiện trên Fig.3, cấu trúc tích chập cộng bao gồm hai khối tích chập chính, tức là, khối trích xuất đặc trưng và khối che mềm (soft masking block).

Trong khi khối trước (khối trích xuất đặc trưng) dùng để trích xuất các đặc trưng phân biệt cho các phép toán sau, thì khối sau (khối che mềm) làm giảm sự đóng góp của các đặc trưng có trong các vùng bị che và dẫn hướng quy trình học đến các vùng khuôn mặt tin cậy hơn.

Tích chập công có hai ưu điểm. Thứ nhất, các phần đóng góp của các vùng bị che được học thích ứng từ dữ liệu qua mỗi cấp của quy trình học. Thứ hai, thay vì ngay lập tức bị bỏ qua từ đầu (tức là, ngưỡng cứng), tất cả thông tin hữu ích (ví dụ, hình dạng khuôn mặt) từ các vùng bị che vẫn có thể đóng góp vào khả năng phân biệt của đặc trưng cuối cùng.

Cụ thể, tích chập công học hệ số của hàm softmax của đặc trưng công x_{ℓ}^g từ đặc trưng đầu vào x_{ℓ} . Sau đó, hàm sigmoid được sử dụng để tinh chỉnh đặc trưng hình ảnh x_{ℓ} như được thể hiện trong phương trình 2.

[Phương trình 2]

$$x_{\ell+1}^g = \mathcal{F}(x_{\ell}, W_{\ell}^g)$$

$$x_{\ell+1} = x_{\ell+1} \times \sigma(x_{\ell+1}^g)$$

Khối tích chập công cung cấp định hướng hiệu quả cho quy trình học khi xử lý các vùng bị che trong các hình ảnh khuôn mặt. Tuy nhiên, tích chập công yêu cầu Mask Prior (tiền vật che hoặc tiền khẩu trang) làm đầu vào và không có thông tin vật che (khẩu trang) giám sát ở lớp sâu hơn.

Do độ phân giải lấy mẫu xuống ở lớp pooling, nên khó để học các vùng bị che mà không có thông tin vật che (khẩu trang) giám sát.

Tích chập công được dẫn hướng

Phương pháp và thiết bị nhận dạng khuôn mặt bị che theo các phương án làm ví dụ của sáng chế có thể thực hiện khối hoặc tích chập công được dẫn hướng để bổ sung cho các phép tích chập khác được mô tả ở trên.

Trong phần mô tả sau đây, thuật ngữ “GiGaNet” có thể được sử dụng đồng nghĩa với tích chập công được dẫn hướng mà phương pháp và thiết bị nhận dạng khuôn mặt bị che theo các phương án làm ví dụ của sáng chế thực hiện.

Thứ nhất, tính khả dụng của tập các khuôn mặt bị che $M = \{M_i\}_{i=1}^m$ được giả định thông qua các nhãn nhận dạng $\mathcal{Y} = \{Y_i\}_{i=1}^m$. Ở đây, m là số lượng các mẫu, các điểm ảnh có giá trị 1 nằm trong vùng bị che và các điểm ảnh có giá trị 0 nằm trong vùng không bị che.

Hình ảnh khuôn mặt bị che tổng hợp $\mathcal{I}^M = \{I_j^M\}_{j=1}^n$ thu được bằng cách phát hiện vùng miệng và cằm trong khuôn mặt đã cho.

Nói cách khác, hình ảnh khuôn mặt bị che tổng hợp được tạo ra từ hình ảnh bị che M_i và hình ảnh không bị che I_j như được thể hiện trong phương trình 3.

[Phương trình 3]

$$I_j^M(I_j, M_i) = I_j \times (1 - Y_i) + M_i \times (Y_i)$$

Phần mô tả ở trên được sử dụng để học trong mạng hoặc mô hình phân vùng.

Mạng hoặc mô hình phân vùng có thể sử dụng các nhãn vật che và các đặc trưng khuôn mặt bị che làm các đầu vào và học hệ số của hàm softmax có trị số cao làm vùng bị che và trị số thấp làm vùng không bị che trong quy trình huấn luyện. Do đó, mô hình phân vùng có thể phân loại một vùng là vùng bị che nếu hệ số của hàm softmax chỉ báo trị số cao và phân loại một vùng là vùng không bị che nếu hệ số của hàm softmax chỉ báo trị số thấp.

Nói cách khác, mạng hoặc mô hình phân vùng được huấn luyện làm mô hình $S(I^M, \theta)$ với tham số θ lấy hình ảnh khuôn mặt bị che I^M làm đầu vào và dự đoán hoặc tính toán đặc trưng của vùng bị che $\hat{x}_{\ell+1}^m = S(I^M, \theta)$ làm đầu ra. Việc này có thể được thực hiện bằng cách tối thiểu hóa hàm mất mát entropi chéo (cross entropy - CE) giữa đặc trưng vùng bị che được tính toán $\hat{x}_{\ell+1}^m$ và nhãn vật che Y .

[Phương trình 4]

$$\mathcal{L}_{CE}(\theta) = \sum_{i=1}^C Y_i \log(S(I^M, \theta)) = \sum_{i=1}^C Y_i \log(\hat{x}_{\ell+1}^m)$$

trong đó C là 2 và biểu thị vùng bị che và vùng không bị che.

Đặc trưng vùng bị che được tính toán $\hat{x}_{\ell+1}^m$ khiến mạng hoặc mô hình tập

trung vào các vùng không bị che và loại bỏ các đặc trưng khuôn mặt bị che khuất trên các vùng bị che. Đặc trưng công được dẫn hướng được xác định bởi phương trình 5.

