



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

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

(11)



1-0032286

(51)<sup>7</sup> G06K 9/00; G06N 3/02

(13) B

(21) 1-2019-07080

(22) 13/12/2019

(30) 10201905273V 10/06/2019 SG

(45) 25/06/2022 411

(43) 25/12/2020 393

(73) Advanced New Technologies Co., Ltd. (KY)

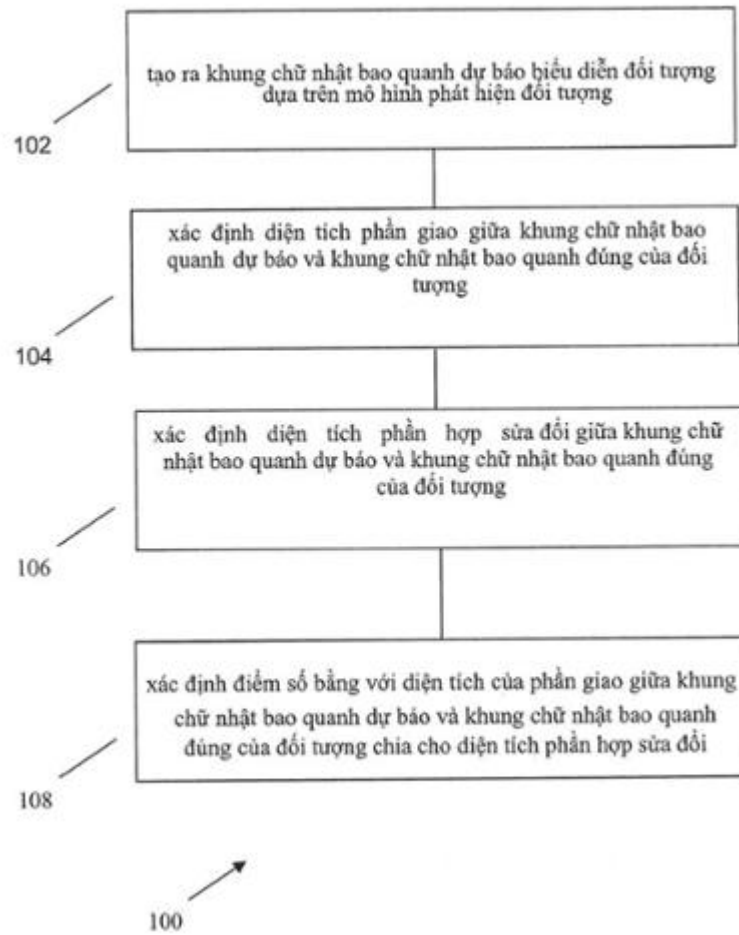
Cayman Corporate Centre, 27 Hospital Road, George Town, Grand Cayman KY1-9008, Cayman Islands

(72) Jiangbo HUANG (CN).

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

(54) PHƯƠNG PHÁP VÀ HỆ THỐNG ĐÁNH GIÁ MÔ HÌNH PHÁT HIỆN ĐỐI TƯỢNG

(57) Sáng chế đề xuất phương pháp, hệ thống và thiết bị đánh giá hiệu quả hoạt động của mô hình phát hiện đối tượng. Phương pháp đánh giá hiệu quả hoạt động của mô hình phát hiện đối tượng bao gồm bước tạo ra khung chữ nhật bao quanh được dự đoán biểu diễn đối tượng dựa trên mô hình phát hiện đối tượng. Đối tượng được định vị gần một hoặc nhiều đối tượng liền kề. Phương pháp cũng bao gồm bước xác định diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng, và xác định diện tích phần hợp sửa đổi giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng. Xác định diện tích phần hợp sửa đổi bao gồm xác định diện tích phần hợp bình quân gia quyền giữa các khung chữ nhật bao quanh được dự đoán và đúng của đối tượng dựa trên một hoặc nhiều trọng số, và thêm vào diện tích phần hợp bình quân gia quyền diện tích của phần giao giữa khung chữ nhật bao quanh được dự đoán và ít nhất một khung chữ nhật bao quanh đúng của một hoặc nhiều đối tượng liền kề. Phương pháp còn bao gồm bước xác định điểm số bằng với diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng chia cho diện tích phần hợp sửa đổi. Điểm số này biểu diễn hiệu quả hoạt động của mô hình phát hiện đối tượng.



### **Lĩnh vực kỹ thuật được đề cập**

Sáng chế đề cập đến các phương pháp và hệ thống đánh giá mô hình phát hiện đối tượng và các phương pháp và thiết bị phát hiện đối tượng.

### **Tình trạng kỹ thuật của sáng chế**

Nhận dạng ký tự quang học (Optical Character Recognition, OCR) là sự chuyển đổi cơ học hoặc điện tử của các hình ảnh được đánh máy, ký tự viết tay hoặc được in thành ký tự được mã hóa bằng máy, dù từ tài liệu được quét, ảnh của tài liệu, ảnh màn hình hoặc ký tự tiêu đề có được chồng lên trên hình ảnh v.v. hay không. Để nhận biết ký tự, bước thứ nhất là phát hiện các khung chữ nhật bao quanh của từng đoạn văn bản. Các thuật toán để phát hiện ký tự thuộc về một lĩnh vực có tên là phát hiện đối tượng trong thị giác máy tính (object detection in computer vision).

