



(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)



1-0052505

(51)^{2020.01} G06T 1/20; G06T 1/60; G06N 3/08

(13) B

(21) 1-2021-03557

(22) 26/07/2019

(86) PCT/KR2019/009298 26/07/2019

(87) WO2020/101143 22/05/2020

(30) 10-2018-0141951 16/11/2018 KR

(45) 27/10/2025 451

(43) 25/10/2021 403A

(73) SAMSUNG ELECTRONICS CO., LTD. (KR)

129, Samsung-ro, Yeongtong-gu, Suwon-si, Gyeonggi-do, 16677, Republic of Korea

(72) HWANG, Chanyoung (KR); LEE, Wonjae (KR); CHO, Insang (KR).

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

(54) THIẾT BỊ XỬ LÝ HÌNH ẢNH VÀ PHƯƠNG PHÁP VẬN HÀNH THIẾT BỊ NÀY

(21) 1-2021-03557

(57) Sáng chế đề cập đến thiết bị xử lý hình ảnh để xử lý hình ảnh bằng cách sử dụng một hoặc nhiều mạng thần kinh nhân tạo, và bao gồm bộ nhớ lưu trữ một hoặc nhiều lệnh và các cấu trúc dữ liệu cho mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ, và bộ xử lý được tạo cấu hình để thực thi một hoặc nhiều lệnh được lưu trữ trong bộ nhớ để xử lý hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính để thu được dữ liệu kết quả trung gian và hình ảnh đầu ra cuối, và để xử lý dữ liệu kết quả trung gian bằng cách sử dụng mạng thần kinh nhân tạo phụ để đưa ra hình ảnh trung gian trong khi hình ảnh đầu vào đang được xử lý bằng cách sử dụng mạng thần kinh nhân tạo chính. Sáng chế cũng đề cập đến phương pháp vận hành thiết bị xử lý hình ảnh này.

[Fig. 2]



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến các thiết bị xử lý hình ảnh và phương pháp vận hành thiết bị này, và cụ thể hơn, đề cập đến thiết bị xử lý hình ảnh mà xử lý hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo bao gồm nhiều lớp tích chập, và phương pháp vận hành thiết bị xử lý hình ảnh.

Tình trạng kỹ thuật của sáng chế

Các hệ thống trí tuệ nhân tạo (Artificial intelligence, AI) là các hệ thống máy tính được tạo cấu hình để nhận dạng trí thông minh cấp độ con người và tự huấn luyện và đưa ra các quyết định đồng thời để trở nên thông minh hơn, trái ngược với các hệ thống thông minh dựa trên quy tắc hiện có. Do tỷ lệ nhận dạng của các hệ thống AI được cải thiện và hệ thống AI hiểu chính xác hơn sở thích của người dùng khi chúng được sử dụng nhiều hơn, các hệ thống thông minh dựa trên quy tắc hiện có đang dần được thay thế bằng các hệ thống AI học sâu.

Công nghệ AI bao gồm học máy (ví dụ, học sâu) và các công nghệ phần tử sử dụng học máy. Học máy là công nghệ thuật toán tự phân loại/học hỏi các đặc tính của dữ liệu đầu vào và sử dụng thuật toán học máy, chẳng hạn như học sâu và bao gồm các lĩnh vực kỹ thuật, chẳng hạn như hiểu ngôn ngữ, hiểu trực quan, suy luận/dự đoán, biểu diễn tri thức, và kiểm soát hoạt động.

Các lĩnh vực khác nhau mà công nghệ AI được áp dụng như sau. Hiểu ngôn ngữ là kỹ thuật nhận dạng ngôn ngữ/ký tự của con người và áp dụng/xử lý ngôn ngữ/ký tự của con người và bao gồm xử lý ngôn ngữ tự nhiên, dịch máy, hệ thống hội thoại, câu hỏi và câu trả lời, nhận dạng/tổng hợp giọng nói, và tương tự. Hiểu thị giác là kỹ thuật nhận dạng và xử lý một đối tượng giống như trong thị giác của con người, và bao gồm nhận dạng đối tượng, theo dõi đối tượng, tìm kiếm hình ảnh, nhận dạng con người, hiểu cảnh, hiểu không gian, cải thiện hình ảnh, và tương tự. Suy luận/dự đoán là công nghệ thực hiện suy luận và dự đoán một cách hợp lý bằng cách xác định thông tin và bao gồm suy luận dựa trên kiến thức/xác suất, dự đoán tối ưu hóa, kế

hoạch dựa trên sở thích, đề xuất, và tương tự. Biểu diễn tri thức là một kỹ thuật xử lý tự động thông tin trải nghiệm của con người dưới dạng dữ liệu tri thức và bao gồm thiết lập tri thức (tạo/phân loại dữ liệu), quản lý tri thức (sử dụng dữ liệu), và dạng tương tự. Điều khiển vận hành là kỹ thuật điều khiển lái xe tự động và các chuyển động của rô bốt, bao gồm điều khiển chuyển động (điều hướng, tránh va chạm và lái xe), điều khiển thao tác (điều khiển hành vi), và tương tự.

Bản chất kỹ thuật của sáng chế

Giải pháp kỹ thuật

Sáng chế đề xuất thiết bị xử lý hình ảnh có thể tạo ra hình ảnh đầu ra trung gian bằng cách sử dụng mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ, và phương pháp vận hành thiết bị xử lý hình ảnh.

Hiệu quả của sáng chế

Thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế có thể tạo ra hình ảnh đầu ra trung gian bằng cách sử dụng mạng thần kinh nhân tạo phụ trước khi tạo ra hình ảnh đầu ra cuối trong mạng thần kinh nhân tạo chính, và vì vậy thời gian chờ từ điểm quan sát của người dùng có thể được giảm xuống.

Thiết bị xử lý hình ảnh có thể tạo ra hình ảnh đầu ra trung gian, và người dùng có thể xác định xem có tạo ra hình ảnh đầu ra cuối trong mạng thần kinh nhân tạo chính, dựa trên hình ảnh đầu ra trung gian hay không.

Thiết bị xử lý hình ảnh có thể bao gồm các mạng thần kinh nhân tạo đưa ra các kết quả khác nhau mặc dù xử lý một phần các hình ảnh tương tự nhau, như mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ, nhờ đó tăng được hiệu quả huấn luyện mạng thần kinh nhân tạo.

Mô tả vắn tắt các hình vẽ

Các khía cạnh, dấu hiệu và ưu điểm nêu trên và khác nữa của các phương án thực hiện nhất định của sáng chế sẽ trở nên rõ ràng hơn qua phần mô tả dưới đây kết

hợp các hình vẽ kèm theo, trong đó:

Fig.1 thể hiện quy trình trong đó thiết bị xử lý hình ảnh theo một phương án thực hiện xử lý hình ảnh bằng cách sử dụng mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ;

Fig.2 là lưu đồ của phương pháp vận hành thiết bị xử lý hình ảnh, theo một phương án thực hiện của sáng chế;

Fig.3 là sơ đồ được sử dụng để giải thích phương pháp trong đó thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế xử lý hình ảnh bằng cách sử dụng mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ;

Fig.4 là sơ đồ thể hiện cấu trúc của mạng thần kinh nhân tạo chính theo một phương án thực hiện của sáng chế;

Fig.5 thể hiện dữ liệu đầu vào, các nhân, và dữ liệu đầu ra của lớp tích chập thứ nhất có trong mạng thần kinh nhân tạo chính theo một phương án thực hiện của sáng chế.

Fig.6 là sơ đồ được sử dụng để mô tả quy trình trong đó lớp tích chập thứ nhất trên Fig.5 thực hiện hoạt động tích chập;

Fig.7 là sơ đồ được sử dụng để mô tả phương pháp trong đó thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế hoạt động;

Fig.8 thể hiện một ví dụ trong đó thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế hiển thị các hình ảnh đầu ra trung gian;

Fig.9 thể hiện một ví dụ trong đó thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế hiển thị các hình ảnh đầu ra trung gian;

Fig.10 thể hiện phương pháp trong đó thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế xác định xem liệu có truy cập vào mạng thần kinh nhân tạo phụ hay không;

Fig.11 là sơ đồ được sử dụng để giải thích phương pháp trong đó thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế tạo ra các hình ảnh đầu ra bằng cách sử dụng mạng thần kinh nhân tạo chính và các mạng thần kinh nhân tạo phụ;

Fig.12 là sơ đồ thể hiện ví dụ trong đó thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế và máy chủ tương tác với nhau để thực hiện việc xử lý hình ảnh;

Fig.13 là sơ đồ khối được sử dụng để giải thích phương pháp huấn luyện mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ theo một phương án thực hiện của sáng chế;

Fig.14 là sơ đồ khối của cấu trúc của thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế;

Fig.15 là sơ đồ khối của bộ xử lý theo một phương án thực hiện của sáng chế;

Fig.16 là sơ đồ thể hiện một ví dụ trong đó thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế và máy chủ tương tác với nhau để huấn luyện và nhận diện dữ liệu; và

Fig.17 là sơ đồ khối của cấu trúc của thiết bị xử lý hình ảnh theo một phương án thực hiện khác của sáng chế.

Mô tả chi tiết sáng chế

Thiết bị xử lý hình ảnh có thể tạo ra hình ảnh đầu ra trung gian bằng cách sử dụng mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ, và phương pháp vận hành thiết bị xử lý hình ảnh được đề xuất.

Các khía cạnh bổ sung sẽ được đưa ra từng phần trong bản mô tả này theo sau và một phần sẽ thấy rõ qua bản mô tả, hoặc có thể biết được bằng cách thực hành các phương án thực hiện được trình bày của sáng chế.

Theo một phương án thực hiện của sáng chế, thiết bị xử lý hình ảnh có thể bao

gồm: bộ nhớ lưu trữ một hoặc nhiều lệnh và các cấu trúc dữ liệu cho mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ; và bộ xử lý được tạo cấu hình để thực thi một hoặc nhiều lệnh được lưu trữ trong bộ nhớ để xử lý hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính để thu được dữ liệu kết quả trung gian và hình ảnh đầu ra cuối, và xử lý dữ liệu kết quả trung gian bằng cách sử dụng mạng thần kinh nhân tạo phụ để đưa ra hình ảnh trung gian trong khi hình ảnh đầu vào đang được xử lý bằng cách sử dụng mạng thần kinh nhân tạo chính để thu được hình ảnh đầu ra cuối.

Mạng thần kinh nhân tạo chính có thể bao gồm nhiều lớp tích chập thứ nhất, và bộ xử lý có thể còn được tạo cấu hình để thực hiện hoạt động tích chập thứ nhất trên hình ảnh đầu vào bằng cách sử dụng một hoặc nhiều nhân trong nhiều lớp tích chập thứ nhất để trích xuất thông tin đặc điểm và tạo ra dữ liệu kết quả trung gian dựa trên thông tin đặc điểm được trích xuất.

Mạng thần kinh nhân tạo phụ có thể bao gồm một hoặc nhiều lớp tích chập thứ hai, và bộ xử lý có thể còn được tạo cấu hình để thực hiện hoạt động tích chập thứ hai bằng cách áp dụng một hoặc nhiều nhân cho dữ liệu kết quả trung gian trong một hoặc nhiều lớp tích chập thứ hai để trích xuất thông tin đặc điểm và đưa ra hình ảnh trung gian dựa trên thông tin đặc điểm được trích xuất.

Mạng thần kinh nhân tạo chính có thể là mạng được huấn luyện để tạo ra hình ảnh đầu ra cuối bằng cách áp dụng kiểu định trước cho hình ảnh đầu vào, và mức độ mà kiểu định trước đã được áp dụng cho hình ảnh trung gian có thể nhỏ hơn mức mà kiểu định trước đã được áp dụng cho hình ảnh đầu ra cuối.

Khoảng thời gian thứ nhất được thực hiện để tạo ra hình ảnh đầu ra cuối từ hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính có thể lớn hơn so với khoảng thời gian thứ hai được thực hiện để tạo ra hình ảnh trung gian từ hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo phụ.

Bộ xử lý còn có thể được tạo cấu hình để dự đoán khoảng thời gian thứ nhất và khoảng thời gian thứ hai và xác định xem liệu có tạo ra hình ảnh trung gian bằng cách sử dụng mạng thần kinh nhân tạo phụ hay không, dựa trên khoảng thời gian thứ nhất đã được dự đoán và khoảng thời gian thứ hai đã được dự đoán.

Thiết bị xử lý hình ảnh có thể còn bao gồm màn hình, trong đó bộ xử lý còn được tạo cấu hình để điều khiển màn hình để hiển thị khoảng thời gian được thực hiện để tạo ra ít nhất một trong số hình ảnh trung gian hoặc hình ảnh đầu ra cuối từ thời điểm hiện tại.

Bộ xử lý còn có thể được tạo cấu hình để dừng việc xử lý hình ảnh đầu vào trước khi hình ảnh đầu ra cuối được tạo ra, dựa vào đầu vào của người dùng.

Mạng thần kinh nhân tạo chính có thể bao gồm lớp đầu vào, lớp đầu ra, nhiều lớp tích chập thứ nhất nằm giữa lớp đầu vào và lớp đầu ra, và ít nhất một lớp đầu ra trung gian nằm giữa hai lớp tích chập tuần tự trong số nhiều lớp tích chập thứ nhất. Bộ xử lý còn có thể được tạo cấu hình để truyền dữ liệu kết quả trung gian đến mạng thần kinh nhân tạo phụ qua ít nhất một lớp đầu ra trung gian trong khi xử lý hình ảnh đầu vào để thu được hình ảnh đầu ra cuối qua nhiều lớp tích chập thứ nhất.

Theo một phương án thực hiện khác của sáng chế, thiết bị xử lý hình ảnh có thể bao gồm: giao diện truyền thông được tạo cấu hình để nhận dữ liệu kết quả trung gian từ máy chủ bao gồm mạng thần kinh nhân tạo chính trong khi máy chủ đang xử lý hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính; bộ nhớ lưu trữ một hoặc nhiều lệnh và cấu trúc dữ liệu cho mạng thần kinh nhân tạo phụ; và bộ xử lý được tạo cấu hình để thực thi một hoặc nhiều lệnh được lưu trữ trong bộ nhớ để đưa ra hình ảnh trung gian bằng cách xử lý dữ liệu kết quả trung gian dựa trên mạng thần kinh nhân tạo phụ.

Mạng thần kinh nhân tạo chính có thể bao gồm lớp đầu vào, lớp đầu ra, nhiều lớp tích chập thứ nhất nằm giữa lớp đầu vào và lớp đầu ra, và ít nhất một lớp đầu ra

trung gian nằm giữa hai lớp tích chập tuần tự trong số nhiều lớp tích chập thứ nhất. Bộ xử lý còn có thể được tạo cấu hình để khiến mạng thần kinh nhân tạo phụ nhận dữ liệu kết quả trung gian từ ít nhất một lớp đầu ra trung gian của mạng thần kinh nhân tạo chính trong khi hình ảnh đầu vào đang được xử lý qua nhiều lớp tích chập thứ nhất của mạng thần kinh nhân tạo chính.

Theo một phương án thực hiện khác của sáng chế, phương pháp vận hành thiết bị xử lý hình ảnh có thể bao gồm các bước: xử lý hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính; thu được dữ liệu kết quả trung gian từ hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính trong khi xử lý hình ảnh đầu vào để thu được hình ảnh đầu ra cuối; và tạo ra hình ảnh trung gian dựa trên dữ liệu kết quả trung gian bằng cách sử dụng mạng thần kinh nhân tạo phụ.

Mạng thần kinh nhân tạo chính có thể bao gồm nhiều lớp tích chập thứ nhất, bước xử lý hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính có thể bao gồm bước trích xuất thông tin đặc điểm bằng cách thực hiện hoạt động tích chập thứ nhất trên hình ảnh đầu vào và một hoặc nhiều nhân trong nhiều lớp tích chập thứ nhất, và bước thu được dữ liệu kết quả trung gian có thể bao gồm bước tạo ra dữ liệu kết quả trung gian, dựa trên thông tin đặc điểm được trích xuất.

Mạng thần kinh nhân tạo phụ có thể bao gồm một hoặc nhiều lớp tích chập thứ hai, và bước tạo ra hình ảnh trung gian có thể bao gồm bước: trích xuất thông tin đặc điểm bằng cách thực hiện hoạt động tích chập thứ hai bằng cách áp dụng dữ liệu kết quả trung gian cho một hoặc nhiều nhân trong một hoặc nhiều lớp tích chập thứ hai; và tạo ra hình ảnh trung gian, dựa trên thông tin đặc điểm được trích xuất.

Mạng thần kinh nhân tạo chính có thể là mạng được huấn luyện để tạo ra hình ảnh cuối bằng cách áp dụng kiểu định trước cho hình ảnh đầu vào, và mức độ mà kiểu định trước đã được áp dụng cho hình ảnh trung gian có thể nhỏ hơn so với mức độ mà kiểu định trước đã được áp dụng cho hình ảnh đầu ra cuối.

Khoảng thời gian thứ nhất được thực hiện để tạo ra hình ảnh đầu ra cuối từ hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính có thể lớn hơn so với khoảng thời gian thứ hai được thực hiện để tạo ra hình ảnh trung gian từ hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo phụ.

Phương pháp này có thể còn bao gồm các bước: dự đoán khoảng thời gian thứ nhất và khoảng thời gian thứ hai; và xác định xem liệu có tạo ra hình ảnh trung gian bằng cách sử dụng mạng thần kinh nhân tạo phụ hay không, dựa trên khoảng thời gian thứ nhất đã được dự đoán và khoảng thời gian thứ hai đã được dự đoán.

Phương pháp này có thể còn bao gồm bước hiển thị, trên màn hình, khoảng thời gian được thực hiện để tạo ra ít nhất một trong số hình ảnh trung gian hoặc hình ảnh đầu ra cuối từ thời điểm hiện tại.

Phương pháp này có thể còn bao gồm bước dùng mạng thần kinh nhân tạo chính xử lý hình ảnh đầu vào, trước khi hình ảnh đầu ra cuối được tạo ra, dựa vào đầu vào của người dùng.

Theo một phương án thực hiện khác của sáng chế, sản phẩm chương trình máy tính bao gồm phương tiện ghi có thể đọc được bằng máy tính không nhất thời đã ghi trên đó chương trình máy tính, khi được thực thi bởi máy tính, có thể thực hiện phương pháp vận hành thiết bị xử lý hình ảnh.

Các phương án thực hiện sáng chế

Xuyên suốt sáng chế, cách diễn tả "ít nhất một trong số a, b hoặc c" chỉ báo chỉ a, chỉ b, chỉ c, cả a và b, cả a và c, cả b và c, tất cả a, b và c, hoặc các biến thể của chúng.

Dưới đây, các thuật ngữ được sử dụng trong bản mô tả sáng chế này sẽ được mô tả ngắn gọn và sau đó sáng chế sẽ được mô tả chi tiết.

Mặc dù các thuật ngữ chung được sử dụng rộng rãi hiện nay đã được lựa chọn

để mô tả sáng chế khi xem xét các chức năng của chúng, nhưng các thuật ngữ chung này có thể thay đổi tùy theo ý định của người có hiểu biết trung bình trong lĩnh vực kỹ thuật này, trường hợp tiền lệ, là sự ra đời của công nghệ mới và những dạng tương tự. Các thuật ngữ được lựa chọn tùy ý bởi Người nộp đơn của sáng chế cũng có thể được sử dụng trong trường hợp cụ thể. Trong trường hợp này, ý nghĩa của chúng cần được đưa ra trong phần mô tả chi tiết. Do đó, các thuật ngữ phải được xác định dựa trên ý nghĩa của chúng và nội dung của toàn bộ bản mô tả, chứ không phải chỉ bằng cách nêu đơn giản các thuật ngữ này.

Các thuật ngữ “gồm có” và/hoặc “gồm” hoặc “bao gồm” và/hoặc “bao hàm” khi được sử dụng trong phần mô tả này, chỉ rõ sự có mặt của các thành phần đã nêu, nhưng không loại trừ sự có mặt hoặc bổ sung một hoặc nhiều thành phần khác. Các thuật ngữ “đơn vị”, “bộ” và “môđun” khi được sử dụng trong bản mô tả này đề cập đến đơn vị trong đó ít nhất một chức năng hoặc hoạt động được thực hiện và có thể được triển khai dưới dạng phần cứng, phần mềm hoặc kết hợp giữa phần cứng và phần mềm.