[Phương trình 5]

$$x_{\ell+1} = x_{\ell+1} \times \hat{x}_{\ell+1}^m$$

Ngoài hàm công của phương trình 2, đặc trưng công được dẫn hướng sử dụng hàm softmax của các vùng bị che được tính toán để tinh chỉnh các đặc trưng của khuôn mặt (hoặc các đặc trưng hình ảnh) phân biệt như được thể hiện trong phương trình 5.

Vùng bị che được xác định hoặc được suy ra bằng cách tối ưu hóa mạng phân vùng S . Do đó, nhãn vật che không cần thiết trong quy trình suy luận (inference).

Do đó, tích chập công được dẫn hướng không yêu cầu vùng bị che được dán nhãn trong quy trình nhận dạng khuôn mặt và do đó có thể được áp dụng hiệu quả cho tất cả các lớp trong các mạng nơron tích chập sâu (DCNN). Fig.3 minh họa sự so sánh giữa tích chập vanilla, tích chập công và tích chập công được dẫn hướng. Do tích chập công được dẫn hướng theo sáng chế không yêu cầu các vùng bị che được dán nhãn trong quy trình nhận dạng, nên nó có thể được áp dụng hiệu quả cho các lớp bất kỳ trong các mạng DCNN.

Kiến trúc mạng sâu dùng để nhận dạng khuôn mặt bị che

Trên cơ sở tích chập công được dẫn hướng nêu trên, phương pháp và thiết bị nhận dạng khuôn mặt bị che theo các phương án làm ví dụ của sáng chế hoạt động như là khung trên cơ sở học sâu.

Do đó, phương pháp và thiết bị nhận dạng khuôn mặt bị che theo các phương án làm ví dụ của sáng chế có thể được triển khai linh hoạt trên mạng nhận dạng cơ sở như ResNet-100 hoặc Mobilefacenet (MFN).

Nhận dạng khuôn mặt bị che bao gồm hai nhánh, tức là, nhánh nhận dạng và nhánh phân vùng.

Nhánh hoặc nút nhận dạng

Nhánh hoặc nút nhận dạng bao gồm n khối lấy mẫu xuống (ví dụ, $n=4$ đối với MFN). Chồng của tất cả các khối cho cùng độ phân giải được coi là một cấp.

Đặc trưng ở lớp ℓ^{th} của cấp s được chỉ báo bởi $x_{\ell,s}$. $s=1$ tương ứng với cấp

thô nhất, $s=n$ tương ứng với cấp nét nhất.

Đầu tiên, $x_{\ell,s}$ được giả định là đặc trưng ở lớp ℓ^{th} của cấp s . Đối với mức phóng đại nhất định hoặc cụ thể s , đặc trưng hình ảnh đầu ra của lớp ℓ^{th} thu được như trong phương trình 6.

[Phương trình 6]

$$x_{\ell+1,s} = \mathcal{F}(x_{\ell}, W_{\ell}) \times \mathcal{S}((\mathcal{F}(x_{\ell}, W_{\ell}^m), Y^s)) = x_{\ell+1,s} \times \hat{x}_{\ell+1,s}^m$$

trong đó Y^s có thể là nhãn của vùng bị che ở độ phóng đại s .

Để tận dụng thông tin khuôn mặt đầy đủ, công quyết định được xác định như trong phương trình 7. Chức năng này có thể được thực hiện bởi bộ tinh chỉnh 140.

[Phương trình 7]

$$\hat{x}_{\ell+1,s}^d = \begin{cases} \hat{x}_{\ell+1,s}^m, & \text{nếu } \sum Y \geq 1 \\ 0, & \text{nếu không thì} \end{cases}$$

Nói cách khác, các đặc trưng khuôn mặt (hình ảnh) đầy đủ và sự tinh chỉnh của nó là giống nhau đối với các khuôn mặt không bị che. Ở đây, giả sử việc phân loại vật che là tốt đủ cho công quyết định trong quy trình suy luận.

Cuối cùng, đặc trưng hình ảnh hoặc đặc trưng khuôn mặt được tinh chỉnh từ phương trình 6 được cập nhật như trong phương trình 8.

[Phương trình 8]

$$x_{\ell+1,s} = x_{\ell+1,s} \times \hat{x}_{\ell+1,s}^d$$

Nhánh hoặc nút phân vùng

Nhánh phân vùng nhằm tính toán đặc trưng của vùng bị che $\hat{x}_{\ell+1}^m$ để tinh chỉnh đặc trưng hình ảnh $x_{\ell,s}$ ở cấp s . Đối với mức phóng đại cụ thể s , nhãn vật che Y được lấy mẫu xuống thành Y^s , và phân vùng bị che được thực hiện đối với độ phóng đại cấp độ điểm ảnh.

Đặc trưng hình ảnh $x_{\ell,s}$ ở lớp ℓ và độ phóng đại s được đưa vào tích chập với trọng số W_{ℓ}^m để thu đặc trưng khuôn mặt bị che $\hat{x}_{\ell+1,s}^m = \mathcal{F}(x_{\ell}, W_{\ell}^m)$. Đặc trưng khuôn mặt bị che được đưa vào tích chập 1×1 để phân loại các vùng thành hai loại, tức là, các vùng bị che và các vùng không bị che.

Hàm mất mát entropi chéo được xác định như trong phương trình 4 để tối thiểu hóa hai loại là các vùng được phân loại và các vùng bị che được dán nhãn.