Trong phát hiện đối tượng, hàm IoU (Intersection over Union) là số liệu tiêu chuẩn phổ biến được dùng để đánh giá độ chính xác của sự lựa chọn bộ phát hiện và mô hình. Hàm IoU truyền thống được xác định là “diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng chia cho diện tích phần hợp giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng”. Hàm này hoạt động tốt xcho hầu hết các trường hợp. Tuy nhiên, trong trường hợp phát hiện ký tự, IoU truyền thống có thể không chọn được các mô hình/thông số tốt nhất và do đó việc sử dụng IoU truyền thống có thể làm giảm đáng kể độ chính xác cuối của việc nhận dạng ký tự.

Ví dụ, hàm IoU truyền thống không tính đến ít nhất hai điều kiện, cụ thể là (1) cả phần giao nhỏ hơn và phần hợp lớn hơn đều dẫn đến IoU nhỏ hơn như nhau; tuy nhiên, để phát hiện ký tự phần giao nhỏ hơn kém hơn so với phần hợp lớn hơn bởi vì nó làm cho một số vùng ký tự bị mất và ảnh hưởng đến kết quả OCR tiếp theo, và (2) IoU truyền thống không xét đến phần giao giữa khung ký tự dự đoán và các khung ký tự đúng của đối tượng khác. Không tính đến hai điều kiện này, giá trị IoU cao hơn có thể không nhất thiết chỉ ra mô hình tốt hơn.

Do đó, cần phải cung cấp các phương pháp và các thiết bị có thể cải thiện việc đánh giá các mô hình phát hiện đối tượng để phát hiện ký tự.

### **Bản chất kỹ thuật của sáng chế**

Khía cạnh thứ nhất của sáng chế đề xuất phương pháp đánh giá hiệu quả hoạt động của mô hình phát hiện đối tượng. Phương pháp này bao gồm các bước: tạo ra khung chữ nhật bao quanh được dự đoán biểu diễn đối tượng dựa trên mô hình phát hiện đối tượng, trong đó đối tượng được định vị gần một hoặc nhiều đối tượng liền kề; xác định diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng; và xác định diện tích phần hợp sửa đổi giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng. Bước xác định diện tích phần hợp sửa đổi bao gồm bước xác định diện tích phần hợp bình quân gia quyền giữa các khung chữ nhật bao quanh được dự đoán và đúng của đối tượng dựa trên một hoặc nhiều trọng số; và thêm vào diện tích phần hợp bình quân gia quyền diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và ít nhất một khung chữ nhật bao quanh đúng của một hoặc nhiều đối tượng liền kề. Phương pháp này còn bao gồm bước xác định điểm số bằng với diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng chia cho diện tích phần hợp sửa đổi. Điểm số này biểu diễn hiệu quả hoạt động của mô hình phát hiện đối tượng.

Khía cạnh thứ hai của sáng chế đề xuất phương pháp phát hiện đối tượng. Phương pháp này bao gồm các bước: thu dữ liệu thực chưa dán nhãn và các mô hình phát hiện đối tượng, trong đó các mô hình phát hiện đối tượng được tạo ra bởi mạng nơron dựa trên dữ liệu huấn luyện được dán nhãn; đánh giá từng mô hình phát hiện đối tượng với dữ liệu hợp lệ bằng cách sử dụng phương pháp như được xác định trong khía cạnh thứ nhất để nhận biết mô hình phát hiện đối tượng có điểm số cao nhất; và phát hiện đối tượng trong dữ liệu thực chưa dán nhãn dựa trên mô hình phát hiện đối tượng đã nhận biết.

Khía cạnh thứ ba của sáng chế đề xuất hệ thống đánh giá mô hình phát hiện đối tượng. Hệ thống này bao gồm bộ xử lý và bộ nhớ đọc được bằng máy tính được nối với bộ xử lý này và có lưu trữ các lệnh trên đó. Các lệnh này được thực thi bởi

bộ xử lý để tạo ra khung chữ nhật bao quanh được dự đoán biểu diễn đối tượng dựa trên mô hình phát hiện đối tượng, trong đó đối tượng được định vị gần một hoặc nhiều đối tượng liền kề; xác định diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng; và xác định diện tích phần hợp sửa đổi giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng. Diện tích phần hợp sửa đổi bằng tổng diện tích phần hợp bình quân gia quyền giữa các khung chữ nhật bao quanh được dự đoán và đúng của đối tượng dựa trên một hoặc nhiều trọng số; và diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và ít nhất một khung chữ nhật bao quanh đúng của một hoặc nhiều đối tượng liền kề. Các lệnh cũng được thực thi bằng bộ xử lý để xác định điểm số bằng với diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng chia cho diện tích phần hợp sửa đổi. Điểm số này biểu diễn hiệu quả hoạt động của mô hình phát hiện đối tượng.