Các phương án thực hiện của sáng chế được mô tả một cách chi tiết ở đây có dựa vào các hình vẽ kèm theo để sáng chế có thể dễ dàng được thực hiện bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật mà sáng chế đề cập đến. Tuy nhiên, sáng chế có thể được thực hiện dưới nhiều dạng khác nhau và không nên được hiểu là giới hạn ở các phương án thực hiện được nêu ở đây. Trong các hình vẽ, các phần không liên quan đến phần mô tả này được bỏ qua để đơn giản hóa việc giải thích, và các số giống nhau biểu thị các thành phần tương tự xuyên suốt.

Fig.1 thể hiện quy trình trong đó thiết bị xử lý hình ảnh 100 theo một phương án thực hiện của sáng chế xử lý hình ảnh bằng cách sử dụng mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ.

Thiết bị xử lý hình ảnh 100 có thể là loại thiết bị điện tử bất kỳ, chẳng hạn như

điện thoại di động, máy tính bảng cá nhân (PC), máy ảnh kỹ thuật số, máy quay phim, máy tính xách tay, máy tính để bàn, thiết bị đầu cuối sách điện tử, thiết bị đầu cuối phát sóng kỹ thuật số, trợ lý kỹ thuật số cá nhân (PDA - Personal Digital Assistant), trình phát đa phương tiện di động (PMP - Portable Multimedia Player), thiết bị điều hướng, máy nghe nhạc MP3 hoặc thiết bị đeo.

Thiết bị xử lý hình ảnh 100 có thể được thực hiện bằng cách sử dụng thiết bị hiển thị hình ảnh bao gồm màn hình. Cụ thể, thiết bị xử lý hình ảnh 100 theo các phương án thực hiện có thể dễ dàng được thực hiện trong thiết bị hiển thị hình ảnh có màn hình lớn, như TV. Tuy nhiên, các phương án thực hiện của sáng chế không giới hạn ở đó. Ví dụ, thiết bị xử lý hình ảnh 100 có thể là thiết bị cố định hoặc di động, hoặc thiết bị thu phát sóng kỹ thuật số. Thiết bị xử lý hình ảnh 100 có thể được thực hiện bằng cách sử dụng không chỉ thiết bị hiển thị phẳng mà còn sử dụng thiết bị hiển thị cong có độ cong hoặc thiết bị hiển thị mềm dẻo với độ cong có thể điều chỉnh được. Ví dụ, độ phân giải đầu ra của thiết bị xử lý hình ảnh 100 có thể là độ phân giải cao (HD), đủ HD, siêu HD, hoặc độ phân giải rõ ràng hơn siêu HD.

Thuật ngữ “người dùng” được sử dụng ở đây biểu thị người điều khiển chức năng hoặc hoạt động của thiết bị xử lý hình ảnh. Các ví dụ về người dùng có thể bao gồm quản lý và kỹ sư cài đặt.

Thiết bị xử lý hình ảnh 100 có thể bao gồm mạng thần kinh nhân tạo chính 200 được huấn luyện để xử lý hình ảnh theo mục đích đặt trước. Mạng thần kinh nhân tạo chính 200 có thể bao gồm, nhưng không giới hạn ở, nhiều lớp tích chập.

Ví dụ, mạng thần kinh nhân tạo chính 200 có thể là một mô hình được huấn luyện để đưa ra hình ảnh chuyển kiểu 20 bằng cách áp dụng kiểu định trước 15 cho hình ảnh đầu vào 10. Trong trường hợp này, kiểu có thể được xác định theo đường nét, màu sắc, vật liệu, và khí quyển xuất hiện trên hình ảnh, nhưng các phương án thực hiện của sáng chế không giới hạn ở đó.

Thiết bị xử lý hình ảnh 100 có thể xử lý hình ảnh đầu vào ví dụ, hình ảnh thứ nhất) 10 bằng cách sử dụng mạng thần kinh nhân tạo chính 200 để thu được hình ảnh đầu ra cuối (ví dụ, hình ảnh thứ hai) 20. Thiết bị xử lý hình ảnh 100 có thể thu được dữ liệu kết quả trung gian mà được đưa ra trong khi hình ảnh đầu vào 10 đang được xử lý và có thể nhập dữ liệu kết quả trung gian vào mạng thần kinh nhân tạo phụ 300. Thiết bị xử lý hình ảnh 100 có thể xử lý dữ liệu kết quả trung gian bằng cách sử dụng mạng thần kinh nhân tạo phụ 300 để tạo ra hình ảnh trung gian (ví dụ, hình ảnh thứ ba). Mạng thần kinh nhân tạo phụ 300 có thể bao gồm, nhưng không giới hạn ở, một hoặc nhiều lớp tích chập.

Thiết bị xử lý hình ảnh 100 có thể thu được hình ảnh đầu ra trung gian 30 khác với hình ảnh đầu ra cuối 20, bằng cách xử lý dữ liệu kết quả trung gian bằng cách sử dụng mạng thần kinh nhân tạo phụ 300.

Khi xử lý hình ảnh bằng cách sử dụng một mạng duy nhất, thiết bị xử lý hình ảnh 100 có thể thu được hình ảnh đầu ra trung gian 30 bằng cách sử dụng dữ liệu kết quả trung gian, trước khi thu được hình ảnh đầu ra cuối 20.

Do đó, thiết bị xử lý hình ảnh 100 có thể thu được hình ảnh đầu ra trung gian 30 trước khi hình ảnh đầu ra cuối được tạo ra, và, xét theo quan điểm của người dùng thiết bị xử lý hình ảnh 100, có thể giảm được thời gian chờ cho tới khi hình ảnh đầu ra cuối 20 được tạo ra.

Fig.2 là lưu đồ của phương pháp vận hành thiết bị xử lý hình ảnh 100, theo một phương án thực hiện của sáng chế.

Tham khảo tới Fig.2, ở bước S210, thiết bị xử lý hình ảnh 100 có thể xử lý hình ảnh thứ nhất bằng cách sử dụng mạng thần kinh nhân tạo chính.

Mạng thần kinh nhân tạo chính theo một phương án thực hiện của sáng chế có thể là mô hình được huấn luyện để xử lý hình ảnh theo mục đích định trước. Mạng thần kinh nhân tạo chính có thể là mạng thần kinh nhân tạo tích chập (CNN -

Convolutional Neural Network) bao gồm nhiều lớp tích chập. Ví dụ, mạng thần kinh nhân tạo chính có thể là mô hình được huấn luyện để đưa ra hình ảnh chuyển kiểu bằng cách áp dụng kiểu định trước cho hình ảnh đầu vào. Một cách tùy ý, mạng thần kinh nhân tạo chính có thể là mô hình được huấn luyện để chuyển đổi hình ảnh có độ phân giải thấp đầu vào thành hình ảnh có độ phân giải cao và đưa ra hình ảnh có độ phân giải cao. Tuy nhiên, các phương án thực hiện của sáng chế không giới hạn ở đó.

Thiết bị xử lý hình ảnh 100 có thể thu được hình ảnh đầu ra cuối bằng cách xử lý hình ảnh đầu vào thứ nhất bằng cách sử dụng mạng thần kinh nhân tạo chính.

Ở bước S220, thiết bị xử lý hình ảnh 100 có thể thu được dữ liệu kết quả trung gian mà được đưa ra trong khi hình ảnh thứ nhất đang được xử lý.

Thiết bị xử lý hình ảnh 100 có thể thu được dữ liệu đã đi qua một số lớp tích chập có trong mạng thần kinh nhân tạo chính.

Ở bước S230, thiết bị xử lý hình ảnh 100 có thể tạo ra hình ảnh thứ hai bằng cách xử lý dữ liệu kết quả trung gian bằng cách sử dụng mạng thần kinh nhân tạo phụ.

Thiết bị xử lý hình ảnh 100 có thể xử lý dữ liệu kết quả trung gian bằng cách sử dụng mạng thần kinh nhân tạo phụ để thu được hình ảnh đầu ra trung gian (ví dụ, hình ảnh thứ ba) khác với hình ảnh đầu ra cuối (ví dụ, hình ảnh thứ hai). Trong trường hợp này, khi mạng thần kinh nhân tạo chính là mô hình được huấn luyện để áp dụng kiểu định trước, mức độ mà kiểu định trước đã được áp dụng trong hình ảnh đầu ra cuối có thể khác với mức độ mà kiểu định trước đã được áp dụng trong hình ảnh đầu ra trung gian. Ví dụ, mức độ mà kiểu định trước đã được áp dụng trong hình ảnh đầu ra trung gian có thể nhỏ hơn so với mức độ mà kiểu định trước đã được áp dụng trong hình ảnh đầu ra cuối. Tuy nhiên, các phương án thực hiện của sáng chế không giới hạn ở đó.

Fig.3 là sơ đồ được sử dụng để giải thích phương pháp, được thực hiện bởi thiết bị xử lý hình ảnh 100 theo một phương án thực hiện của sáng chế, để xử lý hình

ảnh bằng cách sử dụng mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ.

Tham khảo tới Fig.3, thiết bị xử lý hình ảnh 100 có thể bao gồm mạng thần kinh nhân tạo chính 200 và mạng thần kinh nhân tạo phụ 300.

Mạng thần kinh nhân tạo chính 200 có thể là mô hình được huấn luyện để nhận hình ảnh thứ nhất và xử lý hình ảnh đầu vào thứ nhất theo mục đích định trước để đưa ra hình ảnh thứ hai. Ví dụ, như được minh họa trên Fig.3, mạng thần kinh nhân tạo chính 200 có thể là mô hình được huấn luyện để đưa ra hình ảnh chuyển kiểu 302 bằng cách áp dụng kiểu định trước cho hình ảnh đầu vào 301. Tuy nhiên, được huấn luyện cho mục đích chuyển đổi kiểu chỉ là một ví dụ, và mạng thần kinh nhân tạo chính 200 có thể là mô hình được huấn luyện cho các mục đích khác nhau.

Mạng thần kinh nhân tạo chính 200 có thể có cấu trúc trong đó dữ liệu (ví dụ, dữ liệu hình ảnh) được nhập vào và được xử lý trong khi đi qua các lớp ẩn và vì vậy dữ liệu được xử lý được đưa ra. Mạng thần kinh nhân tạo chính 200 có thể là mạng thần kinh nhân tạo sâu (DNN) bao gồm hai hoặc nhiều lớp ẩn 210, 220. Ví dụ, mạng thần kinh nhân tạo chính 200 có thể là CNN bao gồm nhiều lớp tích chập.

Mạng thần kinh nhân tạo phụ 300 có thể nhận dữ liệu kết quả trung gian được đưa ra từ mạng thần kinh nhân tạo chính 200, và xử lý dữ liệu kết quả trung gian theo mục đích định trước để đưa ra hình ảnh thứ ba 303. Ví dụ, mạng thần kinh nhân tạo phụ 300 có cấu trúc trong đó dữ liệu trung gian được nhập vào và được xử lý trong khi đi qua một hoặc nhiều lớp tích chập (các lớp ẩn) và vì vậy dữ liệu được xử lý 303 (ví dụ, hình ảnh thứ ba) được đưa ra.

Các cấu trúc và hoạt động của mạng thần kinh nhân tạo chính 200 và mạng thần kinh nhân tạo phụ 300 sẽ được mô tả dưới đây một cách chi tiết có dựa vào các hình vẽ từ Fig.4 đến Fig.6.

Fig.4 thể hiện cấu trúc của mạng thần kinh nhân tạo chính 200 theo một

phương án thực hiện của sáng chế.

Tham khảo tới Fig.4, mạng thần kinh nhân tạo chính 200 có thể bao gồm lớp đầu vào, lớp đầu ra, và nhiều lớp ẩn, như N lớp tích chập 452 nằm giữa lớp đầu vào và lớp đầu ra. Mạng thần kinh nhân tạo chính 200 có thể nhận dữ liệu đầu vào 451 (ví dụ, hình ảnh thứ nhất) qua lớp đầu vào, và xử lý dữ liệu đầu vào 451 qua N lớp tích chập 452 để tạo ra dữ liệu đầu ra 453 (ví dụ, hình ảnh thứ hai), và cấp dữ liệu đầu ra 453 qua lớp đầu ra. Mạng thần kinh nhân tạo chính 200 có thể là mạng thần kinh nhân tạo tích chập sâu bao gồm hai hoặc nhiều lớp tích chập. Mạng thần kinh nhân tạo chính 200 có thể còn bao gồm các lớp gộp, các lớp được kết nối đầy đủ, và các lớp chuẩn hóa như các lớp ẩn.

Thiết bị xử lý hình ảnh 100 có thể trích xuất “các đặc điểm” như đường viền, đường kẻ và màu sắc từ hình ảnh thứ nhất bằng cách sử dụng mạng thần kinh nhân tạo chính 200. Mỗi trong số N lớp tích chập 452 có trong mạng thần kinh nhân tạo chính 200 có thể nhận dữ liệu, xử lý dữ liệu nhận được, và tạo ra dữ liệu đầu ra. Ví dụ, như được thể hiện trên Fig.3, thiết bị xử lý hình ảnh 100 có thể tạo ra ánh xạ đặc điểm thứ nhất bằng cách tích chập đầu vào hình ảnh (ví dụ, hình ảnh thứ nhất) vào lớp tích chập thứ nhất 210 với một hoặc nhiều nhân hoặc bộ lọc. Thiết bị xử lý hình ảnh 100 cũng có thể tạo ra ánh xạ đặc điểm thứ hai bằng cách nhập ánh xạ đặc điểm thứ nhất vào lớp tích chập thứ hai 220 và tích chập đầu vào ánh xạ đặc điểm thứ nhất vào lớp tích chập thứ hai 220 với một hoặc nhiều nhân hoặc bộ lọc. Hoạt động tích chập được thực hiện trong lớp tích chập sẽ được mô tả một cách chi tiết có dựa vào các hình vẽ từ Fig.4 đến Fig.6.

Các lớp tích chập trước của mạng thần kinh nhân tạo chính 200 có thể hoạt động để trích xuất các đặc điểm ở mức thấp, như các viền hoặc gradien, từ hình ảnh đầu vào. Theo hướng về phía các lớp tích chập sau, các lớp tích chập sau này có thể trích xuất các đặc điểm phức tạp dần (ví dụ, mắt, mũi, miệng và khuôn mặt).

Một hoặc nhiều lớp tích chập mà nhận và đưa ra ánh xạ đặc điểm trong mạng thần kinh nhân tạo chính 200 có thể là các lớp ẩn (ví dụ, các lớp tích chập ẩn). Trong mạng thần kinh nhân tạo chính 200, các hoạt động xử lý khác với hoạt động thực hiện tích chập bằng cách áp dụng một hoặc nhiều nhân cho ánh xạ đặc điểm có thể được thực hiện. Ví dụ, các hoạt động như hàm kích hoạt và gộp có thể được thực hiện. Thiết bị xử lý hình ảnh 100 có thể áp dụng hàm kích hoạt để thay đổi các giá trị của ánh xạ đặc điểm được trích xuất bằng cách thực hiện hoạt động tích chập thành các giá trị phi tuyến tính như “có hoặc không có các đặc điểm”. Các ví dụ về hàm kích hoạt có thể bao gồm hàm tuyến tính trong đó kích hoạt tỷ lệ với đầu vào (tức là tổng trọng số từ nơ-ron), hàm đơn vị tuyến tính lũy thừa (ELU - Exponential Linear Unit) gồm hai công thức, $f(x) = x$ nếu $x \geq 0$ và $f(x) = \alpha (e^x - 1)$ nếu $x < 0$, hàm đơn vị tuyến tính chỉnh lưu (ReLU - Rectified Linear Unit) được xác định là $y = \max(0, x)$, hàm Sigmoid nhận giá trị thực làm đầu vào và đưa ra giá trị khác từ 0 đến 1, hàm Tanh nén một số có giá trị thực trong phạm vi từ -1 đến 1 và hàm Softmax tính toán xác suất của sự kiện trên n sự kiện khác nhau. Thiết bị xử lý ảnh 100 có thể thực hiện lấy mẫu con (gộp) để giảm kích thước của ánh xạ đặc điểm được trích xuất. Trong trường hợp này, gộp tối đa, gộp trung bình, gộp định mức L2, hoặc các phương án tương tự có thể được sử dụng, nhưng các phương án của sáng chế không giới hạn ở đó.

Ví dụ, sau khi hình ảnh thứ nhất 301 (dữ liệu đầu vào) đi qua N lớp tích chập của mạng thần kinh nhân tạo chính 200, hình ảnh thứ nhất 301 có thể trải qua hoạt động, như hàm kích hoạt hoặc gộp, và vì vậy có thể được đưa ra làm hình ảnh thứ hai 302.

Tham khảo tới Fig.3, mạng thần kinh nhân tạo chính 200 có thể bao gồm N lớp tích chập. Trong trường hợp này, hình ảnh đầu vào có thể được tích chập với một hoặc nhiều nhân trong mỗi trong số các lớp tích chập thứ nhất đến thứ M 210, 220, ..., và 255 của mạng thần kinh nhân tạo chính 200, và vì vậy dữ liệu kết quả trung gian có thể được đưa ra. Dữ liệu kết quả trung gian có thể được nhập vào mạng thần

kinh nhân tạo phụ 300.

Mạng thần kinh nhân tạo phụ 300 có thể nhận dữ liệu kết quả trung gian từ mạng thần kinh nhân tạo chính 200. Ví dụ, mạng thần kinh nhân tạo phụ 300 bao gồm lớp đầu vào để nhận dữ liệu trung gian, một hoặc nhiều lớp tích chập để xử lý dữ liệu kết quả trung gian, và lớp đầu ra để đưa ra dữ liệu kết quả trung gian đã được xử lý làm dữ liệu đầu ra (ví dụ, hình ảnh thứ ba 330).

Một hoặc nhiều lớp tích chập có trong mạng thần kinh nhân tạo phụ 300 có thể là các lớp ẩn (ví dụ, các lớp tích chập ẩn). Mạng thần kinh nhân tạo phụ 300 có thể thực hiện các hoạt động xử lý khác với hoạt động thực hiện tích chập bằng cách áp dụng một hoặc nhiều nhân cho ánh xạ đặc điểm. Ví dụ, các hoạt động như hàm kích hoạt và gộp có thể được trích xuất. Thiết bị xử lý hình ảnh 100 có thể áp dụng hàm kích hoạt để thay đổi các giá trị của ánh xạ đặc điểm được trích xuất bằng cách thực hiện hoạt động tích chập vào các giá trị phi tuyến tính như “có hoặc không có các đặc tính”. Trong trường hợp này, hàm Tanh hoặc hàm ReLu có thể được sử dụng, nhưng các phương án thực hiện của sáng chế không giới hạn ở đó. Thiết bị xử lý hình ảnh 100 có thể thực hiện lấy mẫu con (gộp) để giảm kích cỡ của ánh xạ đặc điểm được trích xuất. Trong trường hợp này, gộp tối đa, gộp trung bình, gộp định mức L2, hoặc các phương pháp tương tự có thể được sử dụng, nhưng các phương án thực hiện của sáng chế không giới hạn ở đó.

Ví dụ, sau khi dữ liệu kết quả trung gian đi qua một hoặc nhiều lớp tích chập của mạng thần kinh nhân tạo chính 200, dữ liệu kết quả trung gian có thể trải qua hoạt động như hàm kích hoạt hoặc gộp, và vì vậy có thể được đưa ra làm hình ảnh thứ ba 330.

Fig.5 minh họa dữ liệu đầu vào, các nhân, và dữ liệu đầu ra của lớp tích chập thứ nhất có trong mạng thần kinh nhân tạo chính theo một phương án thực hiện của sáng chế.

Mạng thần kinh nhân tạo chính 200 theo một phương án thực hiện của sáng chế có thể bao gồm N lớp tích chập 452 như được thể hiện trên Fig.4, và tham khảo Fig.5, lớp tích chập thứ nhất trong số N lớp tích chập 452 lúc này sẽ được gọi là lớp tích chập thứ nhất Conv_1.

Dữ liệu đầu vào 510 (ví dụ, dữ liệu đầu vào thứ nhất) trong lớp tích chập thứ nhất Conv_1 có thể là hình ảnh thứ nhất, và kích cỡ của dữ liệu đầu vào 510 có thể là $w*h*C$ (tức là, chiều rộng*chiều cao*kênh), trong đó w, h, và C biểu thị chiều rộng, chiều cao và kênh, một cách tương ứng. Ví dụ, hình ảnh thứ nhất có thể có kích cỡ $w*h$, và số lượng các kênh C có thể là, nhưng không giới hạn ở, 3 (ví dụ, R, G, và B). Một hoặc nhiều nhân 530 (ví dụ, các nhân thứ nhất) trong lớp tích chập thứ nhất Conv_1 có thể có kích cỡ $Kw*Kh$, và số lượng các nhân có thể là $C*D$ (tức là, số lượng các kênh đầu vào*số lượng các kênh đầu ra), trong đó D biểu thị số lượng các kênh đầu ra.