Theo phương án làm ví dụ của sáng chế, việc học nêu trên có thể đề cập đến việc tính toán vật che nhiều cấp trong đó bản đồ vật che đầu ra được sử dụng để cho phép mạng học đặc trưng hợp lệ ở mỗi độ phóng đại. Nói cách khác, mô hình phân vùng có thể được huấn luyện trước, và vùng và đặc trưng khuôn mặt bị che được tính toán thu được bằng cách cung cấp đặc trưng hình ảnh qua mô hình phân vùng. Do đó, không có nhãn nào được yêu cầu cho vùng bị che.

Nhánh phân vùng và nhánh nhận dạng huấn luyện đồng thời

Nhánh nhận dạng và nhánh phân vùng được huấn luyện đồng thời để tìm giải pháp tối ưu. Nói cách khác, các đặc trưng hình ảnh ở các cấp khác nhau được tinh chỉnh bởi các hệ số của hàm softmax từ các đặc trưng thô (các lớp nông) đến các đặc trưng nét hơn (các lớp sâu). Cuối cùng, hàm mục tiêu của mạng trong phương pháp và thiết bị nhận dạng khuôn mặt bị che theo các phương án làm ví dụ của sáng chế được xác định như trong phương trình 9.

[Phương trình 9]

$$\mathcal{L}_{total}(\theta) = \mathcal{L}_{recog}(\theta) + \sum_{i=1}^S \alpha_i \mathcal{L}_{CE}^s(\theta)$$

trong đó α_i biểu thị trọng số để kiểm soát tầm quan trọng của tích chập công được dẫn hướng của các lớp khác nhau. $\mathcal{L}_{recog}(\theta)$ biểu thị mất mát arcface (arcface loss).

Ví dụ thực hiện sáng chế

Các tập dữ liệu

Phương pháp và thiết bị nhận dạng khuôn mặt theo các phương án làm ví dụ của sáng chế được đánh giá toàn diện trên cả khuôn mặt bị che thực và khuôn mặt bị che tổng hợp theo nhiều tiêu chuẩn phổ biến bao gồm các khuôn mặt được dán nhãn trong tự nhiên (labeled faces in the wild - LFW), những người nổi tiếng theo khuôn mặt-tiểu sử trong tự nhiên (celebrities in frontal-profile in the wild - CFP-FP), AgeDB và MegaFace.

Phương pháp và thiết bị nhận dạng khuôn mặt bị che được đề xuất được đánh giá bổ sung trên các khuôn mặt bị che thực của tập dữ liệu khuôn mặt bị che nội bộ mới được thu thập.

Các tập dữ liệu khuôn mặt bị che tổng hợp: Để mở rộng khối lượng và tính đa dạng của việc nhận dạng khuôn mặt bị che trên các tập dữ liệu khuôn mặt quy mô lớn công khai, các vật che (khẩu trang) được đeo trên tất cả các khuôn mặt trên cơ sở 68 điểm quan trọng (landmark points) được trích xuất. Fig.5 thể hiện các hình ảnh khuôn mặt không bị che (hàng thứ nhất) và các hình ảnh khuôn mặt bị che tổng hợp (hàng thứ hai).

Tập dữ liệu khuôn mặt bị che nội bộ: Tập dữ liệu khuôn mặt bị che nội bộ bao gồm 4133 hình ảnh của 600 người đeo khẩu trang, và 3533 hình ảnh của cùng đối tượng mà không đeo khẩu trang. Các biến thể khuôn mặt khác nhau về tư thế, ánh sáng và biểu cảm được bao gồm trong tập dữ liệu khuôn mặt bị che nội bộ.

MegaFace: Tập tham chiếu MegaFace bao gồm một triệu hình ảnh của 690.000 đối tượng, và tập kiểm tra bao gồm 100.000 bức ảnh của 530 cá nhân đơn nhất trong FaceScrub.

Chi tiết phương án thực hiện

Các khuôn mặt đầu vào được cắt và được thay đổi kích thước thành 112×112 điểm ảnh với năm điểm quan trọng. Mobilefacenet và ResNet-100 được dùng trong mạng nhúng. Các thử nghiệm thực hiện MxNet, khung phần mềm học sâu nguồn mở, được sử dụng để huấn luyện và triển khai các mạng nơ-ron sâu. Các mô hình được huấn luyện trên bốn bộ xử lý đồ họa NVIDIA DGX 48GB (GPU). Các mô hình được huấn luyện với thuật toán giảm gradient ngẫu nhiên (stochastic gradient descent - SGD) với sự suy giảm trọng số (weight decay) là $5e^{-4}$ và động lượng (momentum) là 0,9. Tốc độ học bắt đầu từ 0,1 epoch, trở thành 10, 20 và 30 epoch, và kết thúc ở 35 epoch. Đối với lượng mất mát arcface, cài đặt chung được áp dụng, cụ thể là độ phóng đại (scale) được đặt là 64 và lề (margin) được đặt là 0,5. Sau khi mạng được huấn luyện với hàm mất mát softmax đối với 15 epoch, nhánh hoặc nút phân vùng được cố định, và nhánh nhận dạng được huấn luyện với hàm mất mát arcface đối với 20 epoch.

Các nghiên cứu lược bỏ (Ablation Studies)

Tích chập cổng được dẫn hướng so với tích chập vanilla: Trong nhận dạng khuôn mặt bị che, các đầu vào bao gồm cả các vùng với các điểm ảnh hợp lệ và các điểm ảnh không hợp lệ (tức là, vùng không bị che và vùng bị che, tương ứng).

Khác xa với tích chập vanilla, tích chập cổng được dẫn hướng khiến mạng so khớp chỉ học các điểm ảnh hợp lệ bằng cách học để làm nổi bật vùng bị che.

Fig.6 minh họa các đặc trưng của mobilefacenet ở cấp thứ hai ($s=2$) có và không sử dụng tích chập cổng được dẫn hướng từ hình ảnh khuôn mặt bị che ban đầu trên Fig.6 (a).