Khía cạnh thứ tư của sáng chế đề xuất thiết bị bao gồm môđun phát hiện đối tượng được tạo cấu hình để tạo ra khung chữ nhật bao quanh được dự đoán biểu diễn đối tượng dựa trên mô hình phát hiện đối tượng, trong đó đối tượng được định vị gần một hoặc nhiều đối tượng liền kề, và môđun đánh giá. Môđun đánh giá này được tạo cấu hình để xác định diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng; và xác định diện tích phần hợp sửa đổi giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng. Diện tích phần hợp sửa đổi bằng tổng diện tích phần hợp bình quân gia quyền giữa các khung chữ nhật bao quanh được dự đoán và đúng của đối tượng dựa trên một hoặc nhiều trọng số; và diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và ít nhất một khung chữ nhật bao quanh đúng của một hoặc nhiều đối tượng liền kề. Môđun đánh giá còn được tạo cấu hình để xác định điểm số bằng với diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng chia cho diện tích phần hợp sửa đổi; và đánh giá hiệu quả hoạt động của mô hình phát hiện đối tượng dựa trên điểm số.

Khía cạnh thứ năm của sáng chế đề xuất bộ phát hiện đối tượng. Bộ phát hiện đối tượng này bao gồm môđun thu được tạo cấu hình để thu dữ liệu thực chưa dán nhãn và các mô hình phát hiện đối tượng. Các mô hình phát hiện đối tượng được tạo ra bởi mạng nơron dựa trên dữ liệu huấn luyện được dán nhãn. Bộ phát hiện đối tượng này cũng bao gồm các thiết bị như được xác định theo khía cạnh thứ tư được nối với bộ môđun thu và được tạo cấu hình để đánh giá từng mô hình phát hiện đối tượng với dữ liệu hợp lệ để nhận biết mô hình phát hiện đối tượng có điểm số cao nhất. Môđun phát hiện đối tượng này còn được tạo cấu hình để phát hiện đối tượng trong dữ liệu thực chưa dán nhãn dựa trên mô hình phát hiện đối tượng đã nhận biết.

### **Mô tả vắn tắt các hình vẽ**

Các phương án theo sáng chế sẽ được hiểu rõ hơn và rõ ràng với người có hiểu biết trung bình về lĩnh vực kỹ thuật liên quan khi xem xét đến phần mô tả sau đây, chỉ theo cách làm ví dụ, và có liên hệ với các hình vẽ, trong đó:

Fig.1 là lưu đồ minh họa phương pháp đánh giá mô hình phát hiện đối tượng theo phương án.

Fig.2a và Fig.2b là lược đồ minh họa phương án thực hiện để xác định diện tích phần hợp sửa đổi theo phương án trên Fig.1.

Fig.3 là lưu đồ minh họa phương pháp phát hiện đối tượng theo phương án.

Fig.4 là lược đồ minh họa bộ phát hiện đối tượng theo phương án.

Fig.5 là lược đồ minh họa hệ thống máy tính đồng bộ trong phương án thực hiện các phương pháp trên Fig.1 và Fig.3 và bộ phát hiện đối tượng trên Fig.4.

Người có hiểu biết trung bình về lĩnh vực kỹ thuật liên quan sẽ hiểu rõ là các thành phần trong các hình vẽ được minh họa để đơn giản và rõ ràng và không nhất thiết phải được mô tả đúng tỷ lệ. Ví dụ, các kích thước của một số phần tử trong các minh họa, các sơ đồ khối hoặc các lưu đồ có thể được phóng to đối với các thành phần khác để giúp hiểu rõ hơn về các phương án hiện tại.

### **Mô tả chi tiết sáng chế**

Sáng chế đề xuất các phương pháp và các thiết bị trong đó mô hình phát hiện đối tượng được đánh giá dựa trên hàm IoU cải tiến có tính đến các sự cân nhắc như (1) để phát hiện ký tự, phần giao nhỏ hơn kém hơn so với phần hợp lớn hơn bởi vì

nó có thể khiến cho một số vùng của ký tự bị mất và ảnh hưởng đến kết quả OCR tiếp theo, và (2) để phát hiện ký tự, phần giao giữa khung ký tự dự đoán và các khung ký tự đúng của đối tượng khác có thể cũng tác động ngược đến kết quả OCR tiếp theo. Như được mô tả chi tiết hơn ở dưới đây, sự đánh giá dựa trên điểm số không chỉ phạt đối với kết hợp phần giao nhỏ hơn và phần hợp lớn hơn (mà hàm IoU truyền thống thực hiện), mà đồng thời phạt nhiều hơn đối với phần giao nhỏ hơn so với phần hợp lớn hơn. Ngoài ra, nó phạt phần giao giữa khung ký tự dự đoán và các khung ký tự đúng của đối tượng khác. Theo đó, với một số mô hình phát hiện đối tượng khác nhau, có thể nhận biết hoặc lựa chọn mô hình phù hợp nhất để phát hiện ký tự.