Trong lớp tích chập thứ nhất Conv_1, khi hoạt động tích chập đối với dữ liệu đầu vào 510 và các nhân thứ nhất 530 được thực hiện, dữ liệu đầu ra thứ nhất 540 có thể được tạo ra. Cụ thể, dữ liệu đầu ra thứ nhất 540 có thể bao gồm một hoặc nhiều ánh xạ đặc điểm, kích cỡ của dữ liệu đầu ra thứ nhất 540 có thể là $W*H*D$, kích cỡ của một ánh xạ đặc điểm có thể là $W*H$, và số lượng các ánh xạ đặc điểm (hoặc số lượng các kênh) có thể là D.

Fig.6 là sơ đồ được sử dụng để mô tả quy trình trong đó lớp tích chập thứ nhất Conv_1 trên Fig.5 thực hiện hoạt động tích chập.

Để thuận lợi cho việc mô tả, giả định trên Fig.6 rằng dữ liệu đầu vào 610 của lớp tích chập thứ nhất Conv_1 có kích cỡ $5*5$ và số lượng các kênh là n. Cũng giả định rằng kích cỡ của nhân được dùng cho dữ liệu đầu vào 610 là $3*3*n$ và số lượng các nhân là D. Ở đây, n là độ sâu của nhân và vì vậy có giá trị giống như số lượng các kênh của dữ liệu đầu vào 610. Số lượng các kênh của dữ liệu đầu vào 610 có thể

được xác định theo số lượng các nhân được sử dụng bởi lớp trước lớp hiện thời. Nói cách khác, có thể coi rằng một nhân bao gồm n nhân con 631, 632, ..., và 639 từng nhân con có kích cỡ 3×3 , và n nhân con 631, 632, ..., và 639 có thể tương ứng với n kênh của dữ liệu đầu vào 610, một cách tương ứng.

Fig.6 minh họa quy trình trích xuất đặc điểm của dữ liệu đầu vào 610 bằng cách áp dụng nhân thứ nhất 630 bao gồm n nhân con 631, 632, ..., và 639 từ đầu bên trái phía trên của dữ liệu đầu vào 610 đến đầu bên phải phía dưới của nó. Ví dụ, hoạt động tích chập có thể được thực hiện bằng cách áp dụng nhân thứ nhất 630 ($3 \times 3 \times n$) vào các điểm ảnh có trong các vùng đầu bên trái phía trên ($3 \times 3 \times n$) 611, 612, ..., và 619 dữ liệu đầu vào 610. Nói cách khác, các giá trị điểm ảnh có trong các vùng đầu bên trái phía trên ($3 \times 3 \times n$) 611, 612, ..., và 619 có thể được nhân với các giá trị trọng số có trong nhân thứ nhất 630 và được cộng lại, nhờ đó tạo ra giá trị điểm ảnh duy nhất 641 được ánh xạ với các vùng đầu bên trái phía trên ($3 \times 3 \times n$) 611, 612, ..., và 619.

Ngoài ra, các giá trị điểm ảnh có trong các vùng $3 \times 3 \times n$ 621, 622, ..., và 629 đạt được bằng cách di chuyển các vùng đầu bên trái phía trên ($3 \times 3 \times n$) 611, 612, ..., và 619 bởi một điểm ảnh về phía bên phải có thể được nhân với các giá trị trọng số có trong nhân thứ nhất 630 và được cộng lại, nhờ đó tạo ra giá trị điểm ảnh duy nhất 642 được ánh xạ với các vùng $3 \times 3 \times n$ 621, 622, ..., và 629. Sử dụng cùng một phương pháp như mô tả ở trên, trong khi mục đích của hoạt động tích chập là được quét theo các đơn vị điểm ảnh theo hướng từ trái sang phải và từ trên xuống dưới trong dữ liệu đầu vào 610, các đích được nhân bởi các giá trị trọng số có trong nhân thứ nhất 630 được nhân và được cộng lại, và vì vậy các giá trị điểm ảnh được tạo ra. Vì vậy, ánh xạ đặc điểm 3×3 640 (dữ liệu đầu ra) có thể được đưa ra. Dữ liệu là đích của hoạt động tích chập có thể được quét trong khi di chuyển một điểm ảnh tại một thời điểm, nhưng có thể được quét trong khi di chuyển hai hoặc nhiều điểm ảnh tại một thời điểm. Số lượng các điểm ảnh nhờ đó dữ liệu đầu vào di chuyển trong quá trình quét

được xem là một bước nhảy và kích cỡ của ánh xạ đặc điểm mà được đưa ra có thể được xác định theo kích cỡ của bước nhảy này.

Tham khảo tới Fig.6, dữ liệu đầu vào 610 có kích cỡ 5×5 , nhưng dữ liệu đầu ra 640 có kích cỡ 3×3 nhỏ hơn kích cỡ của dữ liệu đầu vào 610. Mạng thần kinh nhân tạo chính bao gồm một vài lớp tích chập, và, trong khi đi qua một vài lớp tích chập, kích cỡ của dữ liệu giảm liên tục. Khi kích cỡ của dữ liệu giảm trước khi đặc điểm được trích xuất một cách đầy đủ, đặc điểm của dữ liệu đầu vào có thể bị mất, và để ngăn chặn điều này, việc đệm có thể được thực hiện. Đệm có nghĩa là tăng kích cỡ của dữ liệu đầu vào bằng cách đưa ra giá trị cụ thể (ví dụ, '0') cho viền của dữ liệu đầu ra để ngăn ngừa dữ liệu đầu ra có kích cỡ giảm. Tuy nhiên, các phương án thực hiện của sáng chế không giới hạn ở đó.

Mặc dù Fig.6 chỉ minh họa kết quả của hoạt động tích chập đối với nhân thứ nhất 630, khi hoạt động tích chập được thực hiện trên D nhân, ánh xạ đặc điểm $3 \times 3 \times D$ có thể được đưa ra. Nói cách khác, số lượng các kênh của dữ liệu đầu ra được xác định theo số D của nhân, và do đó, số lượng các kênh của dữ liệu đầu vào trong lớp tiếp theo có thể được xác định.

Fig.7 là sơ đồ được sử dụng để mô tả phương pháp trong đó thiết bị xử lý hình ảnh hoạt động theo một phương án thực hiện của sáng chế.

Tham khảo tới Fig.7, thiết bị xử lý hình ảnh 100 có thể tạo ra hình ảnh thứ hai 720 từ hình ảnh đầu vào thứ nhất 710 bằng cách sử dụng mạng thần kinh nhân tạo chính 200 được huấn luyện cho mục đích chuyển kiểu. Mạng thần kinh nhân tạo chính 200 có thể là CNN bao gồm N lớp tích chập, và hình ảnh thứ hai 720 có thể là hình ảnh thu được bằng cách áp dụng kiểu định trước 705 cho hình ảnh thứ nhất 710.

Ví dụ, thiết bị xử lý hình ảnh 100 có thể đưa vào hình ảnh thứ nhất 710 vào mạng thần kinh nhân tạo chính 200. Hình ảnh đầu vào thứ nhất 710 có thể bao gồm ba kênh (ví dụ, các kênh R, G, và B), nhưng các phương án thực hiện của sáng chế

không giới hạn ở đó.

Thiết bị xử lý hình ảnh 100 có thể thực hiện hoạt động tích chập dữ liệu hình ảnh đầu vào với một hoặc nhiều nhân trong lớp tích chập thứ nhất 210. Dữ liệu thu được qua tích chập có thể được đưa ra bởi lớp tích chập thứ nhất 210 và đưa vào lớp tích chập thứ hai 220. Theo cùng phương pháp được mô tả ở trên, thiết bị xử lý hình ảnh 100 có thể nhận dữ liệu từ lớp tích chập trước và thực hiện hoạt động tích chập dữ liệu nhận được với một hoặc nhiều nhân hoặc bộ lọc, trong mỗi trong số N lớp tích chập, nhờ đó vào dữ liệu đầu ra thu được qua phép tích chập cho lớp tích chập tiếp theo. Thiết bị xử lý hình ảnh 100 có thể thực hiện hoạt động của hàm kích hoạt trên dữ liệu đã đi qua N lớp tích chập và có thể tạo ra hình ảnh đầu ra cuối (ví dụ, hình ảnh thứ hai 720), dựa trên dữ liệu thu được qua hoạt động của hàm kích hoạt. Hình ảnh đầu ra cuối có thể là hình ảnh bao gồm ba kênh (kênh R, G, và B), nhưng các phương án thực hiện của sáng chế không giới hạn ở đó.

Thiết bị xử lý hình ảnh 100 có thể thu được dữ liệu kết quả trung gian trong khi xử lý hình ảnh thứ nhất 710, bằng cách sử dụng mạng thần kinh nhân tạo chính 200. Thiết bị xử lý hình ảnh 100 có thể thu được dữ liệu đã đi qua lớp tích chập thứ n trong số các lớp tích chập có trong mạng thần kinh nhân tạo chính 200.

Ví dụ, như được thể hiện trên Fig.7, hình ảnh thứ nhất 710 có thể được nhập vào mạng thần kinh nhân tạo chính 200 và, trong khi hình ảnh thứ nhất 710 đang đi qua các lớp tích chập thứ nhất và thứ hai 210 và 220, dữ liệu được tích chập với một hoặc nhiều nhân hoặc bộ lọc có thể được đưa ra làm dữ liệu kết quả trung gian thứ nhất D1. Hình ảnh thứ nhất 710 có thể được nhập vào mạng thần kinh nhân tạo chính 200, và dữ liệu thu được sau khi hình ảnh thứ nhất 710 đi qua các lớp tích chập thứ nhất đến thứ năm 210, 220, 230, 240, và 250 có thể được đưa ra làm dữ liệu kết quả trung gian thứ hai D2. Hình ảnh thứ nhất 710 có thể được nhập vào mạng thần kinh nhân tạo chính 200, và dữ liệu thu được sau khi hình ảnh thứ nhất 710 đi qua các lớp

tích chập thứ nhất đến thứ tám 210, 220, 230, 240, 250, 260, 270, và 280 có thể được đưa ra làm dữ liệu kết quả trung gian thứ ba D3. Hình ảnh thứ nhất 710 có thể được nhập vào mạng thần kinh nhân tạo chính 200, và dữ liệu thu được sau khi hình ảnh thứ nhất 710 đi qua các lớp tích chập thứ nhất đến thứ mười một 210, 220, 230, 240, 250, 260, 270, 280, 290, 291 và 292 có thể được đưa ra làm dữ liệu kết quả trung gian thứ tư D4. Tuy nhiên, các phương án thực hiện của sáng chế không giới hạn ở đó.

Thiết bị xử lý hình ảnh 100 có thể tạo ra các hình ảnh đầu ra trung gian khác với hình ảnh đầu ra cuối 720 (tức là, hình ảnh thứ hai 720), bằng cách xử lý dữ liệu kết quả trung gian thu được bằng cách sử dụng mạng thần kinh nhân tạo phụ.

Ví dụ, thiết bị xử lý hình ảnh 100 có thể bao gồm các mạng thần kinh nhân tạo phụ thứ nhất đến thứ tư 310, 320, 330 và 340. Mạng thần kinh nhân tạo chính 200 có thể bao gồm nhiều lớp đầu ra trung gian đưa ra dữ liệu kết quả trung gian đến các mạng thần kinh nhân tạo phụ 310, 320, 330 và 340. Cụ thể, mạng thần kinh nhân tạo chính 200 có thể bao gồm lớp đầu ra trung gian thứ nhất IO1 nằm giữa lớp tích chập thứ hai 220 và lớp tích chập thứ ba 230 để chuyển dữ liệu kết quả trung gian thứ nhất D1 đến mạng thần kinh nhân tạo phụ thứ nhất 310, lớp đầu ra trung gian thứ hai IO2 nằm giữa lớp tích chập thứ năm 250 và lớp tích chập thứ sáu 260 để chuyển dữ liệu kết quả trung gian thứ hai D2 đến mạng thần kinh nhân tạo phụ thứ hai 320, lớp đầu ra trung gian thứ ba IO3 nằm giữa lớp tích chập thứ tám 280 và lớp tích chập thứ chín 290 để chuyển dữ liệu kết quả trung gian thứ ba D3 đến mạng thần kinh nhân tạo phụ thứ ba 330, và lớp đầu ra trung gian thứ tư IO4 nằm giữa lớp tích chập thứ mười một 292 và lớp tích chập thứ mười hai 293 để chuyển dữ liệu kết quả trung gian thứ tư D4 đến mạng thần kinh nhân tạo phụ thứ tư 340.

Thiết bị xử lý hình ảnh 100 có thể đưa dữ liệu kết quả trung gian thứ nhất thu được D1 vào mạng thần kinh nhân tạo phụ thứ nhất 310. Mạng thần kinh nhân tạo phụ thứ nhất 310 có thể bao gồm một hoặc nhiều lớp tích chập, và một hoặc nhiều

lớp tích chập có thể thực hiện hoạt động tích chập dữ liệu kết quả trung gian thứ nhất đầu vào D1 với một hoặc nhiều nhân hoặc bộ lọc. Thiết bị xử lý hình ảnh 100 có thể thực hiện hoạt động của hàm kích hoạt trên dữ liệu tích chập và có thể tạo ra hình ảnh đầu ra trung gian thứ nhất S1, dựa trên dữ liệu thu được qua hoạt động của hàm kích hoạt. Hình ảnh đầu ra trung gian thứ nhất S1 có thể là hình ảnh bao gồm ba kênh (ví dụ, kênh R, G, và B), nhưng các phương án thực hiện của sáng chế không giới hạn ở đó.

Thiết bị xử lý hình ảnh 100 có thể tạo ra các hình ảnh đầu ra trung gian thứ hai đến thứ tư S2, S3 và S4 bằng cách xử lý dữ liệu kết quả trung gian thứ hai đến thứ tư D2, D3 và D4 bằng cách sử dụng các mạng thần kinh nhân tạo phụ thứ hai đến thứ tư 320, 330 và 340, một cách tương ứng. Các hình ảnh đầu ra trung gian thứ hai đến thứ tư S2, S3 và S4 được tạo ra theo phương pháp tạo ra hình ảnh đầu ra trung gian thứ nhất S1 bằng cách xử lý dữ liệu kết quả trung gian thứ nhất D1 bằng cách sử dụng mạng thần kinh nhân tạo phụ thứ nhất 310, và vì vậy phần mô tả chi tiết của chúng sẽ được bỏ qua.

Mức độ mà kiểu được áp dụng để thu được mỗi trong số các hình ảnh đầu ra trung gian thứ nhất đến thứ tư S1, S2, S3 và S4 có thể nhỏ hơn so với kiểu cho hình ảnh thứ hai 720. Ví dụ, theo chiều từ hình ảnh đầu ra trung gian thứ nhất S1 đến hình ảnh đầu ra trung gian thứ tư S4, mức độ mà kiểu định trước 705 được áp dụng cho hình ảnh thứ nhất 710 có thể tăng lên. Mức độ mà kiểu định trước 705 được áp dụng để thu được hình ảnh thứ hai 720 (tức là, hình ảnh đầu ra cuối) có thể lớn hơn so với mức độ mà kiểu định trước 705 được áp dụng để thu được hình ảnh đầu ra trung gian thứ tư S4.

Fig.8 và Fig.9 là các hình vẽ minh họa một ví dụ trong đó thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế hiển thị các hình ảnh đầu ra trung gian.

Tham khảo tới Fig.8, khoảng thời gian thứ nhất T1 có thể được thực hiện khi thiết bị xử lý hình ảnh 100 tạo ra hình ảnh đầu ra trung gian thứ nhất S1 từ hình ảnh thứ nhất 710 bằng cách sử dụng mạng thần kinh nhân tạo chính 200 và mạng thần kinh nhân tạo phụ thứ nhất 310. Thiết bị xử lý hình ảnh 100 có thể hiển thị hình ảnh đầu ra trung gian thứ nhất đã được tạo ra S1 trên màn hình. Khoảng thời gian thứ hai T1+T2 có thể được thực hiện khi thiết bị xử lý hình ảnh 100 tạo ra hình ảnh đầu ra trung gian thứ hai S2 từ hình ảnh thứ nhất 710 bằng cách sử dụng mạng thần kinh nhân tạo chính 200 và mạng thần kinh nhân tạo phụ thứ hai 320, và khoảng thời gian thứ ba T1+T2+T3 có thể được thực hiện khi thiết bị xử lý hình ảnh 100 tạo ra hình ảnh đầu ra trung gian thứ ba S3 từ hình ảnh thứ nhất 710 bằng cách sử dụng mạng thần kinh nhân tạo chính 200 và mạng thần kinh nhân tạo phụ thứ ba 330.

Khi thiết bị xử lý hình ảnh 100 tạo ra hình ảnh đầu ra cuối 720 (tức là hình ảnh thứ hai) từ hình ảnh thứ nhất 710 bằng cách sử dụng mạng thần kinh nhân tạo chính 200, khoảng thời gian thứ tư T4 có thể được thực hiện. Khoảng thời gian thứ tư T4 có thể dài hơn khoảng thời gian thứ ba T1+T2+T3.

Trước khi thiết bị xử lý hình ảnh 100 tạo ra hình ảnh đầu ra cuối 720 bằng cách xử lý hình ảnh thứ nhất 710 bằng cách sử dụng mạng thần kinh nhân tạo chính 200, thiết bị xử lý hình ảnh 100 có thể tạo ra các hình ảnh đầu ra trung gian thứ nhất đến thứ ba S1, S2 và S3 và hiển thị chúng trên màn hình.

Thiết bị xử lý hình ảnh 100 có thể hiển thị các mẫu thông tin thời gian 810 và 820 từng mẫu biểu thị khoảng thời gian cần được thực hiện cho tới khi hình ảnh đầu ra tiếp theo sẽ được tạo ra, cùng với các hình ảnh đầu ra. Ví dụ, như được minh họa trên Fig.8, trong khi hiển thị hình ảnh đầu ra trung gian thứ nhất S1, thiết bị xử lý hình ảnh 100 cũng có thể hiển thị thông tin thời gian 810 biểu thị khoảng thời gian T2 được thực hiện từ thời điểm hiện tại đến thời điểm khi hình ảnh đầu ra trung gian thứ hai S2 sẽ được tạo ra. Trong khi hiển thị hình ảnh đầu ra trung gian thứ hai S2,

thiết bị xử lý hình ảnh 100 cũng có thể hiển thị thông tin thời gian 820 biểu thị khoảng thời gian T3 được thực hiện cho tới khi hình ảnh đầu ra trung gian thứ ba S3 sẽ được tạo ra.

Thiết bị xử lý hình ảnh 100 có thể hiển thị không chỉ thông tin thời gian biểu thị khoảng thời gian được thực hiện cho tới khi hình ảnh đầu ra trung gian kế tiếp sẽ được tạo ra, mà còn thông tin thời gian biểu thị khoảng thời gian được thực hiện cho tới khi hình ảnh đầu ra cuối sẽ được tạo ra.

Do đó, người dùng có thể kiểm tra các hình ảnh đầu ra trung gian được hiển thị trên màn hình trước khi hình ảnh đầu ra cuối được tạo ra và có thể kiểm tra khoảng thời gian được thực hiện cho tới khi hình ảnh đầu ra tiếp theo sẽ được tạo ra, nhờ đó xác định liệu rằng có tiếp tục thực hiện việc xử lý hình ảnh hay không.

Thiết bị xử lý hình ảnh 100 có thể ngắt quá trình xử lý hình ảnh đối với hình ảnh thứ nhất 710 trong mạng thần kinh nhân tạo chính 200, dựa vào đầu vào của người dùng. Ví dụ, khi hình ảnh đầu ra trung gian thứ ba S3 được hiển thị và ở thời điểm yêu cầu ngắt quá trình xử lý hình ảnh được tiếp nhận, thiết bị xử lý hình ảnh 100 có thể ngắt quá trình xử lý hình ảnh đối với hình ảnh thứ nhất 710 ở mạng thần kinh nhân tạo chính 200. Do đó, thiết bị xử lý hình ảnh 100 có thể không tạo ra các hình ảnh đầu ra trung gian kế tiếp và hình ảnh đầu ra cuối 720.