Tích chập vanilla xử lý tất cả các điểm ảnh như nhau trong cả các vùng bị che và các vùng không bị che, do đó tạo ra các bản đồ đặc trưng được thể hiện trên Fig.6 (b).

Tích chập cổng được dẫn hướng khiến mạng học các điểm ảnh trong các vùng không bị che sử dụng hệ số của hàm softmax (softmax score). Do đó, các bản đồ đặc trưng của tích chập cổng được dẫn hướng trong các vùng bị che có xu hướng có giá trị bằng 0 như được thể hiện trên Fig.6 (c).

Tích chập cổng được dẫn hướng so với tích chập cổng: Tích chập cổng được dẫn hướng theo sáng chế dùng hàm softmax của vùng bị che được tính toán để tinh chỉnh các đặc trưng hình ảnh. Do vùng bị che được tính toán bằng cách tối ưu hóa mạng phân vùng S , nên mạng không yêu cầu nhãn vật che bất kỳ trong quy trình suy luận. Tuy nhiên, tích chập cổng yêu cầu vùng bị che chính xác cho cả thời gian huấn luyện (training) và thời gian suy luận.

Fig.7 minh họa việc so sánh sự phân bố các đặc trưng được kết nối đầy đủ được học trong số các kiểu tích chập khác nhau: tích chập vanilla, tích chập cổng và tích chập cổng được dẫn hướng. Fig.7 thể hiện rằng tất cả các tích chập có thể phân biệt giữa các lớp trong trường hợp các khuôn mặt không bị che (trong hàng thứ nhất). Tuy nhiên, các tích chập hiện có phân loại sai giữa các lớp trong trường hợp khuôn mặt bị che, tức là giữa lớp thứ 2 và lớp thứ 6 đối với tích chập vanilla và lớp thứ 0 và lớp thứ 1 trong tích chập cổng. Trong khi đó, phương pháp theo sáng chế (tức là, tích chập cổng được dẫn hướng) cho phép các cụm tạo với nhau thành các nhóm để phân biệt hơn trong trường hợp các hình ảnh khuôn mặt bị che (trong hàng thứ hai).

Các kết quả định lượng

Thử nghiệm thứ nhất so sánh các kết quả nhận dạng khuôn mặt của mạng theo sáng chế với các phương pháp theo tình trạng kỹ thuật của sáng chế trong cả tập dữ liệu khuôn mặt bị che thực và tổng hợp. Đối với LFW, CFP-FP và AGE-30, các thử nghiệm sử dụng một hình ảnh cho mỗi đối tượng với khuôn mặt không bị che để tạo thành tập tham chiếu, và một hình ảnh khác cho mỗi đối tượng với khuôn mặt bị che để tạo thành tập kiểm tra. Bảng 1 thể hiện độ chính xác của mạng theo sáng chế so với các phương pháp theo tình trạng kỹ thuật trên cả tập dữ liệu khuôn mặt không bị che và khuôn mặt bị che tổng hợp (m biểu thị khuôn mặt bị che). Các độ chính xác cao thu được trên cả hai tập dữ liệu có và không có vật che (khẩu trang) của mạng theo sáng chế thể hiện rằng tích chập công được dẫn hướng cho phép mạng nhận biết thông tin hữu ích để nhận dạng khuôn mặt.

[Bảng 1]

Phương pháp	Giao thức	Độ chính xác ↑						
		LFW	mLFW	CFP-FP	mCFP-FP	Age-30	mAge-30	Avg.
MFNv2-Arcface	Nhỏ	99,60	97,38	95,86	81,01	97,42	88,52	93,29
ShuffleFaceNet	Nhỏ	99,62	-	97,56	-	97,28	-	-
MobileNet	Nhỏ	99,73	96,82	95,16	80,53	96,75	86,65	92,61
Tích chập công	Nhỏ	99,67	97,15	95,63	80,61	97,15	89,33	93,25
GiGaNet	Nhỏ	99,52	98,55	94,90	86,04	96,72	91,78	94,85
R100_Arcface	Lớn	99,77	98,80	98,27	88,90	98,25	93,88	96,31
PDSN	Lớn	99,20	-	-	-	-	-	-
Deng	Lớn	99,60	-	94,05	-	-	-	-
DRGAN	Lớn	-	-	93,41	-	-	-	-
Dream	Lớn	-	-	93,98	-	-	-	-
CurricularFace	Lớn	99,80	98,5	98,42	89,60	98,34	92,66	96,22
Tích chập công	Lớn	99,65	99,40	97,47	93,67	97,57	96,23	97,52
GiGaNet	Lớn	99,70	99,40	97,10	94,11	97,45	96,08	97,30

Để xác minh mạng theo sáng chế trên tập dữ liệu quy mô lớn, thử nghiệm tổng hợp tập dữ liệu khuôn mặt bị che Facescrub. Hình ảnh bị che bao gồm hơn 120 kiểu

khẩu trang khác nhau. Các kết quả trên tập dữ liệu khuôn mặt bị che Facescrub được tổng hợp này được đưa ra trong bảng 2. Không ngạc nhiên khi hiệu quả tương tự giảm trên tập kiểm tra Facescrub ban đầu được quan sát thấy đối với cả mô hình arcface và PDSN. So sánh với mô hình đường cơ sở (PDSN), mạng theo sáng chế vượt trội hơn về tập kiểm tra bị che khuất mà không ảnh hưởng đến hiệu quả trên tập kiểm tra ban đầu.