Các phương án sẽ được mô tả, theo cách làm ví dụ, tham khảo đến các hình vẽ. Các số tham chiếu và các ký tự giống nhau trên các hình vẽ chỉ các phần tử giống nhau hoặc tương đương.

Một số phần của mô tả trong bản mô tả này được trình bày rõ hoặc ngầm định về mặt các thuật toán và chức năng hoặc các biểu tượng đại diện của các hoạt động trên dữ liệu bên trong bộ nhớ máy tính. Các mô tả thuật toán và chức năng hoặc các biểu tượng đại diện này là phương tiện được sử dụng bởi người có hiểu biết trung bình về lĩnh vực kỹ thuật xử lý dữ liệu để truyền đạt có hiệu quả nhất thực chất công việc của họ đến người có hiểu biết trung bình về lĩnh vực kỹ thuật liên quan khác. Thuật toán ở đây và nói chung, được hình thành là chuỗi các bước tự thống nhất dẫn đến kết quả mong muốn. Các bước là những bước đòi hỏi các thao tác vật lý của định lượng vật lý như là các tín hiệu điện, từ hoặc quang có thể được lưu trữ, được phát, được truyền, được kết hợp, được so sánh, và được điều khiển theo cách khác.

Trừ khi được nêu cụ thể khác đi, và rõ ràng từ những điều sau đây, cần phải được hiểu rằng trong suốt bản mô tả này, có sử dụng các thuật ngữ như “quét”, “thu”, “tính toán”, “phát hiện”, “xác định”, “thay thế”, “tạo ra”, “khởi tạo”, “kết xuất”, “đánh giá”, hoặc tương tự, đề cập đến hoạt động hoặc quy trình của hệ thống máy tính, hoặc thiết bị điện tử tương tự, mà các thao tác và chuyển dữ liệu được biểu thị dưới dạng lượng vật lý bên trong hệ thống máy tính thành dữ liệu khác tương tự được biểu thị dưới dạng lượng vật lý bên trong hệ thống máy tính hoặc vật lưu trữ thông tin khác, các thiết bị phát hoặc hiển thị.

Sáng chế cũng bộc lộ thiết bị để biểu diễn các hoạt động của các phương pháp. Thiết bị như vậy có thể được tạo kết cấu đặc biệt cho các mục đích cần thiết, hoặc có thể bao gồm máy tính hoặc thiết bị khác được kích hoạt hoặc được tái tạo cấu hình chọn lọc bởi chương trình máy tính được lưu trữ trong máy tính. Các thuật toán và hiển thị được trình bày ở đây không có liên hệ với bất kỳ máy tính hoặc thiết bị cụ thể nào khác. Các máy khác nhau có thể được sử dụng với các chương trình theo các hướng dẫn trong bản mô tả này. Theo cách khác, kết cấu của thiết bị chuyên dụng hơn để thực hiện các bước của phương pháp cần thiết có thể thích hợp. Kết cấu của máy tính phù hợp để thực thi các phương pháp/quy trình khác nhau được mô tả trong bản mô tả này sẽ xuất hiện từ phần mô tả trong bản mô tả này.

Ngoài ra, sáng chế cũng ngầm bộc lộ chương trình máy tính, theo đó nó sẽ trở nên rõ ràng đối với người có hiểu biết trung bình về lĩnh vực kỹ thuật liên quan rằng các bước riêng lẻ của phương pháp được mô tả trong bản mô tả này có thể có hiệu lực nhờ mã chương trình máy tính. Chương trình máy tính không nhằm giới hạn bất kỳ ngôn ngữ lập trình và phương án thực hiện cụ thể nào của nó. Cần hiểu rằng nhiều ngôn ngữ lập trình và mã hóa của chúng có thể được dùng để thực hiện các hướng dẫn của sáng chế có trong bản mô tả này. Ngoài ra, chương trình máy tính không nhằm giới hạn trong bất kỳ kỹ thuật điều khiển cụ thể nào. Có nhiều biến thể khác của chương trình máy tính, có thể sử dụng các kỹ thuật điều khiển khác nhau mà không vượt ra khỏi bản chất và phạm vi của sáng chế.

Ngoài ra, một hoặc nhiều bước chương trình máy tính có thể thực hiện song song thay vì theo trình tự. Chương trình máy tính như vậy có thể được lưu trữ trên vật ghi đọc được bằng máy tính bất kỳ. Vật ghi đọc được bằng máy tính có thể bao gồm các thiết bị lưu trữ như là các đĩa từ hoặc đĩa quang, chip nhớ, hoặc các thiết bị lưu trữ khác phù hợp để giao tiếp với máy tính. Vật ghi đọc được bằng máy tính có thể cũng bao gồm vật ghi nổi cứng như được lấy làm ví dụ trong hệ thống Internet, hoặc vật ghi không dây như được lấy làm ví dụ trong hệ thống các điện thoại di động GSM, GPRS, 3G hoặc 4G, cũng như các hệ thống không dây khác như là Bluetooth, ZigBee, Wi-Fi. Chương trình máy tính khi được tải và thực thi trên một máy tính như vậy có hiệu quả trong thiết bị triển khai các bước của phương pháp được ưu tiên.