Tham khảo tới Fig.9, thiết bị xử lý hình ảnh 100 có thể thực hiện quá trình xử lý hình ảnh bổ sung trên hình ảnh đầu ra trung gian, dựa vào đầu vào của người dùng. Ví dụ, thiết bị xử lý hình ảnh 100 có thể tạo ra hình ảnh cuối 920 bằng cách tiến hành xử lý hình ảnh (làm sắc nét) điều chỉnh cảm nhận màu sắc, độ sáng, độ trong và dạng tương tự của hình ảnh đầu ra trung gian thứ ba S3. Tuy nhiên, các phương án thực hiện của sáng chế không giới hạn ở đó.

Fig.10 minh họa phương pháp, được thực hiện bởi thiết bị xử lý hình ảnh 100, xác định xem liệu có truy cập vào mạng thần kinh nhân tạo phụ hay không.

Tham khảo tới Fig.10, thiết bị xử lý hình ảnh 100 có thể xác định, làm các điểm mạng phụ, các điểm trong mạng thần kinh nhân tạo chính 200 nơi thu được các mảnh trung gian. Ví dụ, thiết bị xử lý hình ảnh 100 có thể xác định, làm điểm mạng phụ thứ nhất 1010, điểm đầu ra của lớp tích chập thứ hai nơi thu được dữ liệu kết quả trung gian thứ nhất D1 (điểm đầu vào của lớp tích chập thứ ba), xác định, làm điểm mạng phụ thứ hai 1020, điểm đầu ra của lớp tích chập thứ năm nơi thu được dữ liệu kết quả trung gian thứ hai D2 (điểm đầu vào của lớp tích chập thứ sáu), xác định, làm điểm mạng phụ thứ ba 1030, điểm đầu ra của lớp tích chập thứ tám nơi thu được dữ liệu kết quả trung gian thứ ba D3 (điểm đầu vào của lớp tích chập thứ chín), và xác định, làm điểm mạng phụ thứ tư 1040, điểm đầu ra của lớp tích chập thứ mười một nơi thu được dữ liệu kết quả trung gian thứ tư D4 (điểm đầu vào của lớp tích chập thứ mười hai). Tuy nhiên, các phương án thực hiện của sáng chế không giới hạn ở đó.

Thiết bị xử lý hình ảnh 100 có thể xác định liệu có thực hiện việc xử lý hình ảnh bằng cách sử dụng mạng thần kinh nhân tạo phụ, ở các điểm mạng phụ thứ nhất đến thứ tư 1010, 1020, 1030 và 1040. Ví dụ, khi hoạt động tích chập được hoàn tất trong các lớp tích chập thứ nhất và thứ hai và vì vậy dữ liệu kết quả trung gian thứ nhất D1 được đưa ra (ở điểm mạng phụ thứ nhất 1010), thiết bị xử lý hình ảnh 100 có thể đưa dữ liệu kết quả trung gian thứ nhất D1 vào mạng thần kinh nhân tạo phụ thứ nhất 310 và thực hiện việc xử lý hình ảnh, nhờ đó xác định xem liệu có tạo ra hình ảnh đầu ra trung gian thứ nhất S1 hay không.

Thiết bị xử lý hình ảnh 100 có thể tính toán và ước tính khoảng thời gian (ví dụ, khoảng thời gian thứ nhất) được thực hiện để xử lý dữ liệu kết quả trung gian thứ nhất D1 thành hình ảnh đầu ra trung gian thứ nhất S1 bằng cách sử dụng mạng thần kinh nhân tạo phụ thứ nhất 310, và khoảng thời gian (ví dụ, khoảng thời gian thứ hai) được thực hiện để xử lý dữ liệu kết quả trung gian thứ nhất D1 thành hình ảnh đầu ra cuối 720 bằng cách sử dụng mạng thần kinh nhân tạo chính 200. Khi sự chênh lệch giữa các khoảng thời gian thứ nhất và thứ hai bằng hoặc lớn hơn một giá trị định

trước, thiết bị xử lý hình ảnh 100 có thể đi vào mạng thần kinh nhân tạo phụ thứ nhất 310. Ví dụ, thiết bị xử lý hình ảnh 100 có thể tạo ra hình ảnh đầu ra trung gian thứ nhất S1 bằng cách sử dụng dữ liệu kết quả trung gian thứ nhất D1 và mạng thần kinh nhân tạo phụ thứ nhất 310. Mặt khác, khi sự chênh lệch giữa các khoảng thời gian thứ nhất và thứ hai nhỏ hơn giá trị định trước, thiết bị xử lý hình ảnh 100 có thể không đi vào mạng thần kinh nhân tạo phụ thứ nhất 310. Ví dụ, không tạo ra hình ảnh đầu ra trung gian thứ nhất S1, thiết bị xử lý hình ảnh 100 có thể tạo ra hình ảnh đầu ra cuối 720 bằng cách sử dụng dữ liệu kết quả trung gian thứ nhất D1 và mạng thần kinh nhân tạo chính 200.

Tương tự, thiết bị xử lý hình ảnh 100 có thể tính toán khoảng thời gian (ví dụ, khoảng thời gian thứ ba) được thực hiện để xử lý dữ liệu kết quả trung gian thứ hai D2 thành hình ảnh đầu ra trung gian thứ hai S2 bằng cách sử dụng mạng thần kinh nhân tạo phụ thứ hai 320, và khoảng thời gian (ví dụ, khoảng thời gian thứ tư) được thực hiện để xử lý dữ liệu kết quả trung gian thứ hai D2 thành hình ảnh đầu ra cuối 720 bằng cách sử dụng mạng thần kinh nhân tạo chính 200. Thiết bị xử lý hình ảnh 100 có thể xác định liệu có đi vào mạng thần kinh nhân tạo phụ thứ hai 320 hay không, bằng cách so sánh các khoảng thời gian thứ ba và thứ tư với nhau ở điểm mạng phụ thứ hai 1020.

Thiết bị xử lý hình ảnh 100 cũng có thể xác định xem liệu có đi vào mạng thần kinh nhân tạo phụ thứ ba 330 hay không, bằng cách so sánh, ở điểm mạng phụ thứ ba 1030, khoảng thời gian (ví dụ, khoảng thời gian thứ năm) được thực hiện để xử lý dữ liệu kết quả trung gian thứ ba D3 thành hình ảnh đầu ra trung gian thứ ba S3 bằng cách sử dụng mạng thần kinh nhân tạo phụ thứ ba 330 với khoảng thời gian (ví dụ, khoảng thời gian thứ sáu) được thực hiện để xử lý dữ liệu kết quả trung gian thứ ba D3 thành hình ảnh đầu ra cuối 720 bằng cách sử dụng mạng thần kinh nhân tạo chính 200.

Thiết bị xử lý hình ảnh 100 cũng có thể xác định xem liệu có đi vào mạng thần kinh nhân tạo phụ thứ tư 340 hay không, bằng cách so sánh, ở điểm mạng phụ thứ tư 1040, khoảng thời gian (ví dụ, khoảng thời gian thứ bảy) được thực hiện để xử lý dữ liệu kết quả trung gian thứ tư D4 thành hình ảnh đầu ra trung gian thứ tư S4 bằng cách sử dụng mạng thần kinh nhân tạo phụ thứ tư 340 với khoảng thời gian (ví dụ, khoảng thời gian thứ tám) được thực hiện để xử lý dữ liệu kết quả trung gian thứ tư D4 thành hình ảnh đầu ra cuối 720 bằng cách sử dụng mạng thần kinh nhân tạo chính 200.

Fig.11 là sơ đồ được sử dụng để giải thích phương pháp, được thực hiện bởi thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế, tạo ra các hình ảnh đầu ra bằng cách sử dụng mạng thần kinh nhân tạo chính và các mạng thần kinh nhân tạo phụ.

Mạng thần kinh nhân tạo chính 1100 trên Fig.11 có thể tương ứng với mạng thần kinh nhân tạo chính 200 trên Fig.7, mạng thần kinh nhân tạo phụ thứ nhất 1110 trên Fig.11 có thể tương ứng với mạng thần kinh nhân tạo phụ thứ nhất 310 trên Fig.7, và mạng thần kinh nhân tạo phụ thứ hai 1120 trên Fig.11 có thể tương ứng với mạng thần kinh nhân tạo phụ thứ hai 320 trên Fig.7. Do đó, các phần mô tả chi tiết về các cấu trúc và hoạt động của mạng thần kinh nhân tạo chính 1100 và các mạng thần kinh nhân tạo phụ thứ nhất và thứ hai 1110 và 1120 sẽ được bỏ qua.

Tham khảo tới Fig.11, thiết bị xử lý hình ảnh 100 có thể tạo ra hình ảnh thứ hai 720 (tức là, hình ảnh đầu ra cuối) từ hình ảnh thứ nhất 710 bằng cách sử dụng mạng thần kinh nhân tạo chính 1100.

Thiết bị xử lý hình ảnh 100 có thể thu được dữ liệu kết quả trung gian thứ nhất D1 bằng cách xử lý hình ảnh thứ nhất 710 bằng cách sử dụng mạng thần kinh nhân tạo chính 1100 và có thể tạo ra hình ảnh đầu ra trung gian thứ nhất S1 bằng cách sử dụng dữ liệu kết quả trung gian thứ nhất D1 và mạng thần kinh nhân tạo phụ thứ nhất

1110. Thiết bị xử lý hình ảnh 100 cũng có thể thu được dữ liệu kết quả trung gian thứ hai D2 bằng cách xử lý hình ảnh thứ nhất 710 bằng cách sử dụng mạng thần kinh nhân tạo chính 1100 và có thể tạo ra hình ảnh đầu ra trung gian thứ hai S2 bằng cách sử dụng dữ liệu kết quả trung gian thứ hai D2 và mạng thần kinh nhân tạo phụ thứ hai 1120.

Hình ảnh thứ nhất 710 và hình ảnh thứ hai 720 (tức là, hình ảnh đầu ra cuối) có thể là các hình ảnh có độ phân giải cao, và hình ảnh đầu ra trung gian thứ nhất S1 và hình ảnh đầu ra trung gian thứ hai S2 có thể là các hình ảnh có độ phân giải thấp.

Fig.12 là sơ đồ thể hiện một ví dụ nơi thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế tương tác với máy chủ.

Tham khảo tới Fig.12, máy chủ 2000 có thể bao gồm mạng thần kinh nhân tạo chính 2030, và mạng thần kinh nhân tạo chính 2030 có thể là mô hình được huấn luyện để xử lý hình ảnh theo mục đích định trước. Ví dụ, mạng thần kinh nhân tạo chính 2030 có thể là mô hình được huấn luyện nhằm mục đích chuyển kiểu.

Máy chủ 2000 có thể tạo ra hình ảnh thứ hai 720 (tức là, hình ảnh cuối, hình ảnh được chuyển kiểu) bằng cách áp dụng kiểu định trước cho hình ảnh đầu vào thứ nhất 710 bằng cách sử dụng mạng thần kinh nhân tạo chính 2030.

Do mạng thần kinh nhân tạo chính 2030 trên Fig.12 có thể tương ứng với mạng thần kinh nhân tạo chính 200 trên Fig.3, các phần mô tả chi tiết về cấu trúc và hoạt động của mạng thần kinh nhân tạo chính 2030 sẽ được bỏ qua ở đây. Máy chủ 2000 có thể truyền hình ảnh thứ hai 720 đến thiết bị xử lý hình ảnh 100.

Máy chủ 2000 có thể thu được dữ liệu kết quả trung gian còn hình ảnh thứ nhất 710 đang được xử lý trong mạng thần kinh nhân tạo chính 2030. Ví dụ, dữ liệu kết quả trung gian có thể bao gồm ảnh xạ đặc điểm (ví dụ, thông tin đặc điểm) được trích xuất từ lớp cụ thể trong số nhiều lớp tích chập có trong mạng thần kinh nhân tạo chính 2030.

Máy chủ 2000 có thể truyền dữ liệu kết quả trung gian đến thiết bị xử lý hình ảnh 100.

Thiết bị xử lý hình ảnh 100 có thể bao gồm mạng thần kinh nhân tạo phụ 2050. Mạng thần kinh nhân tạo phụ 2050 có thể bao gồm một hoặc nhiều lớp tích chập, và hoạt động tích chập có thể được thực hiện trong một hoặc nhiều lớp tích chập. Ví dụ, thiết bị xử lý hình ảnh 100 có thể đưa vào dữ liệu kết quả trung gian được tiếp nhận từ máy chủ 2000 đến mạng thần kinh nhân tạo phụ 2050, và dữ liệu kết quả trung gian có thể được tích chập với một hoặc nhiều nhân trong khi đi qua một hoặc nhiều lớp tích chập. Thiết bị xử lý hình ảnh 100 có thể thực hiện hoạt động của hàm kích hoạt trên dữ liệu tích chập và có thể tạo ra hình ảnh thứ ba 1250 (ví dụ, hình ảnh đầu ra trung gian), dựa trên dữ liệu thu được qua hoạt động của hàm kích hoạt.

Thiết bị xử lý hình ảnh 100 có thể hiển thị hình ảnh thứ hai 720 được tiếp nhận từ máy chủ 2000 và tạo ra hình ảnh thứ ba 1250 trên màn hình. Mức độ mà kiểu định trước được áp dụng cho hình ảnh thứ ba 1250 (ví dụ, hình ảnh đầu ra trung gian) có thể nhỏ hơn so với mức độ mà kiểu định trước được áp dụng cho hình ảnh thứ hai 720 (ví dụ, hình ảnh đầu ra cuối).

Như được minh họa trên Fig.12, hoạt động của mạng thần kinh nhân tạo chính 2030 được thực hiện trong máy chủ 2000 và hoạt động của mạng thần kinh nhân tạo phụ 2050 được thực hiện trong thiết bị xử lý hình ảnh 100, và vì vậy băng thông bộ nhớ của thiết bị xử lý hình ảnh 100 và khoảng thời gian được thực hiện để thu được hình ảnh đầu ra có thể được giảm xuống.

Fig.13 là sơ đồ khối được sử dụng để giải thích phương pháp huấn luyện mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ, theo một phương án thực hiện của sáng chế.

Tham khảo tới Fig.13, thiết bị xử lý hình ảnh 100 có thể bao gồm hình ảnh bộ trích xuất đặc điểm 1310, bộ phân tích kiểu 1320, bộ phân tích nội dung 1330, và bộ

tinh chỉnh mạng 1340. Bộ trích xuất đặc điểm hình ảnh 1310 có thể nhận hình ảnh Y , hình ảnh đích kiểu Y_s , và hình ảnh đích nội dung Y_c từ mạng thần kinh nhân tạo chính 200. Hình ảnh đích kiểu Y_s có thể là hình ảnh bao gồm kiểu mà cần được chuyển, bằng cách sử dụng mạng thần kinh nhân tạo chính 200. Hình ảnh đích nội dung Y_c , làm hình ảnh gốc trước khi kiểu được chuyển, có thể là hình ảnh giống với hình ảnh X mà được đưa vào mạng thần kinh nhân tạo chính 200.

Bộ trích xuất đặc điểm hình ảnh 1310 có thể bao gồm CNN và có thể trích xuất các ánh xạ đặc điểm (thông tin đặc điểm) của hình ảnh đích kiểu Y_s nhận được, hình ảnh Y nhận được, và hình ảnh đích nội dung Y_c nhận được bằng cách sử dụng CNN. CNN có thể bao gồm nhiều lớp tích chập, và hoạt động tích chập có thể được thực hiện trong nhiều lớp tích chập. Bộ trích xuất đặc điểm hình ảnh 1310 có thể thực hiện các hoạt động, như hàm kích hoạt và gộp, ngoài hoạt động tích chập trong nhiều lớp tích chập, và có thể trích xuất ánh xạ đặc điểm trong mỗi trong số các lớp tích chập. Ví dụ, bộ trích xuất đặc điểm hình ảnh 1310 có thể trích xuất ánh xạ đặc điểm của hình ảnh đích kiểu Y_s , ánh xạ đặc điểm của hình ảnh đích nội dung Y_c , và ánh xạ đặc điểm của hình ảnh Y trong mỗi trong số các lớp tích chập có trong bộ trích xuất đặc điểm hình ảnh 1310.

Bộ trích xuất đặc điểm hình ảnh 1310 có thể trích xuất các đặc điểm ở mức thấp, như các viền hoặc các gradien của hình ảnh đầu vào trong các lớp tích chập trước, và có thể trích xuất các ánh xạ đặc điểm phức tạp dần lên ở các lớp tích chập sau so với các lớp tích chập trước. Tuy nhiên, các phương án thực hiện của sáng chế không giới hạn ở đó.

Bộ trích xuất đặc điểm hình ảnh 1310 có thể trích xuất các đặc điểm ở mức thấp, như các viền hoặc các gradien của hình ảnh đầu vào trong các lớp tích chập trước đó, và có thể trích xuất các ánh xạ đặc điểm phức tạp dần lên trong các lớp tích chập sau so với các lớp tích chập trước đó. Tuy nhiên, các phương án thực hiện của

sáng chế không giới hạn ở đó.

Bộ phân tích kiểu 1320 có thể phân tích sự khác nhau giữa các kiểu của hình ảnh đích kiểu Ys và hình ảnh Y. Ví dụ, bộ phân tích kiểu 1320 có thể phân tích sự khác nhau giữa ánh xạ đặc điểm của hình ảnh đích kiểu Ys và ánh xạ đặc điểm của hình ảnh Y.

Bộ phân tích nội dung 1330 có thể phân tích sự khác nhau giữa nội dung của hình ảnh đích nội dung Yc và nội dung của hình ảnh Y. Ví dụ, bộ phân tích nội dung 1330 có thể phân tích sự khác nhau giữa ánh xạ đặc điểm của hình ảnh đích nội dung Yc và ánh xạ đặc điểm của hình ảnh Y.

Bộ tinh chỉnh mạng 1340 có thể điều chỉnh các trọng số của các nhân của mạng thần kinh nhân tạo chính 200 sao cho sự khác nhau giữa ánh xạ đặc điểm về hình ảnh đích kiểu và ánh xạ đặc điểm của hình ảnh đầu ra giảm xuống. Bộ tinh chỉnh mạng 1340 cũng có thể điều chỉnh các trọng số của các nhân của mạng thần kinh nhân tạo chính 200 sao cho sự khác nhau giữa ánh xạ đặc điểm về hình ảnh đích nội dung và ánh xạ đặc điểm của hình ảnh đầu ra giảm xuống. Tuy nhiên, các phương án thực hiện của sáng chế không giới hạn ở đó.

Mặc dù Fig.13 minh họa và mô tả chỉ phương pháp huấn luyện mạng thần kinh nhân tạo chính 200 theo một phương án thực hiện của sáng chế, mạng thần kinh nhân tạo phụ theo một phương án thực hiện của sáng chế có thể được huấn luyện bằng cách sử dụng phương pháp giống như được mô tả ở trên.

Cụ thể, thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế có thể bao gồm các mạng thần kinh nhân tạo đưa ra các kết quả khác nhau mặc dù xử lý một phần các hình ảnh theo cách thức tương tự, làm mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ. Ví dụ, một trong số mạng thần kinh nhân tạo thứ nhất và mạng thần kinh nhân tạo thứ hai có thể làm mạng thần kinh nhân tạo chính, và chỉ một phần của mạng thần kinh nhân tạo còn lại có xử lý khác với mạng thần

kinh nhân tạo chính có thể làm mạng thần kinh nhân tạo phụ. Tuy nhiên, các phương án thực hiện của sáng chế không giới hạn ở đó. Do đó, thiết bị xử lý hình ảnh có thể làm tăng hiệu suất huấn luyện mạng thần kinh nhân tạo bằng cách không lặp lại việc huấn luyện đối với xử lý tương tự.

Fig.14 là sơ đồ khối của cấu trúc của thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế.

Tham khảo tới Fig.14, thiết bị xử lý hình ảnh 100 theo một phương án thực hiện của sáng chế có thể bao gồm bộ xử lý 120 và bộ nhớ 130.

Bộ xử lý 120 có thể thực hiện một hoặc nhiều chương trình được lưu trữ trong bộ nhớ 130. Bộ xử lý 120 có thể bao gồm nhân đơn, nhân kép, ba nhân, bốn nhân, hoặc đa nhân của nó. Bộ xử lý 120 có thể bao gồm nhiều bộ xử lý. Ví dụ, bộ xử lý 120 có thể được thực hiện bằng cách sử dụng bộ xử lý chính và bộ xử lý phụ hoạt động ở chế độ ngủ.