[Bảng 2]

Phương pháp	Giao thức	#FLOP ↓	Độ chính xác ↑		
			Megaface	mMegaface	Tập dữ liệu nội bộ
MFNv2-Arcface	Nhỏ	0,933G	95,05	47,99	94,67
ShuffleFaceNet	Nhỏ	1,05G	77,4	-	-
MobileNet	Nhỏ	-	91,3	-	-
SphereFace	Nhỏ	-	72,73	-	-
CosFace	Nhỏ	-	77,11	-	-
GiGaNet	Nhỏ	0,937G	94,36	83,53	96,67
R100_ArcFace	Lớn	24,5G	98,35	81,94	96,67
PDSN	Lớn	-	74,40	56,34	-
CurricularFace	Lớn	24,5G	97,95	78,57	96,83
GiGaNet	Lớn	24,9G	97,56	90,90	97,50

Theo một phương án làm ví dụ của sáng chế, có thể triển khai phương pháp và thiết bị nhận dạng khuôn mặt bị che trong đó tích chập được thực hiện mà không có nhân vật che.

Theo một phương án làm ví dụ của sáng chế, có thể triển khai phương pháp và thiết bị nhận dạng khuôn mặt bị che tập trung vào vùng khuôn mặt quan trọng, tức là, vùng không bị che, khi nhận dạng khuôn mặt.

Theo một phương án làm ví dụ của sáng chế, có thể triển khai phương pháp và thiết bị nhận dạng khuôn mặt bị che trong đó chức năng nhận dạng vùng bị che được thực hiện tự động để mang lại hiệu quả nhận dạng khuôn mặt cao.

Theo một phương án làm ví dụ của sáng chế, có thể triển khai phương pháp và thiết bị nhận dạng khuôn mặt bị che mà thể hiện hiệu quả được cải thiện trong cả các

khuôn mặt không bị che và các khuôn mặt bị che với kết quả định tính và định lượng ngay cả khi được đánh giá trên các tập dữ liệu khuôn mặt bị che tổng hợp, các tập dữ liệu khuôn mặt bị che thực và các tập dữ liệu khuôn mặt quy mô lớn tiêu chuẩn.

Các ưu điểm và lợi ích khác nhau của sáng chế không bị giới hạn ở các ưu điểm và lợi ích được mô tả ở trên và có thể được hiểu dễ dàng trong phần mô tả chi tiết các phương án của sáng chế.

Thuật ngữ “bộ phận/đơn vị” được sử dụng trong phương án làm ví dụ của sáng chế có nghĩa là bộ phận/đơn vị phần mềm hoặc phần cứng, chẳng hạn như mảng cổng lập trình được định cấu hình (field-programmable gate array - FPGA) hoặc mạch tích hợp chuyên dụng (application-specific integrated circuit - ASIC), và “bộ phận/đơn vị” thực hiện một vai trò cụ thể. Tuy nhiên, “bộ phận/đơn vị” không bị giới hạn ở phần mềm hoặc phần cứng. “Bộ phận/đơn vị” có thể được tạo cấu hình để nằm trong vật ghi khả năng lập địa chỉ và cũng có thể được tạo cấu hình để chạy một hoặc nhiều bộ xử lý. Do đó, theo một ví dụ, “bộ phận/đơn vị” bao gồm các thành phần, chẳng hạn như các thành phần phần mềm, các thành phần phần cứng hướng đối tượng, các thành phần lớp và các thành phần tác vụ, các quy trình, các chức năng, các thuộc tính, các thủ tục, các chương trình con, các đoạn mã chương trình, các bộ điều khiển, phần sụn, vi mã, các mạch, dữ liệu, cơ sở dữ liệu, các cấu trúc dữ liệu, các bảng, các mảng, và các biến thể. Các thành phần và các chức năng được cung cấp trong “các bộ phận/đơn vị” có thể được ghép với số lượng các thành phần và “các bộ phận/đơn vị” nhỏ hơn hoặc có thể được chia nhỏ thành số lượng các thành phần và “các bộ phận/đơn vị” lớn hơn. Ngoài ra, các thành phần và “các bộ phận/đơn vị” có thể được triển khai để chạy một hoặc nhiều bộ xử lý trung tâm (CPU) trong thiết bị hoặc thẻ đa phương tiện bảo mật.

Tất cả hoặc ít nhất là một số thành phần của thiết bị nêu trên có thể được triển khai dưới dạng mô đun phần cứng, mô đun phần mềm hoặc kết hợp của mô đun phần cứng và mô đun phần mềm. Mô đun phần mềm có thể được hiểu là, ví dụ, lệnh được thực thi bởi bộ xử lý để điều khiển việc tính toán trong thiết bị, và lệnh có thể có dạng được cài đặt trong bộ nhớ của thiết bị.

Thiết bị và phương pháp theo các phương án làm ví dụ của sáng chế có thể được triển khai dưới dạng lệnh chương trình mà có thể được thực hiện qua các phương tiện tính toán khác nhau và có thể được ghi trong vật ghi đọc được bằng máy tính. Vật

ghi đọc được bằng máy tính có thể bao gồm lệnh chương trình, tệp dữ liệu, cấu trúc dữ liệu, v.v ở dạng riêng rẽ hoặc kết hợp. Lệnh chương trình được ghi trong vật ghi có thể được thiết kế và xây dựng đặc biệt cho sáng chế hoặc có thể đã được biết đến và sử dụng bởi người có hiểu biết trung bình trong lĩnh vực phần mềm máy tính. Các ví dụ về vật ghi đọc được bằng máy tính bao gồm phương tiện từ tính như đĩa cứng, đĩa mềm và băng từ, phương tiện quang như đĩa compac (CD)-bộ nhớ chỉ đọc (ROM) và đĩa vạn năng số (DVD), phương tiện quang-từ như đĩa mềm, và các thiết bị phần cứng như ROM, bộ nhớ truy cập ngẫu nhiên (RAM), và bộ nhớ nhanh, được tạo cấu hình đặc biệt để lưu trữ và thực thi lệnh chương trình. Các ví dụ về lệnh chương trình bao gồm không chỉ mã ngôn ngữ máy được tạo ra bởi trình biên dịch mà còn bao gồm mã ngôn ngữ cấp cao có thể thực thi được bởi máy tính qua trình thông dịch và dạng tương tự. Thiết bị phần cứng nêu trên có thể được tạo cấu hình để hoạt động như một hoặc nhiều môđun phần mềm để thực hiện hoạt động của sáng chế, và ngược lại.