Sáng chế có thể cũng được triển khai ở dạng môđun phần cứng. Cụ thể hơn là, trong cảm biến phần cứng, môđun là bộ phận phần cứng chức năng được thiết kế để sử dụng với các bộ phận hoặc môđun khác. Ví dụ, môđun có thể được triển khai sử dụng các bộ phận điện tử rời rạc, hoặc nó có thể tạo ra một phần của toàn bộ mạch điện như là mạch điện tích hợp ứng dụng đặc thù (Application Specific Integrated Circuit, ASIC) mảng cổng lập trình được dạng trường (Field Programmable Gate Array, FPGA). Tồn tại vô số các khả năng khác. Người có hiểu biết trung bình về lĩnh vực kỹ thuật liên quan sẽ đánh giá rằng hệ thống cũng có thể được triển khai ở dạng kết hợp của các môđun phần cứng và phần mềm.

Theo các phương án khác nhau, "mạch" có thể được hiểu là bất kỳ loại đối tượng thực thi logic nào, có thể là mạch mục đích đặc biệt hoặc phần mềm thực thi bộ xử lý được lưu trữ trong bộ nhớ, phần sụn, hoặc tổ hợp của chúng. Do đó, theo một phương án, "mạch" có thể là mạch logic nổi cứng hoặc mạch logic có thể lập trình như là bộ xử lý có thể lập trình, ví dụ bộ xử lý cực nhỏ (ví dụ, bộ xử lý máy tính với tập lệnh phức tạp (Complex Instruction Set Computer, CISC) hoặc bộ xử lý máy tính với tập lệnh đơn giản hóa (Reduced Instruction Set Computer, RISC)). "Mạch" có thể cũng là phần mềm thực thi bộ xử lý, ví dụ, bất kỳ loại chương trình máy tính nào, ví dụ, chương trình máy tính sử dụng mã máy ảo ví dụ, Java. Bất kỳ loại phương án thực hiện nào khác của các chức năng tương ứng có thể được mô tả chi tiết hơn ở đây có thể cũng được hiểu là "mạch" theo phương án thay thế.

Fig.1 là lưu đồ minh họa 100 minh họa phương pháp đánh giá hiệu quả hoạt động của mô hình phát hiện đối tượng theo phương án.

Tại bước 102, khung chữ nhật bao quanh được dự đoán biểu diễn đối tượng được tạo ra dựa trên mô hình phát hiện đối tượng. Đối tượng thông thường được định vị gần một hoặc nhiều đối tượng liền kề. Ví dụ, đối tượng có thể là chuỗi ký tự (ví dụ từ hoặc cụm từ) gần với các chuỗi ký tự khác (ví dụ các từ và cụm từ khác ở trong câu hoặc đoạn văn). Các thông số như là kích thước và vị trí của khung chữ nhật bao quanh được dự đoán có thể thay đổi từ mô hình phát hiện đối tượng sang mô hình phát hiện đối tượng khác. Hiệu quả hoạt động của mô hình phát hiện đối tượng do đó phụ thuộc vào khung chữ nhật bao quanh được dự đoán.

Tại bước 104, diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng được xác định. Thông thường, diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng lớn hơn biểu thị mức độ chồng lên nhau lớn hơn giữa các khung chữ nhật bao quanh và có thể được mong muốn hơn so với diện tích nhỏ hơn.

Tại bước 106, diện tích phần hợp sửa đổi giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng được xác định bằng cách (1) xác định diện tích phần hợp bình quân gia quyền giữa các khung chữ nhật bao quanh được dự đoán và đúng của đối tượng dựa trên một hoặc nhiều trọng số và (2) thêm vào diện tích phần hợp bình quân gia quyền, diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và ít nhất một khung chữ nhật bao quanh đúng của một hoặc nhiều đối tượng liên kề.

Như được minh họa kỹ dưới đây tham khảo Fig.2a, do diện tích phần hợp giữa hai đối tượng là tổng diện tích phần giao và diện tích của phần không giao nhau của chúng, theo các ví dụ hiện tại, diện tích phần hợp bình quân gia quyền giữa các khung chữ nhật bao quanh được dự đoán và đúng của đối tượng có thể được xác định bằng cách chỉ định các trọng số thích hợp đối với các diện tích cấu thành tương ứng, bao gồm diện tích của phần không giao nhau giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng.

Diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và ít nhất một khung chữ nhật bao quanh đúng của một hoặc nhiều đối tượng liên kề được minh họa thêm tham chiếu đến Fig.2b.