Bộ nhớ 130 có thể lưu trữ các phần dữ liệu, chương trình hoặc ứng dụng khác nhau để dẫn động và điều khiển thiết bị xử lý hình ảnh 100.

Chương trình được lưu trữ trong bộ nhớ 130 có thể bao gồm một hoặc nhiều lệnh. Chương trình (một hoặc nhiều lệnh) hoặc ứng dụng được lưu trữ trong bộ nhớ 130 có thể được trích xuất bởi bộ xử lý 120.

Bộ xử lý 120 có thể thực thi một hoặc nhiều lệnh được lưu trữ trong bộ nhớ 130 nhờ đó xử lý hình ảnh thứ nhất bằng cách sử dụng mạng thần kinh nhân tạo chính. Mạng thần kinh nhân tạo chính có thể là mô hình được huấn luyện để xử lý hình ảnh theo điều kiện định trước. Mạng thần kinh nhân tạo chính có thể là CNN bao gồm nhiều lớp tích chập. Ví dụ, mạng thần kinh nhân tạo chính có thể là mô hình được huấn luyện để đưa ra hình ảnh chuyển kiểu bằng cách áp dụng kiểu định trước cho hình ảnh đầu vào. Một cách tùy ý, mạng thần kinh nhân tạo chính có thể là mô hình được huấn luyện để chuyển hình ảnh có độ phân giải thấp đầu vào thành hình ảnh có

độ phân giải cao và đưa ra hình ảnh có độ phân giải cao. Tuy nhiên, các phương án thực hiện của sáng chế không giới hạn ở đó.

Bộ xử lý 120 có thể xử lý hình ảnh thứ nhất bằng cách sử dụng mạng thần kinh nhân tạo chính để thu được hình ảnh đầu ra cuối (ví dụ, hình ảnh thứ hai).

Bộ xử lý 120 có thể thu được dữ liệu kết quả trung gian trong khi xử lý hình ảnh thứ nhất và xử lý dữ liệu kết quả trung gian thu được bằng cách sử dụng mạng thần kinh nhân tạo phụ. Bộ xử lý 120 có thể thu được hình ảnh đầu ra trung gian (ví dụ, hình ảnh thứ ba) khác với hình ảnh đầu ra cuối (ví dụ, hình ảnh thứ hai), bằng cách xử lý dữ liệu kết quả trung gian bằng cách sử dụng mạng thần kinh nhân tạo phụ.

Cụ thể, khi mạng thần kinh nhân tạo chính là mô hình được huấn luyện để áp dụng kiểu định trước, mức độ mà kiểu định trước đã được áp dụng để thu được hình ảnh đầu ra cuối có thể là khác với mức độ mà kiểu định trước đã được áp dụng để thu được hình ảnh đầu ra trung gian. Ví dụ, mức độ mà kiểu định trước đã được áp dụng để thu được hình ảnh đầu ra trung gian có thể nhỏ hơn so với mức độ mà kiểu định trước đã được áp dụng để thu được hình ảnh đầu ra cuối. Tuy nhiên, các phương án thực hiện của sáng chế không giới hạn ở đó.

Bộ xử lý 120 có thể dự đoán khoảng thời gian (ví dụ, khoảng thời gian thứ nhất) được thực hiện để tạo ra hình ảnh đầu ra trung gian từ hình ảnh thứ nhất (ví dụ, hình ảnh đầu vào) và khoảng thời gian (ví dụ, khoảng thời gian thứ hai) được thực hiện để tạo ra hình ảnh đầu ra cuối từ hình ảnh thứ nhất (ví dụ, hình ảnh đầu vào). Bộ xử lý 120 có thể xác định xem có tạo ra hình ảnh đầu ra trung gian, bằng cách sử dụng mạng thần kinh nhân tạo phụ, bằng cách so sánh sự khác nhau giữa các khoảng thời gian thứ nhất và thứ hai đã dự đoán với thời gian định trước.

Bộ xử lý 120 có thể điều khiển màn hình để hiển thị khoảng thời gian được thực hiện để tạo ra hình ảnh đầu ra trung gian từ thời điểm hiện tại hoặc khoảng thời gian được thực hiện để tạo ra hình ảnh đầu ra cuối từ thời điểm hiện tại.

Bộ xử lý 120 có thể kết thúc quá trình xử lý hình ảnh thứ nhất bằng cách sử dụng mạng thần kinh nhân tạo chính, trước khi hình ảnh đầu ra cuối được tạo ra, dựa vào đầu vào của người dùng. Bộ xử lý 120 có thể cho phép người dùng dừng việc xử lý hình ảnh thứ nhất qua giao diện người dùng trong khi cấp hình ảnh đầu ra trung gian đến người dùng. Ví dụ, khi hình ảnh đầu ra trung gian được hiển thị và ở thời điểm này yêu cầu ngắt quá trình xử lý hình ảnh được nhận, bộ xử lý 120 có thể ngắt quá trình xử lý hình ảnh thứ nhất trong mạng thần kinh nhân tạo chính. Do đó, bộ xử lý 120 có thể không tạo ra hình ảnh đầu ra trung gian kế tiếp và hình ảnh đầu ra cuối.

Fig.15 là sơ đồ khối của bộ xử lý 120 theo một phương án thực hiện của sáng chế.

Tham khảo tới Fig.15, bộ xử lý 120 có thể bao gồm bộ học dữ liệu 1410 và bộ xử lý dữ liệu 1420.

Bộ học dữ liệu 1410 có thể học tiêu chuẩn xử lý hình ảnh theo mục đích định trước để huấn luyện mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ theo một phương án thực hiện của sáng chế. Ví dụ, bộ học dữ liệu 1410 có thể học tiêu chuẩn liên quan đến thông tin nào (ví dụ, thông tin đặc điểm) của hình ảnh đầu vào được sử dụng để áp dụng kiểu định trước vào hình ảnh đầu vào.

Bộ học dữ liệu 1410 có thể học tiêu chuẩn liên quan đến cách áp dụng kiểu định trước, bằng cách sử dụng thông tin đặc điểm của hình ảnh. Bộ học dữ liệu 1410 có thể học tiêu chuẩn áp dụng kiểu định trước vào hình ảnh đầu vào, bằng cách thu được dữ liệu (ví dụ, hình ảnh) để sử dụng trong quá trình học và áp dụng dữ liệu thu được vào các mô hình xử lý dữ liệu (mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ).

Các mô hình xử lý dữ liệu (ví dụ, mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ) có thể được thiết lập xem xét, ví dụ, lĩnh vực ứng dụng của mô hình nhận dạng, mục đích học, hoặc hiệu suất máy tính của thiết bị. Các mô hình xử

lý dữ liệu có thể là, ví dụ, các mô hình dựa trên mạng thần kinh nhân tạo. Ví dụ, mô hình, như mạng thần kinh nhân tạo sâu (DNN), mạng thần kinh nhân tạo hồi quy (RNN - Recurrent Neural Network), hoặc DNN hồi quy hai chiều (BRDNN), có thể được sử dụng như mô hình xử lý dữ liệu, nhưng các phương án thực hiện của sáng chế không giới hạn ở đó.

Bộ học dữ liệu 1410 có thể huấn luyện các mô hình xử lý dữ liệu bằng cách sử dụng thuật toán học hoặc dạng tương tự bao gồm, ví dụ, lan truyền ngược lỗi hoặc giảm gradient.

Bộ học dữ liệu 1410 có thể huấn luyện mô hình xử lý dữ liệu thông qua việc học có giám sát bằng cách sử dụng, ví dụ, dữ liệu học làm giá trị đầu vào. Bộ học dữ liệu 1410 có thể huấn luyện mô hình xử lý dữ liệu qua việc học không giám sát để tìm ra tiêu chuẩn xử lý dữ liệu, ví dụ, bằng cách tự huấn luyện loại dữ liệu cần để xử lý dữ liệu mà không giám sát. Bộ học dữ liệu 1410 có thể huấn luyện mô hình xử lý dữ liệu qua học củng cố bằng cách sử dụng phản hồi về, ví dụ, việc liệu kết quả xử lý dữ liệu theo huấn luyện có đúng không.

Khi mô hình xử lý dữ liệu được huấn luyện, bộ học dữ liệu 1410 có thể lưu trữ mô hình xử lý dữ liệu được huấn luyện. Trong trường hợp này, bộ học dữ liệu 1410 có thể lưu trữ các mô hình xử lý dữ liệu được huấn luyện trong bộ nhớ của thiết bị xử lý hình ảnh. Một cách tùy ý, bộ học dữ liệu 1410 có thể lưu trữ mô hình xử lý dữ liệu được huấn luyện trong bộ nhớ của máy chủ mà được nối với thiết bị xử lý hình ảnh qua mạng không dây hoặc có dây.

Cụ thể, bộ nhớ lưu trữ các mô hình xử lý dữ liệu được huấn luyện cũng có thể lưu trữ, ví dụ, lệnh hoặc dữ liệu liên quan đến ít nhất một thành phần khác của thiết bị xử lý hình ảnh. Bộ nhớ cũng có thể lưu trữ phần mềm và/hoặc chương trình. Chương trình có thể bao gồm, ví dụ, nhân, phần mềm trung gian, một giao diện lập trình ứng dụng (API), và/hoặc chương trình ứng dụng (hoặc ứng dụng).

Bộ xử lý dữ liệu 1420 có thể đưa vào hình ảnh vào mô hình xử lý dữ liệu bao gồm mạng thần kinh nhân tạo chính được huấn luyện, và mô hình xử lý dữ liệu có thể đưa ra, làm giá trị kết quả, hình ảnh được xử lý theo mục đích định trước. Ví dụ, mô hình xử lý dữ liệu có thể đưa ra hình ảnh thu được bằng cách áp dụng kiểu định trước vào hình ảnh đầu vào. Giá trị kết quả đầu ra (hình ảnh) có thể được sử dụng để cập nhật mô hình xử lý dữ liệu bao gồm mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ.

Ít nhất một trong số bộ học dữ liệu 1410 hoặc bộ xử lý dữ liệu 1420 có thể được sản xuất dưới dạng ít nhất một chip phần cứng và có thể được gắn trên thiết bị xử lý hình ảnh. Ví dụ, ít nhất một trong số bộ học dữ liệu 1410 hoặc bộ xử lý dữ liệu 1420 có thể được sản xuất dưới dạng chip phần cứng chuyên dụng cho trí tuệ nhân tạo (AI - Artificial Intelligence) hoặc có thể được sản xuất như một phần của bộ xử lý đa năng hiện có (ví dụ, bộ xử lý trung tâm (CPU - Central Processing Unit) hoặc bộ xử lý ứng dụng (AP - Application Processor)) hoặc bộ xử lý dành riêng cho đồ họa, ví dụ, bộ xử lý đồ họa (GPU - Graphics Processing Unit)) và có thể được gắn trên các thiết bị xử lý hình ảnh bất kỳ đã đề cập ở trên.

Bộ học dữ liệu 1410 và bộ xử lý dữ liệu 1420 có thể được gắn trên một thiết bị xử lý hình ảnh, hoặc có thể được gắn tương ứng trên các thiết bị xử lý hình ảnh độc lập. Ví dụ, một trong số bộ học dữ liệu 1410 và bộ xử lý dữ liệu 1420 có thể có trong thiết bị xử lý hình ảnh, và thiết bị còn lại có thể có trong máy chủ. Bộ học dữ liệu 1410 và bộ xử lý dữ liệu 1420 có thể được nối với nhau qua dây dẫn hoặc không dây, và vì vậy thông tin mô hình được thiết lập bởi bộ học dữ liệu 1410 có thể được tạo ra cho bộ xử lý dữ liệu 1420 và đầu vào dữ liệu cho bộ xử lý dữ liệu 1420 có thể được tạo ra là dữ liệu học bổ sung cho bộ học dữ liệu 1410.

Ít nhất một trong số bộ học dữ liệu 1410 hoặc bộ xử lý dữ liệu 1420 có thể được thực hiện làm môđun phần mềm. Khi ít nhất một trong số bộ học dữ liệu 1410

hoặc bộ xử lý dữ liệu 1420 được thực hiện bằng cách sử dụng môđun phần mềm (hoặc môđun chương trình bao gồm các lệnh), môđun phần mềm có thể được lưu trữ trong phương tiện ghi có thể đọc được bằng máy tính không nhất thời. Trong trường hợp này, ít nhất một môđun phần mềm có thể được tạo ra bởi hệ điều hành (OS) hoặc bởi ứng dụng nhất định. Một cách tùy ý, một số trong số ít nhất một môđun phần mềm có thể được tạo ra bởi OS và các môđun khác có thể được tạo ra bởi một ứng dụng nhất định.

Fig.16 là sơ đồ thể hiện ví dụ về nơi thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế và máy chủ tương tác với nhau để học và nhận dạng dữ liệu.

Tham khảo tới Fig.16, máy chủ 2000 có thể huấn luyện mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ bằng cách học tiêu chuẩn xử lý hình ảnh theo mục đích định trước.

Trong trường hợp này, máy chủ 2000 có thể thực hiện chức năng của bộ học dữ liệu 1410 trên Fig.15. Máy chủ 2000 có thể học tiêu chuẩn liên quan đến dữ liệu học nào được sử dụng để xử lý hình ảnh đầu vào theo mục đích định trước. Máy chủ 2000 có thể học tiêu chuẩn để xử lý hình ảnh theo mục đích định trước bằng cách sử dụng dữ liệu.

Máy chủ 2000 có thể huấn luyện ít nhất một trong số mạng thần kinh nhân tạo chính hoặc mạng thần kinh nhân tạo phụ bằng cách sử dụng dữ liệu học.

Thiết bị xử lý hình ảnh 100 có thể truyền dữ liệu đến máy chủ 2000 và có thể yêu cầu máy chủ 2000 xử lý dữ liệu bằng cách áp dụng dữ liệu cho các mô hình xử lý dữ liệu (mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ). Ví dụ, máy chủ 2000 có thể xử lý hình ảnh theo mục đích định trước bằng cách sử dụng các mô hình xử lý dữ liệu (mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ). Ví dụ, máy chủ 2000 có thể đưa ra hình ảnh chuyển kiểu bằng cách áp dụng

kiểu định trước cho hình ảnh đầu vào.

Một cách tùy ý, thiết bị xử lý hình ảnh 100 có thể nhận các mô hình xử lý dữ liệu được tạo ra bởi máy chủ 2000 từ máy chủ 2000, và có thể xử lý dữ liệu bằng cách sử dụng các mô hình xử lý dữ liệu nhận được. Ví dụ, thiết bị xử lý hình ảnh 100 có thể xử lý hình ảnh theo mục đích định trước bằng cách sử dụng các mô hình xử lý dữ liệu nhận được (mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ). Ví dụ, thiết bị xử lý hình ảnh 100 có thể đưa ra hình ảnh chuyển kiểu bằng cách áp dụng kiểu định trước cho hình ảnh đầu vào.

Fig.17 là sơ đồ khối của cấu trúc của thiết bị xử lý hình ảnh 1700 theo một phương án thực hiện khác của sáng chế. Thiết bị xử lý hình ảnh 1700 trên Fig.17 có thể là phương án thực hiện của thiết bị xử lý hình ảnh 100 trên Fig.14.

Tham khảo tới Fig.17, thiết bị xử lý hình ảnh 1700 có thể bao gồm bộ dò 1740, bộ điều khiển 1710, màn hình 1720, giao diện truyền thông 1750, bộ cảm biến 1730, giao diện đầu vào/đầu ra (I/O) 1770, bộ xử lý video 1780, bộ xử lý âm thanh 1715, bộ lưu trữ 1790, và nguồn điện 1760.

Bộ dò 1740 có thể chỉ dò và chọn tần số của kênh mà thiết bị xử lý hình ảnh 1700 muốn nhận trong số nhiều thành phần sóng vô tuyến thu được thông qua khuếch đại, trộn, cộng hưởng hoặc tương tự tín hiệu phát sóng có dây hoặc không dây. Tín hiệu phát sóng bao gồm âm thanh, video và thông tin bổ sung (ví dụ, hướng dẫn chương trình điện tử (EPG - Electronic Program Guide)).

Bộ dò 1740 có thể nhận tín hiệu phát sóng trong dải tần số tương ứng với số kênh theo đầu vào người dùng (ví dụ, tín hiệu điều khiển được tiếp nhận từ thiết bị điều khiển, ví dụ, đầu vào số kênh, đầu vào lên-xuống kênh, và đầu vào kênh trên hình ảnh màn hình EPG).

Bộ dò 1740 có thể nhận tín hiệu phát sóng từ nhiều nguồn khác nhau, chẳng hạn như phát sóng mặt đất, phát sóng cáp, phát sóng vệ tinh và phát sóng Internet. Bộ

dò 1740 cũng có thể nhận tín hiệu phát sóng từ một nguồn như phát sóng tương tự hoặc phát sóng số.

Giao diện truyền thông 1750 có thể truyền hoặc nhận dữ liệu hoặc tín hiệu đến hoặc từ thiết bị bên ngoài hoặc máy chủ dưới sự điều khiển của bộ điều khiển 1710. Bộ điều khiển 1710 có thể truyền/nhận nội dung đến/từ thiết bị bên ngoài được kết nối qua giao diện truyền thông 1750, tải xuống ứng dụng từ thiết bị bên ngoài hoặc thực hiện duyệt web. Giao diện truyền thông 1750 có thể truyền hoặc nhận dữ liệu hoặc tín hiệu theo ít nhất một phương pháp giữa mạng cục bộ không dây (LAN - Local Area Network) 1751 (ví dụ: Wi-Fi), mạng Bluetooth 1752 hoặc mạng Ethernet có dây 1753 tương ứng với hiệu suất và cấu trúc của thiết bị xử lý hình ảnh 1700.

Bộ xử lý video 1780 xử lý dữ liệu video được nhận bởi thiết bị xử lý hình ảnh 1700. Bộ xử lý video 1780 có thể thực hiện nhiều loại xử lý hình ảnh, chẳng hạn như giải mã, chia tỷ lệ, lọc nhiễu, chuyển đổi tốc độ khung hình và chuyển đổi độ phân giải trên dữ liệu video đã nhận.

Bộ cảm biến 1730 cảm nhận giọng nói của người dùng, hình ảnh của người dùng, hoặc tương tác với người dùng, và có thể bao gồm micrô 1731, camera 1732, và bộ thu ánh sáng 1733.

Micrô 1731 nhận giọng nói của người dùng. Micrô 1731 có thể chuyển giọng nói nhận được thành tín hiệu điện và xuất tín hiệu điện tới bộ điều khiển 1710. Giọng nói người dùng có thể bao gồm, ví dụ, giọng nói tương ứng với trình đơn hoặc chức năng của thiết bị xử lý hình ảnh 1700.

Camera 1732 có thể nhận hình ảnh (ví dụ, các khung hình liên tiếp) tương ứng với chuyển động của người dùng bao gồm cử chỉ trong phạm vi nhận dạng của camera 1732. Bộ điều khiển 1710 có thể chọn trình đơn hiển thị trên thiết bị xử lý hình ảnh 1700 bằng cách sử dụng kết quả của việc nhận dạng chuyển động đã nhận, hoặc thực hiện điều khiển tương ứng với kết quả nhận dạng chuyển động.

Bộ thu ánh sáng 1733 nhận tín hiệu quang (bao gồm tín hiệu điều khiển) từ thiết bị điều khiển bên ngoài qua cửa sổ ánh sáng hoặc phản tương tự của khung viền của màn hình 1720. Bộ thu ánh sáng 1733 có thể nhận tín hiệu quang tương ứng với đầu vào người dùng (ví dụ, chạm, nhấn, cử chỉ chạm, giọng nói hoặc chuyển động) từ thiết bị điều khiển. Tín hiệu điều khiển có thể là được trích xuất từ tín hiệu quang nhận được dưới sự điều khiển của bộ điều khiển 1710.

Giao diện I/O 1770 nhận video (ví dụ, hình ảnh chuyển động), âm thanh (ví dụ, giọng nói hoặc âm nhạc) và thông tin bổ sung (ví dụ, EPG) từ bên ngoài thiết bị xử lý hình ảnh 1700 dưới sự điều khiển của bộ điều khiển 1710. Giao diện I/O 1770 có thể bao gồm cổng giao diện đa phương tiện độ nét cao (HDMI - High-Definition Multimedia Interface) 1771, jack cắm tổ hợp 1772, cổng PC 1773 hoặc cổng USB 1774. Giao diện I/O 1770 có thể bao gồm sự kết hợp của cổng HDMI 1771, jack cắm tổ hợp 1772, cổng PC 1773 và cổng USB 1774.