Mặc dù sáng chế đã được mô tả ở trên với các phương án làm ví dụ, người có hiểu biết trung bình trong lĩnh vực kỹ thuật này cần hiểu rằng các biến thể và các thay đổi khác nhau có thể được thực hiện từ phần mô tả ở trên mà không nằm ngoài phạm vi của sáng chế được xác định trong yêu cầu bảo hộ kèm theo.

YÊU CẦU BẢO HỘ

1. Phương pháp nhận dạng khuôn mặt bao gồm các bước:

thu hình ảnh khuôn mặt bị che bao gồm vùng bị che và vùng không bị che;

thu đặc trưng hình ảnh từ hình ảnh khuôn mặt bị che;

nhập đặc trưng hình ảnh vào mô hình phân vùng được huấn luyện trước để tự động tính toán đặc trưng của vùng bị che; và

tinh chỉnh đặc trưng hình ảnh sử dụng đặc trưng được tính toán của vùng bị che,

trong đó bước tinh chỉnh bao gồm bước tập trung vào đặc trưng của vùng không bị che và loại bỏ đặc trưng được tính toán của vùng bị che, và

trong đó mô hình phân vùng được huấn luyện bằng cách sử dụng nhãn vật che và đặc trưng khuôn mặt bị che làm các đầu vào và học hệ số của hàm softmax để phân loại một vùng là vùng bị che nếu hệ số của hàm softmax chỉ báo trị số cao và phân loại một vùng là vùng không bị che nếu hệ số của hàm softmax chỉ báo trị số thấp.

2. Phương pháp nhận dạng khuôn mặt theo điểm 1, trong đó phương pháp này còn bao gồm bước tối thiểu hóa hàm mất mát entropi chéo giữa đặc trưng được tính toán của vùng bị che và nhãn vật che.

3. Phương pháp nhận dạng khuôn mặt theo điểm 1, trong đó hình ảnh khuôn mặt bị che được chọn từ nhóm bao gồm hình ảnh khuôn mặt bị che và hình ảnh khuôn mặt không bị che.

4. Phương pháp nhận dạng khuôn mặt theo điểm 1, trong đó phương pháp nhận dạng khuôn mặt dựa vào mạng nơron tích chập sâu và phương pháp nhận dạng khuôn mặt được thực hiện trên nhiều lớp của mạng nơron tích chập sâu.

5. Phương pháp nhận dạng khuôn mặt theo điểm 4, trong đó đặc trưng hình ảnh được lấy mẫu xuống từ cấp thô nhất đến cấp nét nhất trên nhiều lớp.

6. Phương pháp nhận dạng khuôn mặt theo điểm 5, trong đó nhãn vật che được lấy mẫu xuống.

7. Vật ghi đọc được bằng máy tính không khả biến chứa chương trình để thực thi phương pháp nhận dạng khuôn mặt theo điểm 1.