Kết quả thực của bước 106 là diện tích phần hợp sửa đổi lớn hơn nếu khung chữ nhật bao quanh được dự đoán không che phủ tốt khung chữ nhật bao quanh đúng của đối tượng và/hoặc che phủ với (các) khung chữ nhật bao quanh đúng của (các) đối tượng liên kề. Thông thường, diện tích phần hợp sửa đổi lớn hơn là ít được mong muốn so với diện tích phần hợp nhỏ hơn.

Tại bước 108, điểm số bằng với diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng chia cho diện tích phần hợp sửa đổi được xác định. Điểm số này biểu diễn hiệu quả hoạt

động của mô hình phát hiện đối tượng. Ví dụ, nếu mô hình phát hiện đối tượng làm phát sinh diện tích phân hợp sửa đổi lớn hơn, thì điểm số sẽ thấp hơn. Mặt khác, nếu mô hình phát hiện đối tượng làm phát sinh diện tích phân hợp sửa đổi nhỏ hơn, thì điểm số sẽ cao hơn. Điểm số gần với 1 biểu thị độ chính xác phát hiện đối tượng lớn hơn.

Fig.2a và Fig.2b là lược đồ minh họa phương án thực hiện để xác định diện tích phân hợp sửa đổi theo phương án trên Fig.1. Theo phương án thực hiện này, đối tượng và các đối tượng liền kề là các chuỗi ký tự có chứa các ký tự và các chữ số, và các khung chữ nhật bao quanh được hoán đổi cho các khung ký tự. Tuy nhiên, cần phải được hiểu bởi người có hiểu biết trung bình về lĩnh vực kỹ thuật liên quan rằng việc phát hiện ký tự chỉ là một ví dụ, và rằng phương pháp có thể được áp dụng cho các dạng phát hiện đối tượng khác.

Tham khảo Fig.2a và Fig.2b, khung chữ nhật bao quanh được dự đoán có thể được dán nhãn như tập hợp A, khung chữ nhật bao quanh đúng của đối tượng có thể được dán nhãn như tập hợp B, và các khung chữ nhật bao quanh đúng của đối tượng của các đối tượng liền kề có thể được dán nhãn như các tập hợp C và D tương ứng. Được cung cấp dữ liệu hợp lệ được dán nhãn, các kích thước và vị trí của B, C và D được biết, trong khi kích thước và vị trí của A có thể thay đổi dựa trên mô hình phát hiện đã sử dụng. Khi các khung chữ nhật bao quanh của hai đối tượng liền kề được biểu diễn ở ví dụ này, cần phải được hiểu rằng phép tính có thể được áp dụng cho số lượng lớn các đối tượng liền kề.

Diện tích phân hợp sửa đổi, *modified (AUB)*, trong bước 106 của Fig.1 có thể được biểu diễn dưới dạng:

$$\text{modified (AUB)} = \text{weighted (AUB)} + A \cap C + A \cap D \quad (I)$$

trong đó *weighted (AUB)* biểu diễn diện tích phân hợp bình quân gia quyền giữa các tập hợp A và B,  $A \cap C$  biểu diễn phần giao giữa các tập hợp A và C, và  $A \cap D$  biểu diễn phần giao giữa các tập hợp A và D. Các diện tích tương ứng với  $A \cap C$  và  $A \cap D$  được biểu diễn trên Fig.2b.

Tham khảo Fig.2a, trong khi diện tích phân hợp tiêu chuẩn là tổng của diện tích phần giao ( $A \cap B$ ) và diện tích của phần không giao ( $A1$  và  $B1$ ), giá trị *weighted (AUB)* theo phương trình (I) được xác định bằng cách xem xét  $B1$  lớn

hơn nghĩa là phần giao nhỏ hơn với khung chữ nhật bao quanh đúng của đối tượng, trong khi A1 lớn hơn dẫn đến phần hợp lớn hơn. Do đó, theo các phương án hiện tại, B1 bị phạt nhiều hơn A1.

Ví dụ,  $weighted(A \cup B)$  có thể được biểu thị như sau:

$$weighted(A \cup B) = w_1 A_1 + w_2 B_1 + (A \cap B) \quad (II)$$

Theo phương trình (II), các trọng số  $w_1$  và  $w_2$  được dùng để điều chỉnh các hình phạt cho A1 và B1. B1 lớn hơn dẫn đến phần giao nhỏ hơn, trong khi A1 lớn hơn dẫn đến phần hợp lớn hơn. Do đó, B1 nên bị phạt nhiều hơn, có nghĩa là  $w_1 < w_2$ . Giá trị chính xác của  $w_1$  và  $w_2$  có thể được lựa chọn dựa trên ví dụ dữ liệu thực tế và có thể thay đổi dựa theo tình hình. Theo một ví dụ không giới hạn,  $w_1 + w_2 = 2$ .