Bộ điều khiển 1710 điều khiển toàn bộ hoạt động của thiết bị xử lý hình ảnh 1700 và truyền tín hiệu giữa các bộ phận trong của thiết bị xử lý hình ảnh 1700 và xử lý dữ liệu. Khi có đầu vào của người dùng hoặc thỏa mãn các điều kiện định trước được lưu trữ, bộ điều khiển 1710 có thể thực thi hệ điều hành và các ứng dụng khác nhau được lưu trữ trong bộ lưu trữ 1790.

Bộ điều khiển 1710 có thể bao gồm bộ nhớ truy cập ngẫu nhiên (RAM - Random-Access Memory) 1781 lưu trữ tín hiệu hoặc dữ liệu được đưa vào bởi nguồn bên ngoài của thiết bị xử lý hình ảnh 1700 hoặc được sử dụng làm vùng bộ nhớ cho các hoạt động khác nhau được thực hiện bởi thiết bị xử lý hình ảnh 1700, bộ nhớ chỉ đọc (ROM - Read-Only Memory) 1782 lưu trữ chương trình điều khiển để điều khiển thiết bị xử lý hình ảnh 1700, và bộ xử lý 1783.

Bộ xử lý đồ họa 1784 tạo ra hình ảnh trên màn hình bao gồm các đối tượng khác nhau, chẳng hạn như biểu tượng, hình ảnh và văn bản, bằng cách sử dụng bộ số

học và bộ kết xuất. Bộ số học tính toán các giá trị thuộc tính, chẳng hạn như giá trị tọa độ, hình dạng, kích thước, màu sắc và những thứ tương tự mà mỗi đối tượng sẽ được hiển thị theo bố cục của hình ảnh trên màn hình, dựa trên tương tác của người dùng được cảm nhận bởi cảm biến 1730. Bộ kết xuất tạo ra hình ảnh màn hình của nhiều bố cục khác nhau bao gồm các đối tượng, dựa trên các giá trị thuộc tính được tính bằng bộ số học. Các hình ảnh màn hình do đơn vị kết xuất tạo ra được hiển thị trên vùng hiển thị của màn hình 1720.

Các giao diện thứ nhất 1785-1 đến thứ n 1785-n được kết nối với các bộ phận được mô tả ở trên của thiết bị xử lý hình ảnh 1700. Một trong số các giao diện thứ nhất 1285-1 đến thứ n 1285-n có thể là giao diện mạng được kết nối với thiết bị bên ngoài qua mạng.

RAM 1781, ROM 1782, bộ xử lý 1783, bộ xử lý đồ họa 1784, và các giao diện thứ nhất 1785-1 đến thứ n 1785-n có thể được nối với nhau qua bus trong 1786.

Thuật ngữ “bộ điều khiển của thiết bị xử lý hình ảnh” được sử dụng trong phương án thực hiện này của sáng chế bao gồm bộ xử lý 1783, ROM 1782, và RAM 1781.

Màn hình 1720 có thể tạo ra tín hiệu điều khiển bằng cách chuyển đổi tín hiệu hình ảnh, tín hiệu dữ liệu, tín hiệu hiển thị trên màn hình (OSD - On-Screen Display), và tín hiệu điều khiển được xử lý bởi bộ điều khiển 1710. Màn hình 1720 có thể là màn hình plasma (PDP - Plasma Display Panel), màn hình tinh thể lỏng (LCD - Liquid-Crystal Display), thiết bị phát sáng hữu cơ (OLED - Organic Light-Emitting Device), màn hình mềm dẻo hoặc màn hình 3 chiều (3D - 3-Dimensional). Màn hình 1720 có thể được tạo cấu hình như một màn hình cảm ứng và do đó có thể dùng làm thiết bị đầu vào cũng như thiết bị đầu ra.

Bộ xử lý âm thanh 1715 xử lý dữ liệu âm thanh. Bộ xử lý âm thanh 1715 có thể thực hiện nhiều loại xử lý, chẳng hạn như giải mã, khuếch đại hoặc lọc tiếng ồn

trên dữ liệu âm thanh. Bộ xử lý âm thanh 1715 có thể bao gồm nhiều môđun xử lý âm thanh để xử lý âm thanh tương ứng với nhiều phần nội dung.

Giao diện đầu ra âm thanh 1725 xuất ra âm thanh có trong tín hiệu phát sóng nhận được qua bộ dò 1740, dưới sự điều khiển của bộ điều khiển 1710. Giao diện đầu ra âm thanh 1725 cũng có thể xuất ra âm thanh (ví dụ, giọng nói hoặc âm thanh) được đưa vào qua giao diện truyền thông 1750 hoặc giao diện I/O 1770. Giao diện đầu ra âm thanh 1725 cũng có thể xuất âm thanh được lưu trữ trong bộ lưu trữ 1790 dưới sự điều khiển của bộ điều khiển 1710. Giao diện đầu ra âm thanh 1725 có thể bao gồm ít nhất một giao diện được chọn từ loa 1726, cổng đầu ra tai nghe 1727 và cổng đầu ra giao diện số Sony/Philips (S/PDIF - Sony / Philips Digital Interface) 1728. Giao diện đầu ra âm thanh 1725 có thể bao gồm sự kết hợp của loa 1726, cổng đầu ra tai nghe 1727 và cổng đầu ra S/PDIF 1728.

Nguồn cấp điện 1760 cấp điện được đưa vào từ nguồn điện bên ngoài tới các bộ phận trong của thiết bị xử lý hình ảnh 1700, dưới sự điều khiển của bộ điều khiển 1710. Nguồn cấp điện 1760 cũng có thể cấp điện được phát ra bởi một hoặc nhiều pin nằm trong thiết bị xử lý hình ảnh 1700, đến các bộ phận trong thiết bị xử lý hình ảnh 1700, dưới sự điều khiển của bộ điều khiển 1710.

Bộ lưu trữ 1790 có thể lưu trữ các dữ liệu, các chương trình hoặc các ứng dụng khác nhau để dẫn động và điều khiển thiết bị xử lý hình ảnh 1700 dưới sự điều khiển của bộ điều khiển 1710. Bộ lưu trữ 1790 có thể bao gồm một môđun thu phát sóng, một môđun điều khiển kênh, một môđun điều khiển âm lượng, môđun điều khiển truyền thông, môđun nhận dạng giọng nói, môđun nhận dạng chuyển động, môđun nhận ánh sáng, môđun điều khiển màn hình, môđun điều khiển âm thanh, môđun điều khiển đầu vào bên ngoài, môđun điều khiển công suất, môđun điều khiển công suất của thiết bị bên ngoài kết nối không dây (ví dụ, Bluetooth), cơ sở dữ liệu thoại (DB - voice database) hoặc DB chuyển động. Các môđun này và các DB của bộ lưu trữ

1790 có thể được triển khai dưới dạng phần mềm để thực hiện chức năng điều khiển thu phát sóng của thiết bị xử lý hình ảnh 1700, chức năng điều khiển kênh, chức năng điều khiển âm lượng, chức năng điều khiển truyền thông, chức năng nhận dạng giọng nói, chức năng nhận dạng chuyển động, chức năng điều khiển nhận ánh sáng, chức năng điều khiển màn hình, chức năng điều khiển âm thanh, chức năng điều khiển đầu vào bên ngoài, chức năng điều khiển công suất, hoặc chức năng điều khiển công suất của thiết bị bên ngoài kết nối không dây (ví dụ, Bluetooth). Bộ điều khiển 1710 có thể thực hiện các chức năng này bằng cách sử dụng phần mềm được lưu trữ trong bộ lưu trữ 1790.

Các sơ đồ khối của các thiết bị xử lý hình ảnh 100 và 1700 được thể hiện trên Fig.14 và Fig.17 chỉ là các phương án thực hiện làm ví dụ của sáng chế. Các bộ phận được thể hiện trên Fig.14 và Fig.17 có thể được kết hợp hoặc bỏ qua khỏi phần mô tả của các thiết bị xử lý hình ảnh 100 và 1700 khi được thực hiện trên thực tế, hoặc các bộ phận bổ sung có thể có trong các sơ đồ khối trên Fig.14 và Fig.17. Nói cách khác, hai hoặc nhiều bộ phận có thể được kết hợp thành một bộ phận duy nhất, hoặc một bộ phận duy nhất có thể được chia thành hai hoặc nhiều bộ phận. Chức năng được thực hiện trong mỗi khối chỉ là một ví dụ để minh họa các phương án thực hiện của sáng chế, và hoạt động hoặc thiết bị chi tiết của từng khối không giới hạn phạm vi của các phương án thực hiện của sáng chế.

Phương pháp vận hành thiết bị xử lý hình ảnh, theo một phương án thực hiện của sáng chế, có thể được thể hiện dưới dạng các lệnh chương trình có thể thực thi được bằng các phương tiện máy tính khác nhau và có thể được ghi lại trên phương tiện ghi đọc được bằng máy tính. Phương tiện ghi đọc được bằng máy tính có thể bao gồm các lệnh chương trình, tệp dữ liệu, cấu trúc dữ liệu và những thứ tương tự riêng biệt hoặc kết hợp. Các lệnh chương trình được ghi lại trên phương tiện ghi đọc được bằng máy tính có thể được thiết kế và được tạo cấu hình đặc biệt cho các phương án thực hiện của sáng chế hoặc có thể đã được biết đến và có thể sử dụng được bởi người

có hiểu biết trung bình trong lĩnh vực phần mềm máy tính. Các ví dụ về phương tiện ghi đọc được bằng máy tính bao gồm phương tiện từ tính như đĩa cứng, đĩa mềm hoặc băng từ, phương tiện quang học như bộ nhớ chỉ đọc đĩa nhỏ gọn (CD-ROM) hoặc đĩa đa năng kỹ thuật số (DVD), phương tiện quang-từ như đĩa quang và thiết bị phần cứng được tạo cấu hình đặc biệt để lưu trữ và thực thi các lệnh chương trình như ROM, bộ nhớ truy cập ngẫu nhiên (RAM) hoặc bộ nhớ nhanh. Ví dụ về các lệnh chương trình là các mã ngôn ngữ nâng cao có thể được thực thi bởi máy tính bằng cách sử dụng trình thông dịch hoặc tương tự cũng như mã ngôn ngữ máy do trình biên dịch tạo ra.

Ngoài ra, thiết bị xử lý hình ảnh hoặc phương pháp vận hành thiết bị xử lý hình ảnh theo các phương án thực hiện được bộc lộ có thể được tạo ra trong sản phẩm chương trình máy tính. Sản phẩm chương trình máy tính có thể được mua bán như một hàng hóa giữa người bán và người mua.

Sản phẩm chương trình máy tính có thể bao gồm chương trình phần mềm và phương tiện lưu trữ đọc được bằng máy tính có chương trình phần mềm được lưu trữ trên đó. Ví dụ, sản phẩm chương trình máy tính có thể bao gồm sản phẩm ở dạng chương trình phần mềm (ví dụ, ứng dụng có thể tải xuống) được phân phối điện tử qua nhà sản xuất thiết bị điện tử hoặc chợ điện tử (ví dụ, Cửa hàng Google Play, AppStore). Đối với phân phối điện tử, ít nhất một phần của chương trình phần mềm có thể được lưu trữ trên phương tiện lưu trữ hoặc có thể được tạo tạm thời. Trong trường hợp này, phương tiện lưu trữ có thể là máy chủ của nhà sản xuất, máy chủ của chợ điện tử hoặc phương tiện lưu trữ của máy chủ chuyển tiếp để lưu trữ tạm thời chương trình SW.

Sản phẩm chương trình máy tính có thể bao gồm, trong hệ thống bao gồm máy chủ và thiết bị khách, phương tiện lưu trữ của máy chủ hoặc phương tiện lưu trữ của thiết bị khách. Một cách tùy ý, khi có thiết bị thứ ba (ví dụ, điện thoại thông minh)

truyền thông với máy chủ hoặc thiết bị khách, sản phẩm chương trình máy tính có thể bao gồm phương tiện lưu trữ của thiết bị thứ ba. Một cách tùy ý, sản phẩm chương trình máy tính có thể bao gồm chính chương trình S/W được truyền từ máy chủ đến thiết bị khách hoặc thiết bị thứ ba, hoặc được truyền từ thiết bị thứ ba đến thiết bị khách.

Trong trường hợp này, một trong số máy chủ, thiết bị khách, và thiết bị thứ ba có thể thực thi sản phẩm chương trình máy tính để thực hiện các phương pháp theo các phương án thực hiện được bộc lộ. Một cách tùy ý, ít nhất hai trong số máy chủ, thiết bị khách và thiết bị thứ ba có thể thực thi sản phẩm chương trình máy tính để phân phối và thực hiện các phương pháp theo các phương án thực hiện được bộc lộ.

Ví dụ, máy chủ (ví dụ, máy chủ đám mây hoặc máy chủ trí tuệ nhân tạo) có thể thực thi sản phẩm chương trình máy tính được lưu trữ trên máy chủ để điều khiển thiết bị khách truyền thông với máy chủ để thực hiện các phương pháp theo các phương án thực hiện được bộc lộ.

Thiết bị xử lý hình ảnh theo một phương án thực hiện của sáng chế có thể tạo ra hình ảnh đầu ra trung gian bằng cách sử dụng mạng thần kinh nhân tạo phụ trước khi tạo ra hình ảnh đầu ra cuối trong mạng thần kinh nhân tạo chính, và vì vậy thời gian chờ từ điểm quan sát của người dùng có thể được giảm xuống.

Thiết bị xử lý hình ảnh có thể tạo ra hình ảnh đầu ra trung gian, và người dùng có thể xác định xem có tạo ra hình ảnh đầu ra cuối trong mạng thần kinh nhân tạo chính hay không, dựa trên hình ảnh đầu ra trung gian.

Thiết bị xử lý hình ảnh có thể bao gồm các mạng thần kinh nhân tạo xuất ra các kết quả khác nhau mặc dù xử lý một phần các hình ảnh tương tự nhau, làm mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ, nhờ đó làm tăng hiệu suất huấn luyện mạng thần kinh nhân tạo.

Mặc dù không giới hạn ở đó, nhưng phương án thực hiện làm ví dụ có thể được

thể hiện dưới dạng mã có thể đọc được bằng máy tính trên phương tiện ghi có thể đọc được bằng máy tính. Phương tiện ghi có thể đọc được bằng máy tính là thiết bị lưu trữ dữ liệu bất kỳ có thể lưu trữ dữ liệu mà sau đó có thể đọc được bằng hệ thống máy tính. Các ví dụ về phương tiện ghi có thể đọc được bằng máy tính bao gồm bộ nhớ chỉ đọc (ROM), bộ nhớ truy cập ngẫu nhiên (RAM), CD-ROM, băng từ, đĩa mềm và thiết bị lưu trữ dữ liệu quang học. Phương tiện ghi đọc được bằng máy tính cũng có thể được phân phối qua các hệ thống máy tính kết hợp mạng để mã đọc được bằng máy tính được lưu trữ và thực thi theo kiểu phân tán. Ngoài ra, phương án thực hiện làm ví dụ có thể được viết dưới dạng chương trình máy tính được truyền qua phương tiện truyền dẫn đọc được bằng máy tính, chẳng hạn như sóng mang, và được nhận và triển khai trong các máy tính số thông dụng hoặc chuyên dùng thực thi các chương trình. Hơn nữa, cần hiểu rằng trong các phương án thực hiện làm ví dụ, một hoặc nhiều bộ phận của phương tiện và thiết bị được mô tả ở trên có thể bao gồm mạch điện, bộ xử lý, bộ vi xử lý, v.v. và có thể thực thi chương trình máy tính được lưu trữ trong phương tiện có thể đọc được bằng máy tính.

Các phương án thực hiện làm ví dụ nêu trên chỉ nhằm làm ví dụ và không được hiểu là giới hạn. Hướng dẫn này có thể dễ dàng áp dụng cho các loại thiết bị khác. Ngoài ra, phần mô tả các phương án thực hiện làm ví dụ nhằm minh họa và không nhằm giới hạn phạm vi bảo hộ của các điểm yêu cầu bảo hộ, và nhiều phương án thay thế, cải biến và biến đổi sẽ là hiển nhiên đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này.

YÊU CẦU BẢO HỘ

1. Thiết bị xử lý hình ảnh bao gồm:

bộ nhớ lưu trữ một hoặc nhiều lệnh, và các cấu trúc dữ liệu cho mạng thần kinh nhân tạo chính và mạng thần kinh nhân tạo phụ; và

bộ xử lý được tạo cấu hình để thực thi một hoặc nhiều lệnh được lưu trữ trong bộ nhớ, để xử lý hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính để thu được dữ liệu kết quả trung gian, và để xử lý dữ liệu kết quả trung gian bằng cách sử dụng mạng thần kinh nhân tạo phụ để đưa ra hình ảnh trung gian trong khi hình ảnh đầu vào đang được xử lý bằng cách sử dụng mạng thần kinh nhân tạo chính.

2. Thiết bị xử lý hình ảnh theo điểm 1, trong đó:

mạng thần kinh nhân tạo chính bao gồm nhiều lớp tích chập thứ nhất, và

bộ xử lý còn được tạo cấu hình để thực hiện hoạt động tích chập thứ nhất trên hình ảnh đầu vào bằng cách sử dụng một hoặc nhiều nhân trong nhiều lớp tích chập thứ nhất để trích xuất thông tin đặc điểm và đưa ra dữ liệu kết quả trung gian dựa trên thông tin đặc điểm được trích xuất.

3. Thiết bị xử lý hình ảnh theo điểm 1, trong đó:

mạng thần kinh nhân tạo phụ bao gồm một hoặc nhiều lớp tích chập thứ hai, và

bộ xử lý còn được tạo cấu hình để thực hiện hoạt động tích chập thứ hai bằng cách áp dụng một hoặc nhiều nhân cho dữ liệu kết quả trung gian trong một hoặc nhiều lớp tích chập thứ hai để trích xuất thông tin đặc điểm và tạo ra hình ảnh trung gian dựa trên thông tin đặc điểm được trích xuất.

4. Thiết bị xử lý hình ảnh theo điểm 1, trong đó bộ xử lý còn được tạo cấu hình để xử lý hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính để thu được hình ảnh đầu ra cuối.

5. Thiết bị xử lý hình ảnh theo điểm 4, trong đó mạng thần kinh nhân tạo chính là mạng được huấn luyện để tạo ra hình ảnh đầu ra cuối bằng cách áp dụng kiểu định trước cho hình ảnh đầu vào, và mức độ mà kiểu định trước đã được áp dụng cho hình ảnh trung gian nhỏ hơn mức độ mà kiểu định trước đã được áp dụng cho hình ảnh đầu ra cuối.

6. Thiết bị xử lý hình ảnh theo điểm 4, trong đó khoảng thời gian thứ nhất được thực hiện để tạo ra hình ảnh đầu ra cuối từ hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính lớn hơn khoảng thời gian thứ hai được thực hiện để tạo ra hình ảnh trung gian từ hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo phụ.

7. Thiết bị xử lý hình ảnh theo điểm 6, trong đó bộ xử lý còn được tạo cấu hình để dự đoán khoảng thời gian thứ nhất và khoảng thời gian thứ hai và xác định xem liệu có tạo ra hình ảnh trung gian bằng cách sử dụng mạng thần kinh nhân tạo phụ hay không, dựa trên khoảng thời gian thứ nhất đã được dự đoán và khoảng thời gian thứ hai đã được dự đoán.

8. Thiết bị xử lý hình ảnh theo điểm 4, trong đó thiết bị này còn bao gồm màn hình, trong đó bộ xử lý còn được tạo cấu hình để điều khiển màn hình để hiển thị khoảng thời gian được thực hiện để tạo ra ít nhất một trong số hình ảnh trung gian hoặc hình ảnh đầu ra cuối từ thời điểm hiện tại.

9. Thiết bị xử lý hình ảnh theo điểm 4, trong đó bộ xử lý còn được tạo cấu hình để dừng việc xử lý hình ảnh đầu vào trước khi hình ảnh đầu ra cuối được tạo ra, dựa vào đầu vào của người dùng.