Fig.1

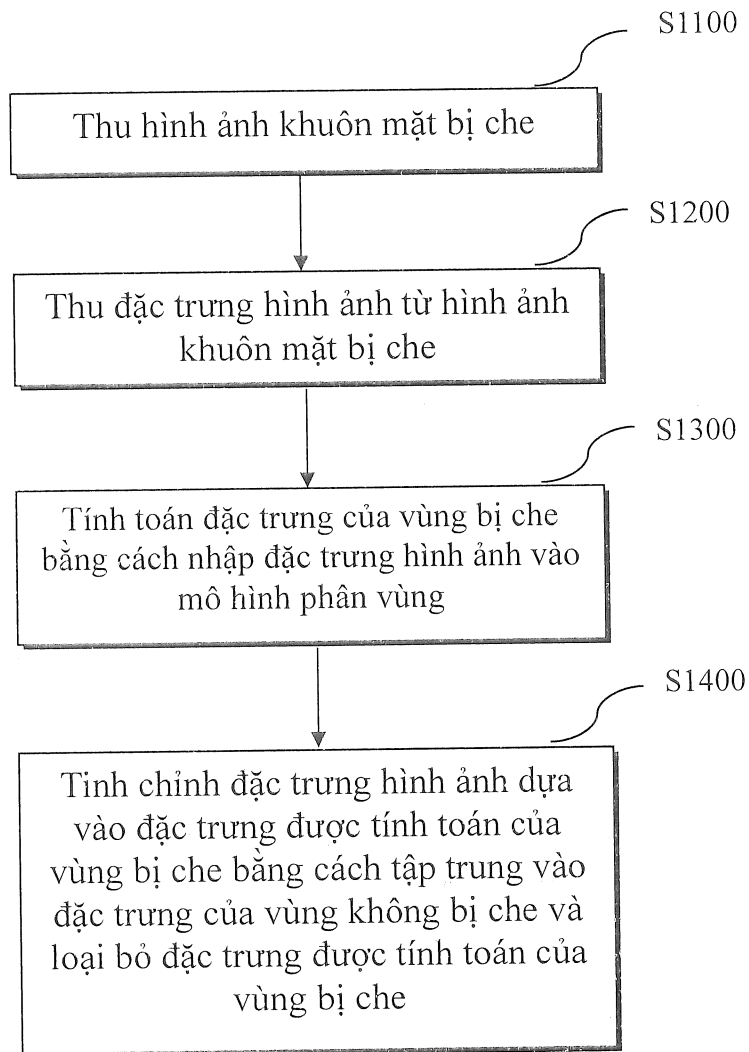


Fig.2

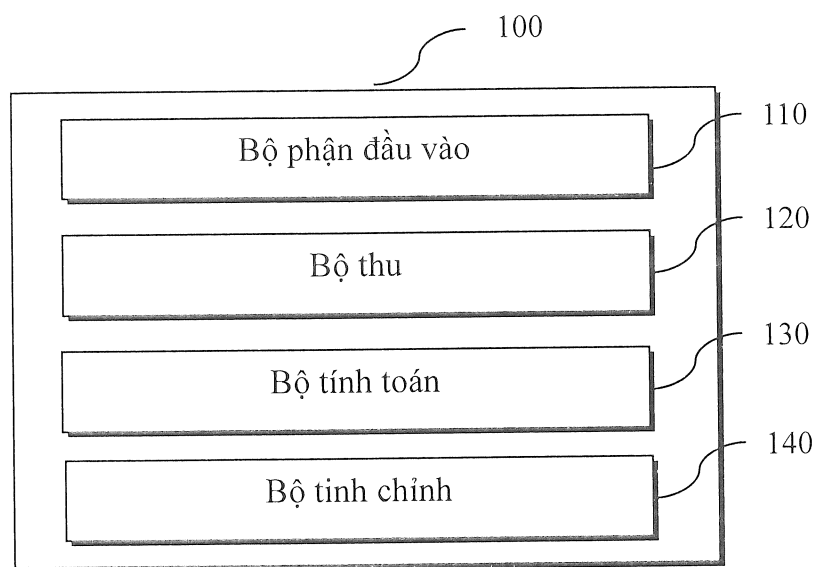


Fig.3

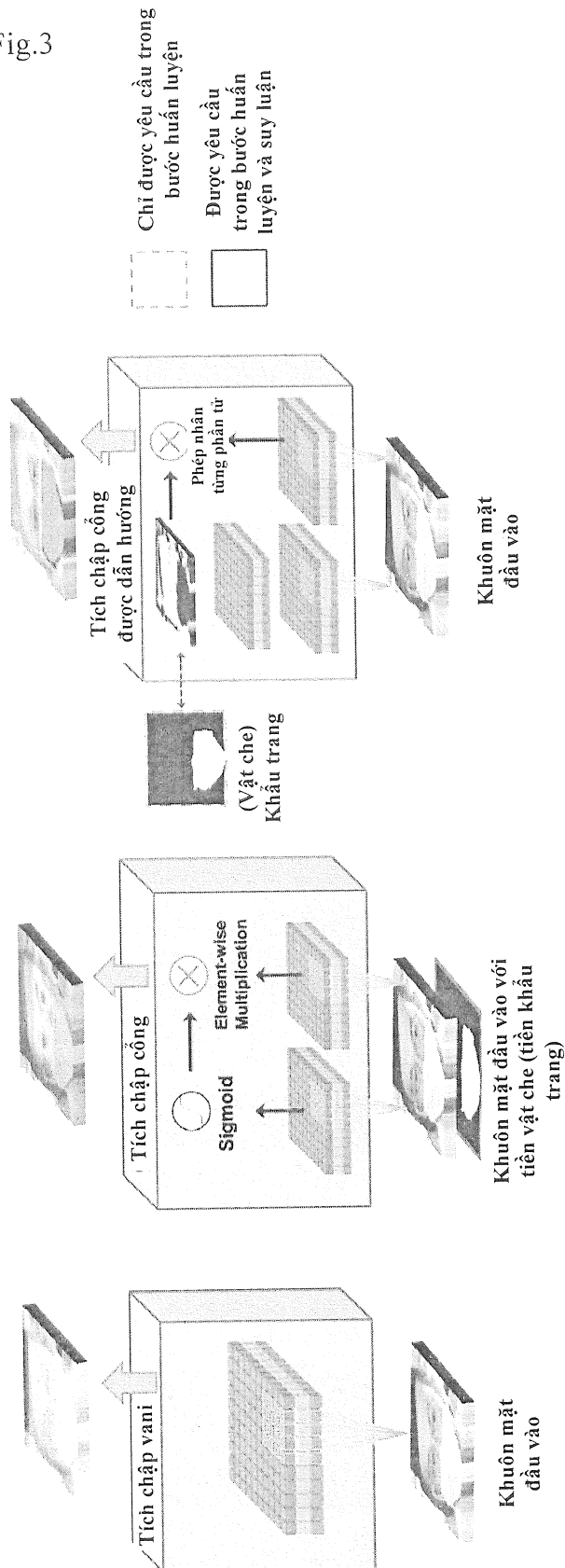