2. Các diện tích của A1, B1 và  $A \cap B$  tương ứng được biểu diễn trên Fig.2a.

Bằng cách kết hợp hai phương trình (I) và (II) ở trên:

$$modified(A \cup B) = w_1 A_1 + w_2 B_1 + (A \cap B) + A \cap C + A \cap D \quad (III)$$

Dựa vào bước 108 theo Fig.1, điểm số có thể được tính như sau:

$$\text{Điểm số} = \frac{A \cap B}{weighted(A \cup B)} = \frac{A \cap B}{w_1 A_1 + w_2 B_1 + (A \cap B) + A \cap C + A \cap D} \quad (IV)$$

Theo các phương trình (III) và (IV),  $w_1 < w_2$  và theo một ví dụ không giới hạn,  $w_1 + w_2 = 2$ .

Như có thể thấy từ phương trình (IV) ở trên, trong khi tử số giống như tử số được dùng để xác định IoU truyền thống, thì mẫu số sửa đổi để phạt (1) nhiều hơn ở phần giao giữa các tập hợp A và B nhỏ và ít hơn ở phần hợp giữa các tập hợp A và B lớn hơn (như là  $w_1 < w_2$ ), và (2) phần giao giữa các tập hợp A và C bất kỳ hoặc giữa các tập hợp A và D (như là các phần  $A \cap C$  và  $A \cap D$  đều khác không). Nói cách khác, điểm số thấp hơn nếu một hoặc cả hai điều kiện (1) và (2) biểu thị việc xảy ra chất lượng thấp của mô hình phát hiện đối tượng. Ngược lại, nếu phần giao giữa tập hợp A và B lớn, và không có phần giao giữa tập hợp A và C hoặc giữa tập hợp A và D, thì điểm số sẽ cao. Điểm số cao hơn (tức là gần với 1) thể hiện mô hình phát hiện đối tượng có độ chính xác tốt hơn.

Phương pháp đánh giá như được mô tả ở trên có thể được dùng để phát hiện đối tượng bằng cách nhận biết trước tiên mô hình thích hợp để phát hiện tác vụ

trước khi sử dụng mô hình này cho dữ liệu thực được phân tích. Fig.3 là lưu đồ minh họa 300 minh họa phương pháp phát hiện đối tượng theo phương án.

Tại bước 302, dữ liệu thực chưa dán nhãn và các mô hình phát hiện đối tượng được nhận. Các mô hình phát hiện đối tượng có thể được tạo ra bằng mạng noron, ví dụ mạng noron chập, dựa trên dữ liệu huấn luyện được dán nhãn. Dữ liệu thực chưa dán nhãn có thể dưới dạng dữ liệu ảnh của ảnh quang học của đối tượng. Ví dụ, ảnh quang học có thể là tài liệu được quét hoặc hình ảnh của tài liệu được tải lên bởi khách hàng có chứa thông tin gốc mà khách hàng muốn cung cấp.

Tại bước 304, với dữ liệu hợp lệ được dán nhãn đã cho, mỗi một mô hình phát hiện đối tượng được đánh giá sử dụng phương pháp như được mô tả ở trên, tham khảo Fig.1 và Fig.2a, Fig.2b, để nhận biết mô hình phát hiện đối tượng có điểm số cao nhất. Tại bước 306, mô hình phát hiện đối tượng đã nhận biết được dùng để phát hiện đối tượng trong dữ liệu thực chưa dán nhãn.

Theo phương án của phương pháp phát hiện đối tượng theo Fig.3, các mô hình phát hiện đối tượng trong bước 302 có thể được liên hợp các siêu tham số, và mô hình phát hiện đối tượng với các siêu tham số được liên hợp có điểm số cao nhất được nhận biết trong bước 306. Các mô hình đã được nhận biết và các siêu tham số được liên hợp sau đó được dùng để phát hiện đối tượng trong dữ liệu thực chưa dán nhãn.

Fig.4 là lược đồ minh họa bộ phát hiện đối tượng 400 theo phương án. Bộ phát hiện đối tượng bao gồm môđun thu 402 được nối với thiết bị 404. Thiết bị 404 bao gồm môđun phát hiện đối tượng 406 và môđun đánh giá 408. Thông thường, môđun đánh giá 408 được cung cấp với tập hợp dữ liệu hợp lệ được dán nhãn 410. Môđun thu 402 được tạo cấu hình để thu dữ liệu thực chưa dán nhãn và các mô hình phát hiện đối tượng. Các mô hình phát hiện đối tượng được tạo ra bởi mạng noron dựa trên dữ liệu huấn luyện được dán nhãn. Thiết bị 404 có thể đánh giá từng mô hình phát hiện đối tượng dựa trên dữ liệu hợp lệ để nhận biết mô hình phát hiện đối tượng có điểm số cao nhất.