10. Phương pháp vận hành thiết bị xử lý hình ảnh, phương pháp bao gồm các bước:

xử lý hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính;

thu được dữ liệu kết quả trung gian từ hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính trong khi xử lý hình ảnh đầu vào để thu được hình ảnh

đầu ra cuối; và

tạo ra hình ảnh trung gian dựa trên dữ liệu kết quả trung gian bằng cách sử dụng mạng thần kinh nhân tạo phụ.

11. Phương pháp theo điểm 10, trong đó:

mạng thần kinh nhân tạo chính bao gồm nhiều lớp tích chập thứ nhất,

bước xử lý hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính bao gồm bước trích xuất thông tin đặc điểm bằng cách thực hiện hoạt động tích chập thứ nhất trên hình ảnh đầu vào và một hoặc nhiều nhân trong nhiều lớp tích chập thứ nhất, và

bước thu được dữ liệu kết quả trung gian bao gồm bước tạo ra dữ liệu kết quả trung gian, dựa trên thông tin đặc điểm được trích xuất.

12. Phương pháp theo điểm 10, trong đó:

mạng thần kinh nhân tạo phụ bao gồm một hoặc nhiều lớp tích chập thứ hai, và

bước tạo ra hình ảnh trung gian bao gồm các bước:

trích xuất thông tin đặc điểm bằng cách thực hiện hoạt động tích chập thứ hai bằng cách áp dụng dữ liệu kết quả trung gian cho một hoặc nhiều nhân trong một hoặc nhiều lớp tích chập thứ hai; và

tạo ra hình ảnh trung gian, dựa trên thông tin đặc điểm được trích xuất.

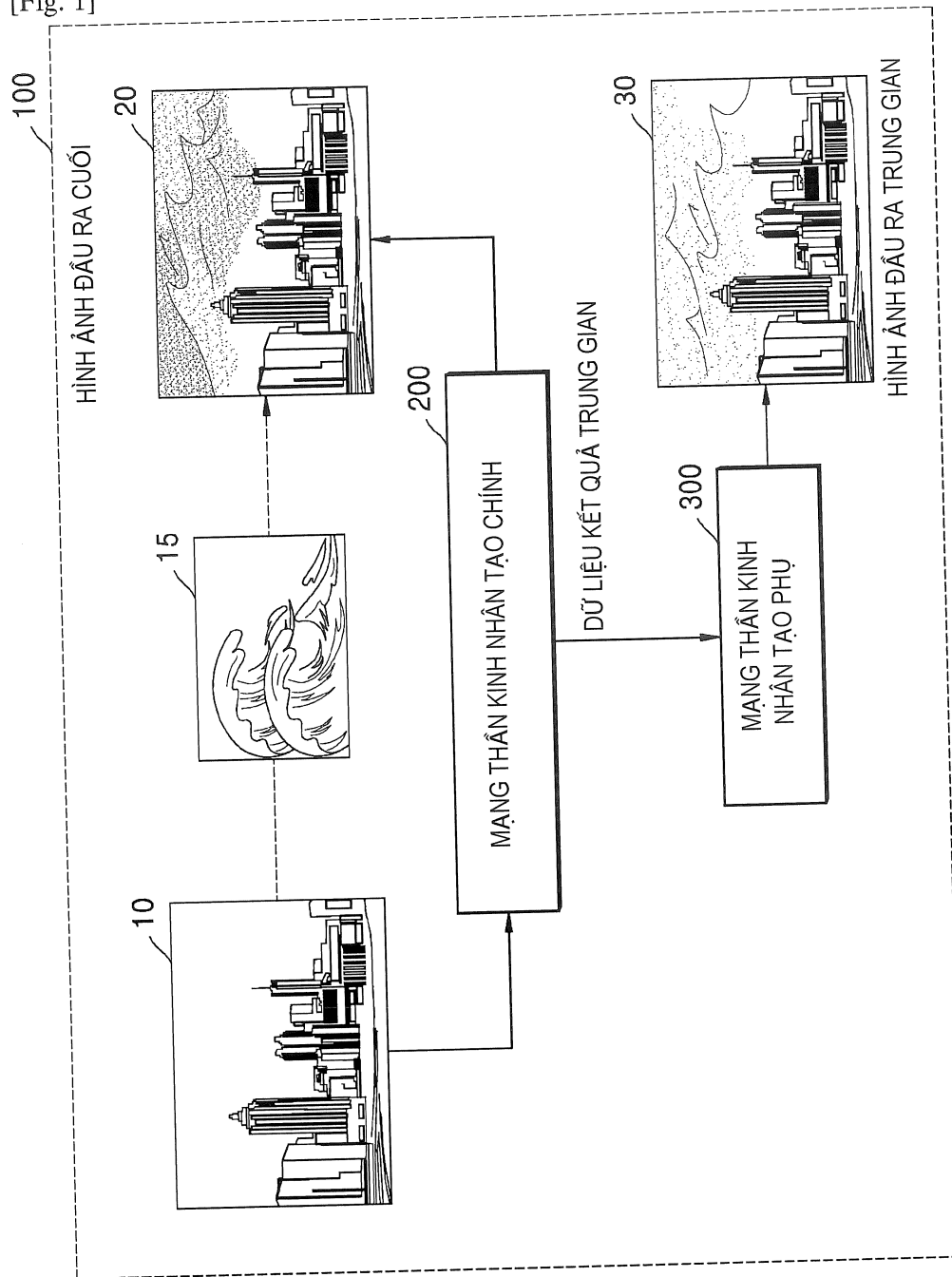
13. Phương pháp theo điểm 10, trong đó phương pháp này còn bao gồm bước tạo ra hình ảnh đầu ra cuối bằng cách xử lý hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính.

14. Phương pháp theo điểm 13, trong đó mạng thần kinh nhân tạo chính là mạng được huấn luyện để tạo ra hình ảnh đầu ra cuối bằng cách áp dụng kiểu định trước cho hình

ảnh đầu vào, và mức độ mà kiểu định trước đã được áp dụng cho hình ảnh trung gian nhỏ hơn mức độ mà kiểu định trước đã được áp dụng cho hình ảnh đầu ra cuối.

15. Phương pháp theo điểm 13, trong đó khoảng thời gian thứ nhất được thực hiện để tạo ra hình ảnh đầu ra cuối từ hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo chính lớn hơn khoảng thời gian thứ hai được thực hiện để tạo ra hình ảnh trung gian từ hình ảnh đầu vào bằng cách sử dụng mạng thần kinh nhân tạo phụ.

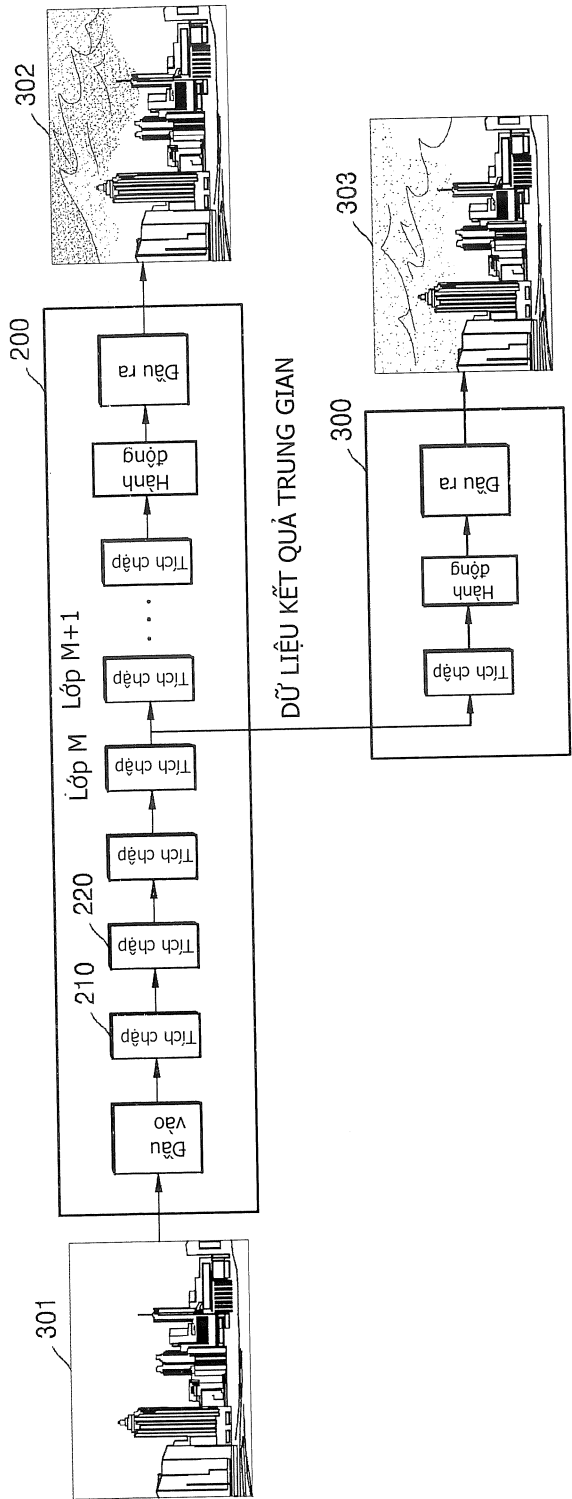
[Fig. 1]



[Fig. 2]

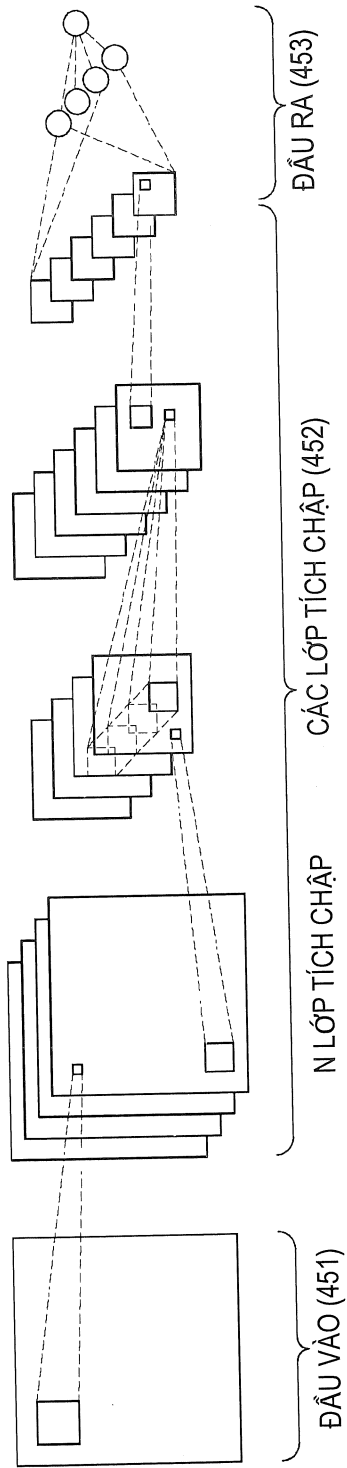


[Fig. 3]



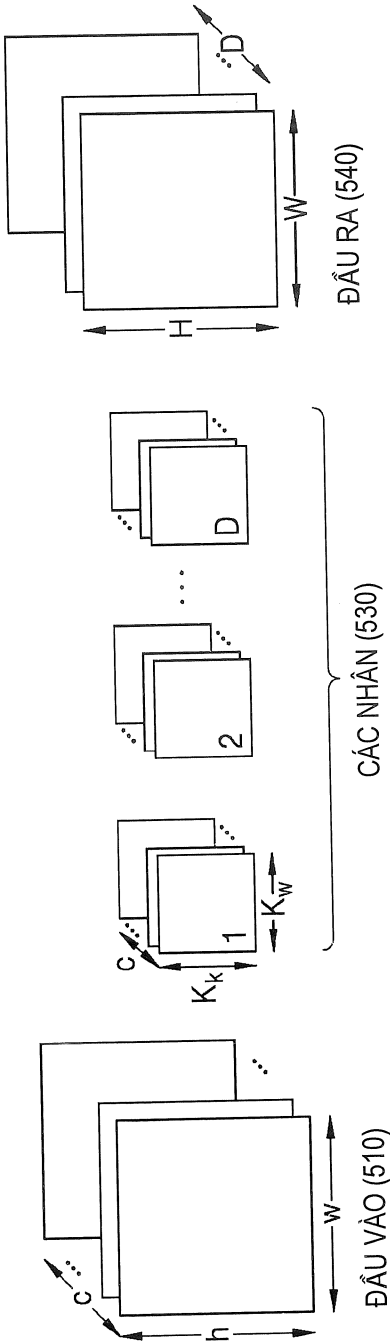
[Fig. 4]

200

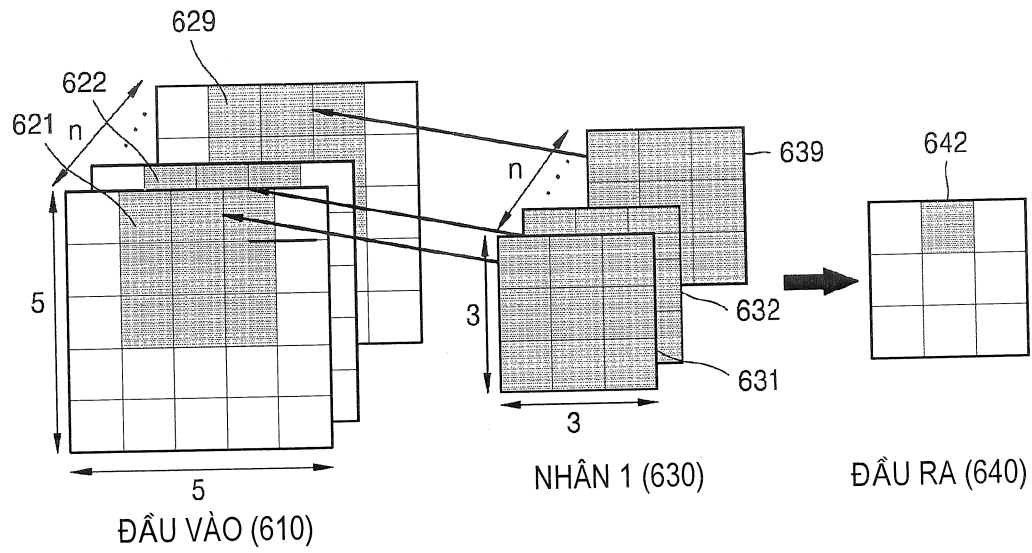
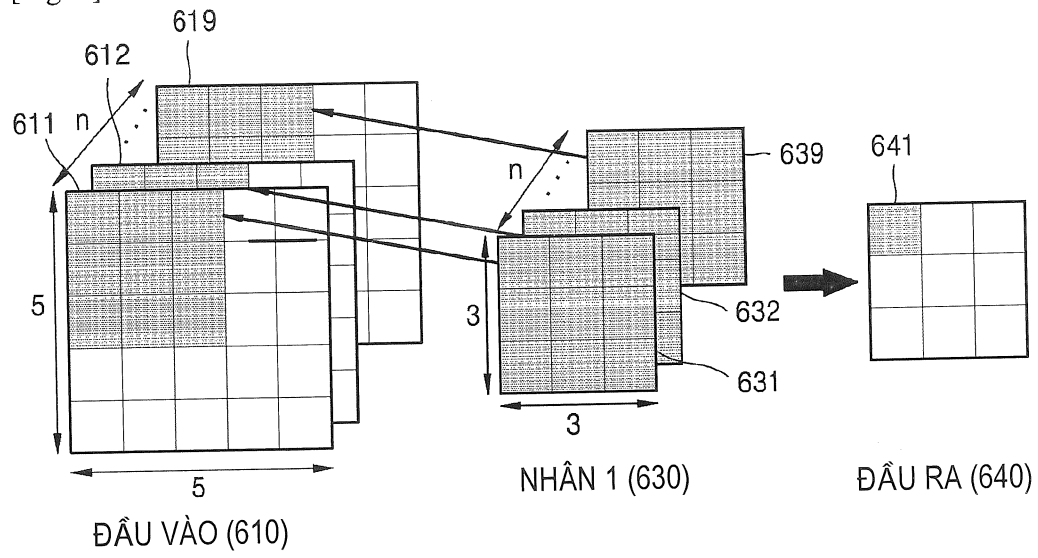


[Fig. 5]

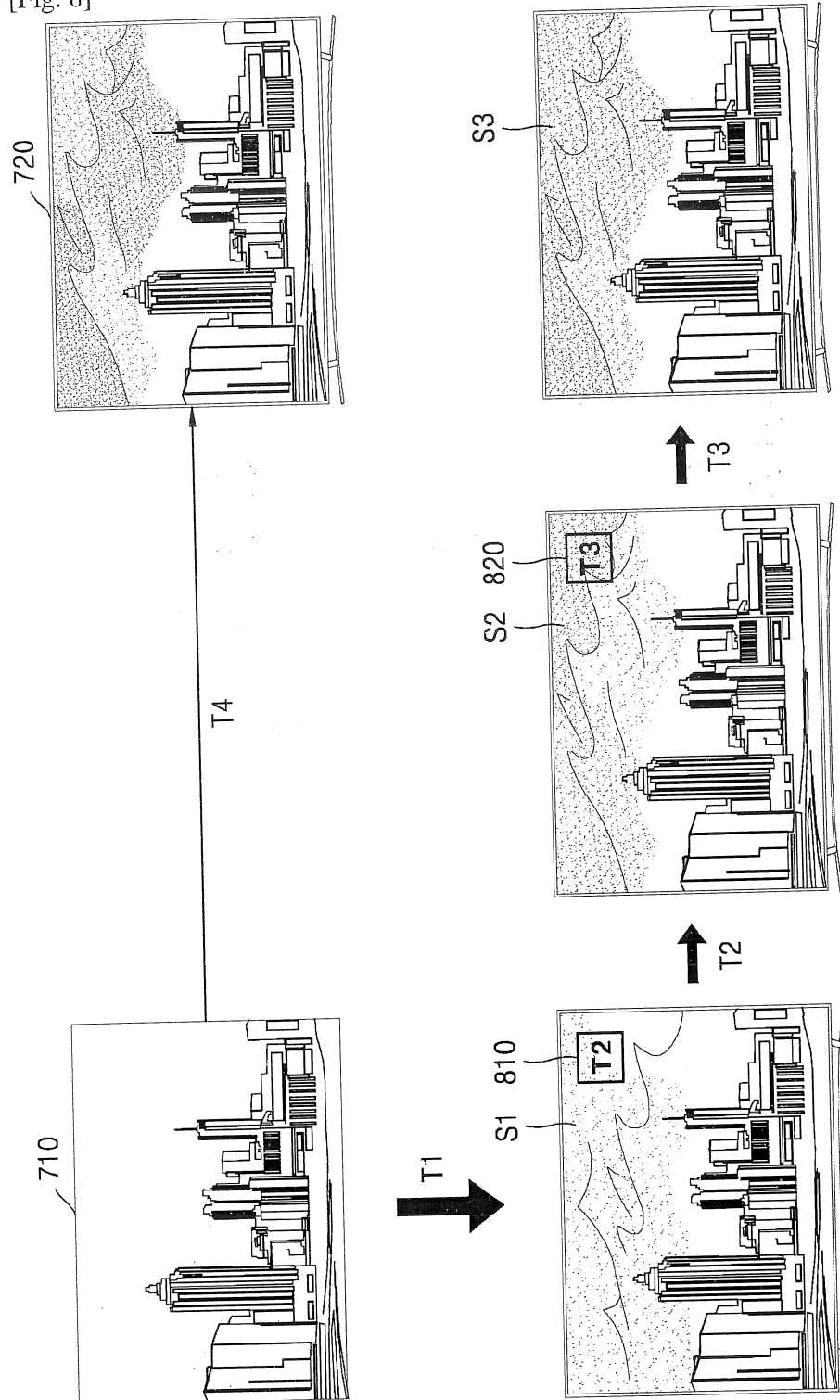
Tích chấp_1
NHÂN: (K_w, K_h, C, D)
(w, h, C) → (W, H, D)