Fig.4

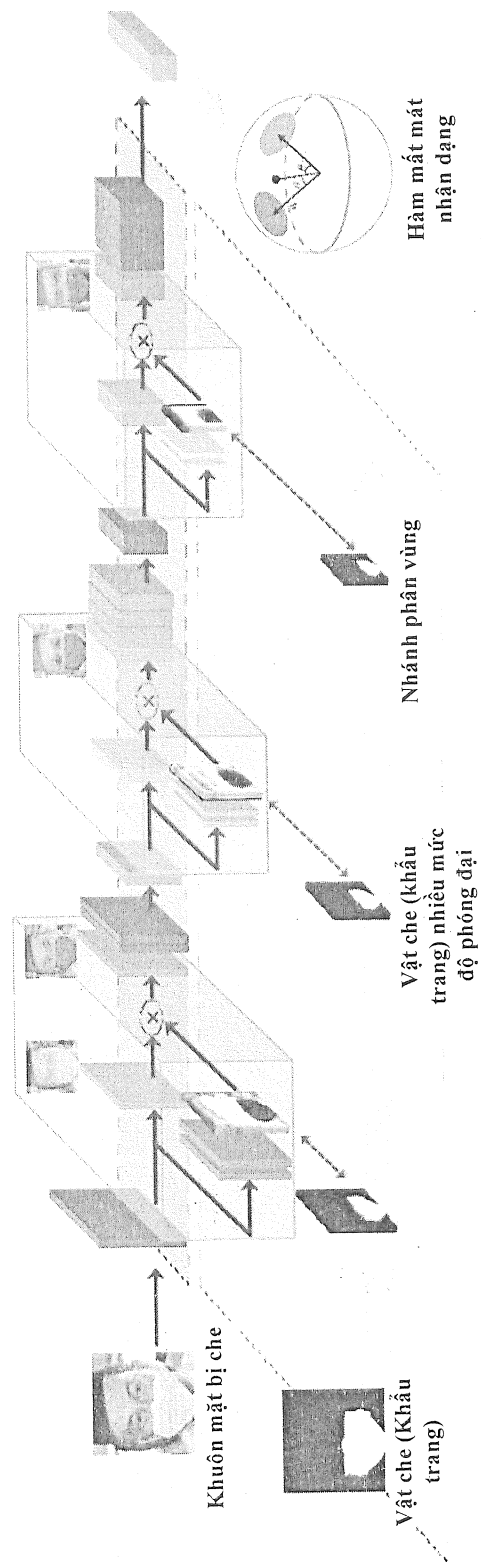


Fig.5

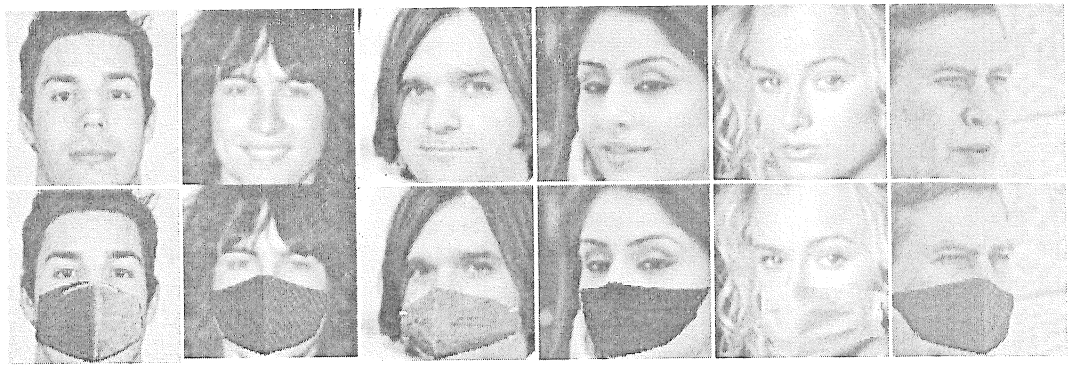


Fig.6

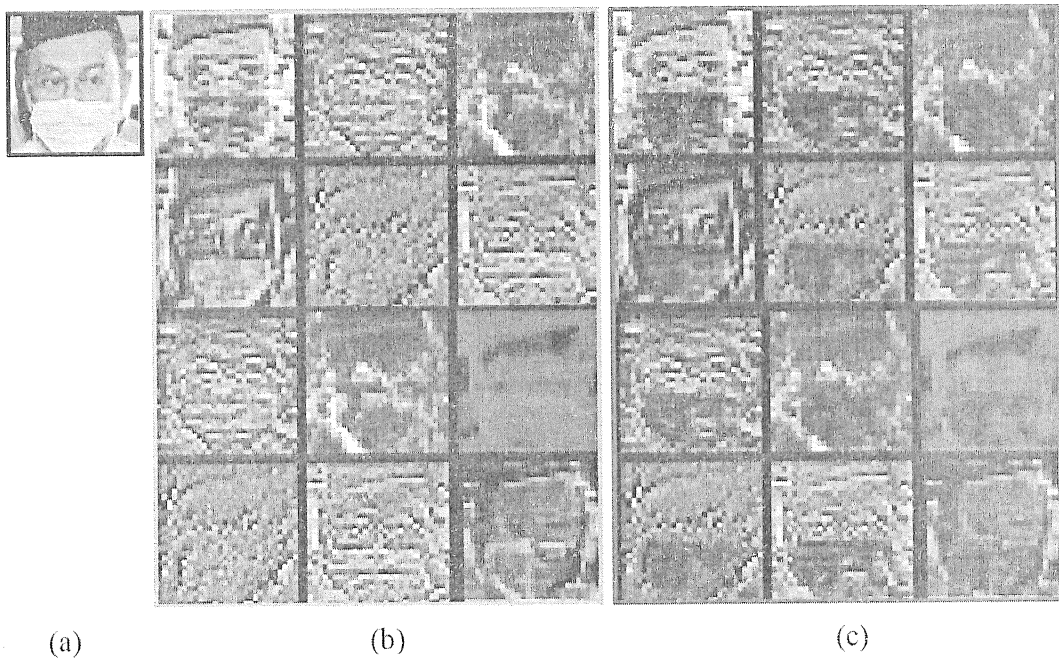


Fig.7