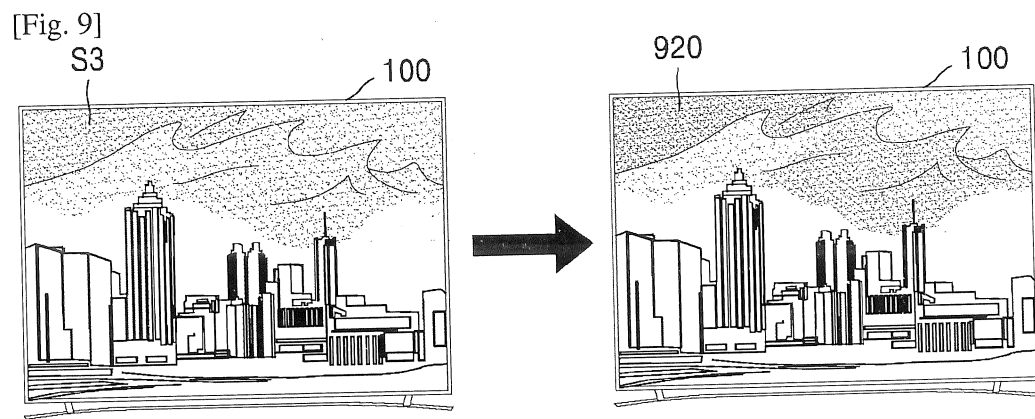


[Fig. 6]

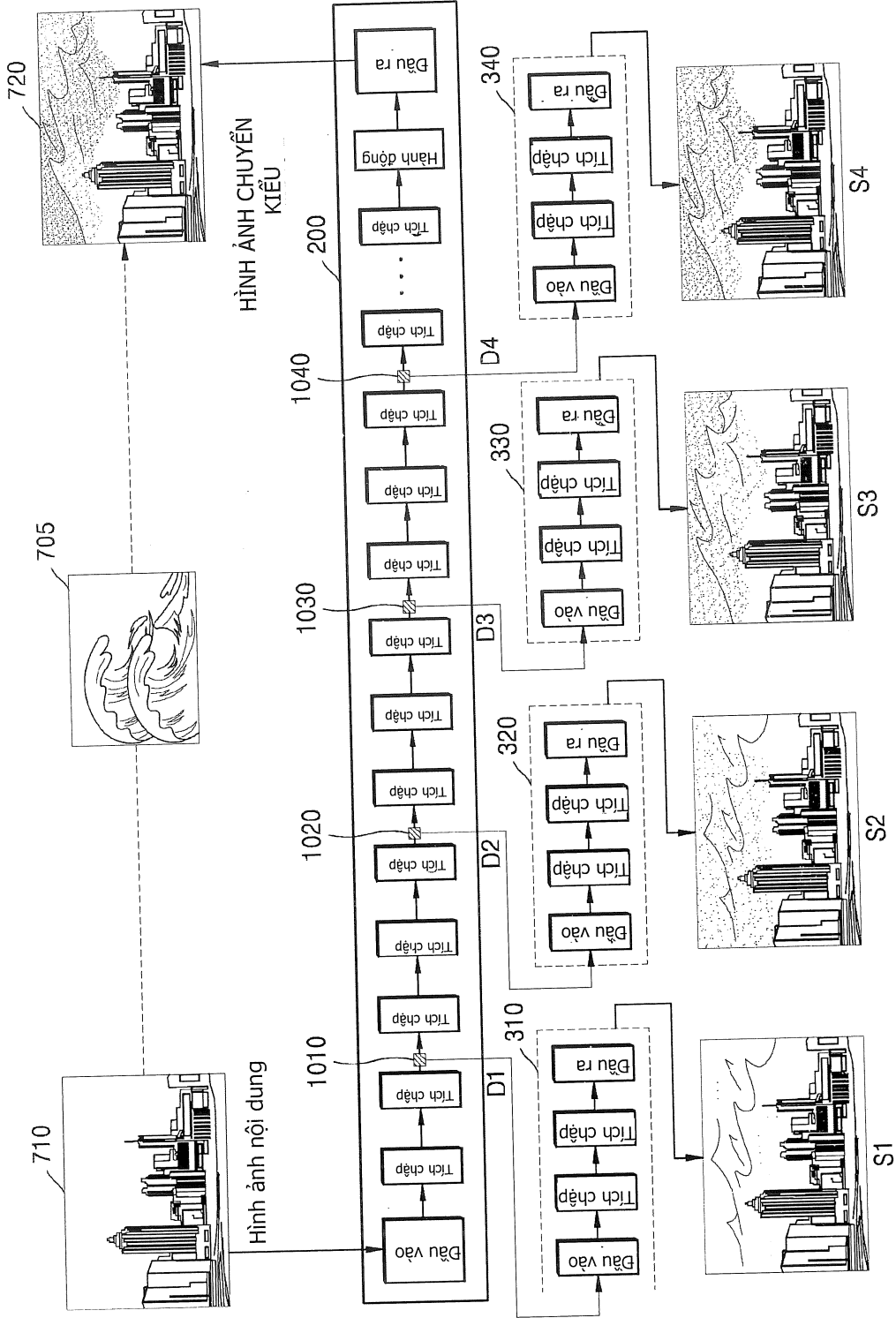


[Fig. 8]

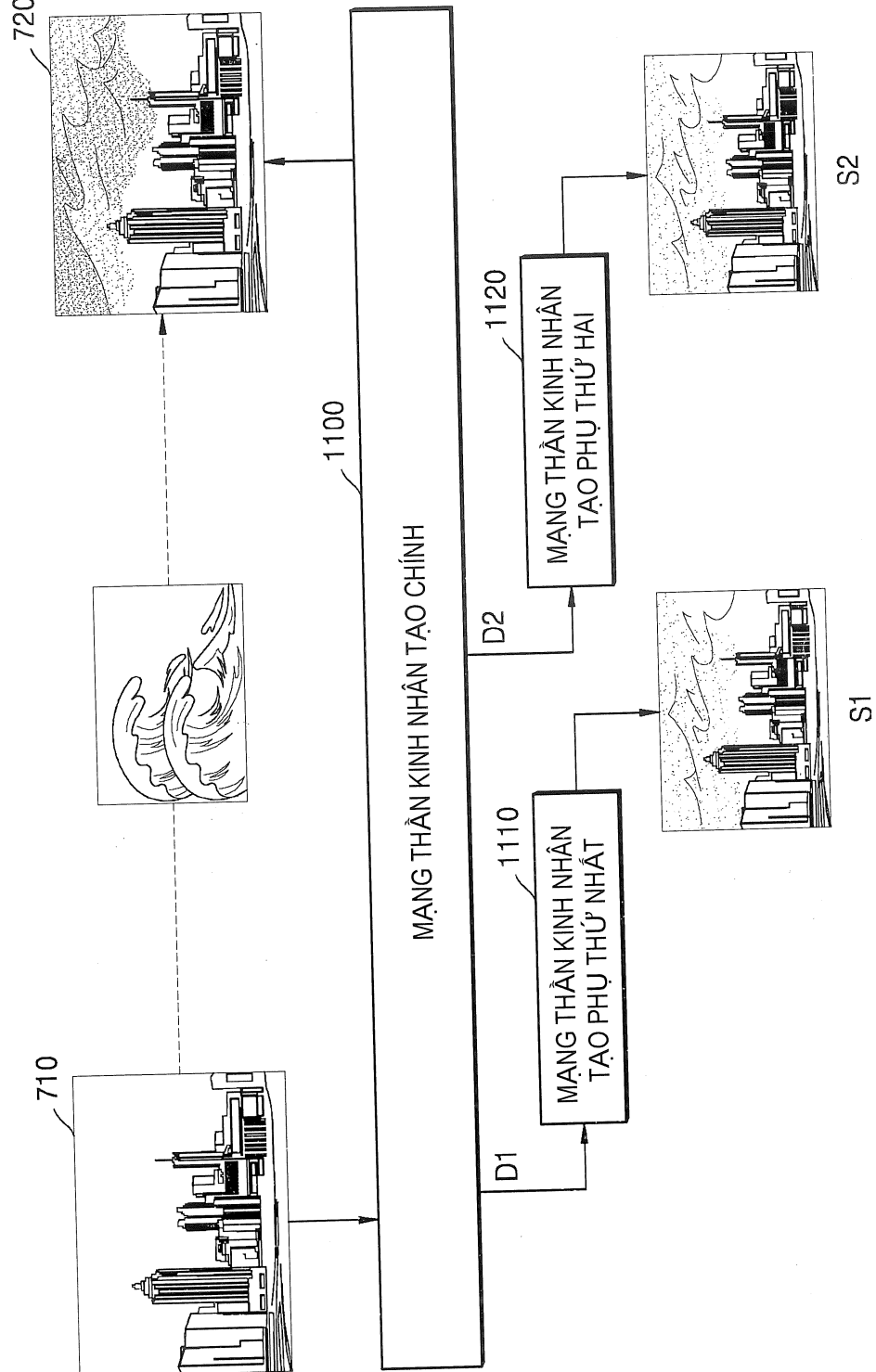




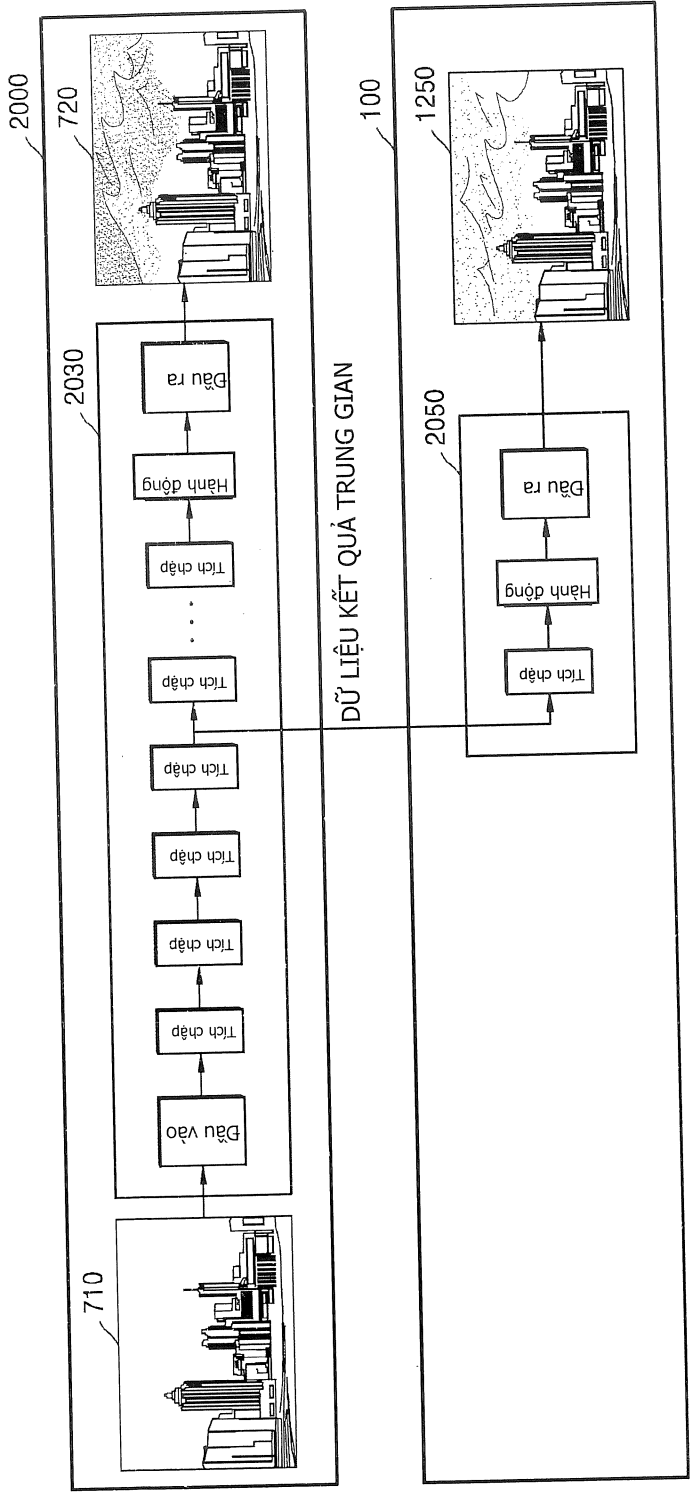
[Fig. 10]



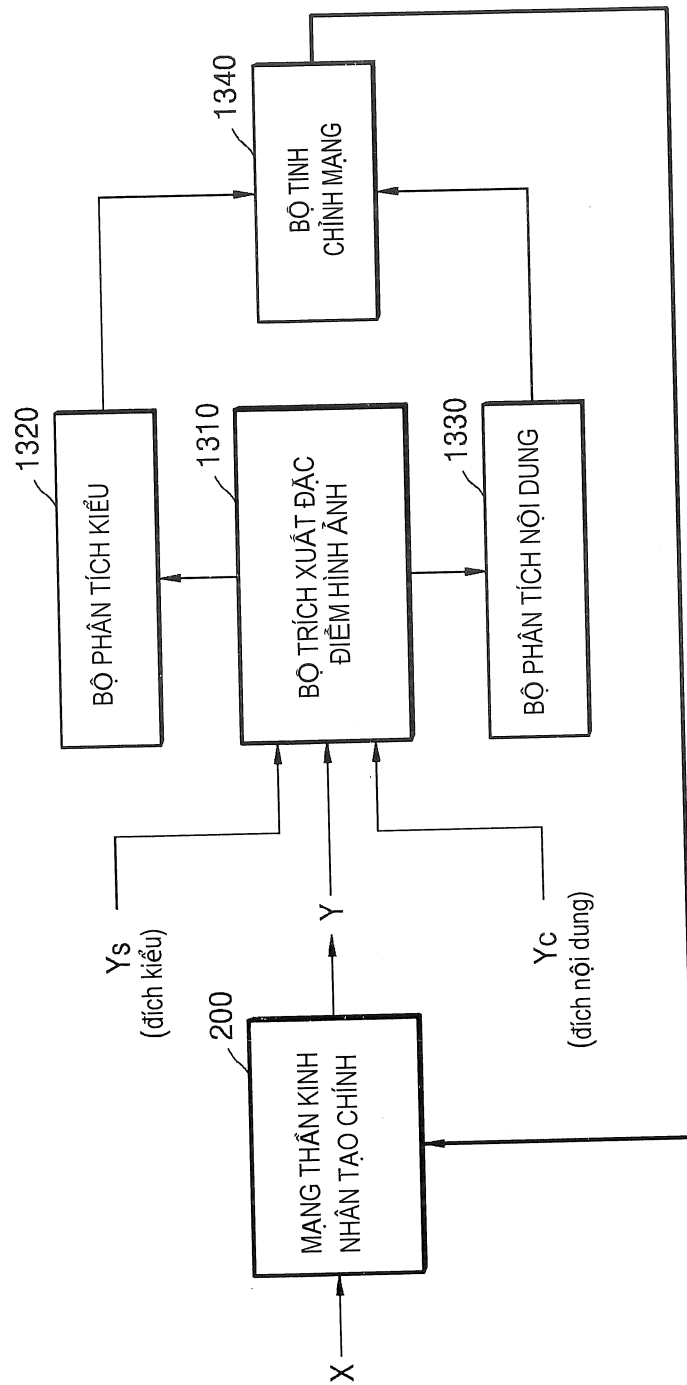
[Fig. 11]



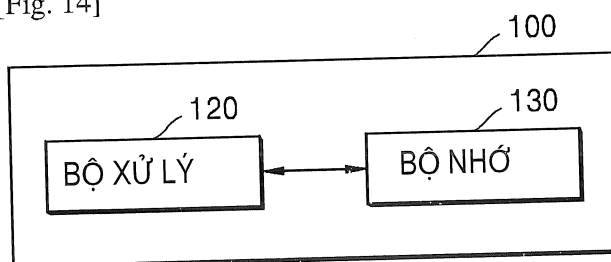
[Fig. 12]



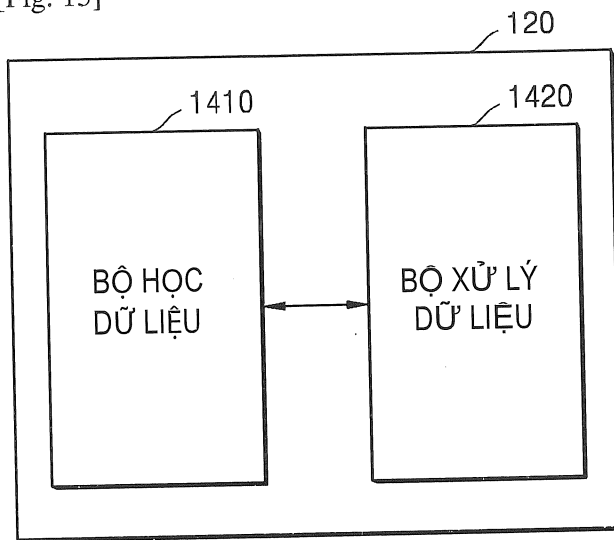
[Fig. 13]



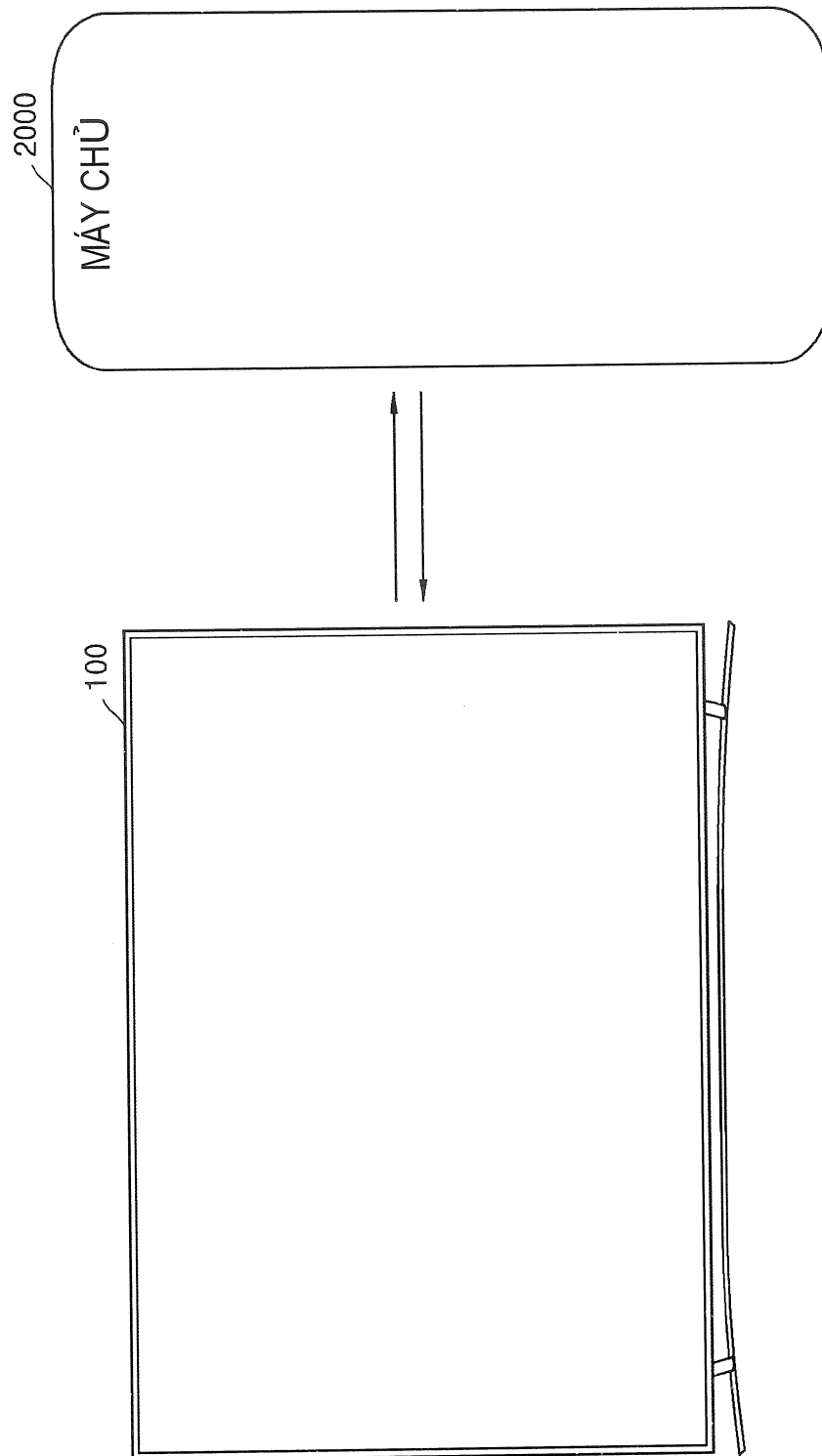
[Fig. 14]



[Fig. 15]



[Fig. 16]



[Fig. 17]