Theo ví dụ, môđun phát hiện đối tượng 406 được tạo cấu hình để tạo ra khung chữ nhật bao quanh được dự đoán biểu diễn đối tượng dựa trên mô hình phát hiện đối tượng. Đối tượng được định vị gần một hoặc nhiều đối tượng liền kề. Môđun

đánh giá 408 được tạo cấu hình để xác định diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng; và xác định diện tích phần hợp sửa đổi giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng. Diện tích phần hợp sửa đổi bằng tổng diện tích phần hợp bình quân gia quyền giữa các khung chữ nhật bao quanh được dự đoán và đúng của đối tượng dựa trên một hoặc nhiều trọng số và diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và ít nhất một khung chữ nhật bao quanh đúng của một hoặc nhiều đối tượng liên kề. Môđun đánh giá 408 còn được tạo cấu hình để xác định điểm số bằng với diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng chia cho diện tích phần hợp sửa đổi; và đánh giá hiệu quả hoạt động của mô hình phát hiện đối tượng dựa trên điểm số. Các bước được thực hiện cho từng mô hình phát hiện đối tượng để nhận biết mô hình phát hiện đối tượng có điểm số cao nhất.

Môđun phát hiện đối tượng 406 sau đó có thể phát hiện đối tượng trong dữ liệu thực chưa dán nhãn dựa trên mô hình phát hiện đối tượng đã nhận biết.

Các phương pháp, hệ thống và thiết bị như được mô tả khi được áp dụng để phát hiện ký tự có thể cải thiện chất lượng lựa chọn mô hình trong phát hiện khung ký tự và tăng độ chính xác cuối của OCR. Mô hình phát hiện cung cấp phần giao giữa khung ký tự dự đoán và khung ký tự đúng của đối tượng của chuỗi ký tự mục tiêu nhỏ xuất dữ liệu có thể không đầy đủ hoặc bị cắt xén. Tương tự, mô hình phát hiện cung cấp một số phạm vi của phần giao giữa khung ký tự dự đoán và các khung ký tự đúng của đối tượng của các chuỗi ký tự liên kề xuất dữ liệu có thể nhiều hoặc không đúng. Theo các phương án hiện tại, các mô hình phát hiện như vậy có điểm số đánh giá thấp và sẽ không được lựa chọn. Để thay thế, mô hình cung cấp phần giao giữa khung ký tự dự đoán và khung ký tự đúng của đối tượng của chuỗi ký tự mục tiêu lớn, và không có phần giao giữa khung ký tự dự đoán và các khung ký tự đúng của đối tượng của các chuỗi ký tự liên kề, có điểm số đánh giá cao và sẽ được lựa chọn. Khi được dùng để phát hiện các khung ký tự trên ảnh quang học thật/có thực, mô hình được lựa chọn có thể xuất dữ liệu chính xác có thể giúp thay đổi độ chính xác của các bước OCR tiếp theo.

Fig.5 mô tả thiết bị điện toán mẫu 500, sau đây được gọi là hệ thống máy tính 500, trong đó một hoặc nhiều thiết bị điện toán 500 như vậy có thể được sử dụng cho bộ phát hiện đối tượng 400 theo Fig.4, hoặc cho phương án thực hiện một số hoặc tất cả các bước của các phương pháp theo Fig.1 và Fig.3. Mô tả sau đây của thiết bị điện toán 500 được cung cấp để làm ví dụ và không nhằm giới hạn.

Như được biểu diễn trên Fig.5, thiết bị điện toán mẫu 500 bao gồm bộ xử lý 504 để thực thi các đoạn chương trình phần mềm. Mặc dù bộ xử lý đơn được thể hiện cho mục đích rõ ràng, nhưng thiết bị điện toán 500 có thể cũng bao gồm hệ thống đa bộ xử lý. Bộ xử lý 504 được kết nối với cơ sở hạ tầng truyền thông 506 để trao đổi thông tin với các bộ phận khác của thiết bị điện toán 500. Ví dụ, cơ sở hạ tầng truyền thông 506 có thể bao gồm bus truyền thông, thanh chéo, hoặc mạng.

Thiết bị điện toán 500 còn bao gồm bộ nhớ chính 508, như là bộ nhớ truy cập ngẫu nhiên (random access memory, RAM), và bộ nhớ phụ 510. Ví dụ, bộ nhớ phụ 510 có thể bao gồm ổ đĩa cứng 512 và/hoặc ổ lưu trữ di động 514, có thể bao gồm ổ đĩa mềm, ổ băng từ, ổ đĩa quang, hoặc loại tương tự. Ổ lưu trữ di động 514 đọc từ và/hoặc ghi vào bộ lưu trữ di động 518 theo cách đã biết. Bộ lưu trữ di động 518 có thể bao gồm đĩa mềm, băng từ, đĩa quang, hoặc loại tương tự, được đọc và ghi vào bởi ổ lưu trữ di động 514. Như sẽ được hiểu bởi người có hiểu biết trung bình về lĩnh vực kỹ thuật liên quan, bộ lưu trữ di động 518 bao gồm phương tiện lưu trữ có thể đọc được bằng máy tính được lưu trữ trong đó các lệnh và/hoặc dữ liệu mã chương trình thực thi của máy tính.

Theo phương án thay thế, bộ nhớ phụ 510 có thể bổ sung hoặc thay thế bao gồm các phương tiện tương tự khác để cho phép các chương trình máy tính hoặc các lệnh khác được tải vào thiết bị điện toán 500. Ví dụ, các phương tiện như vậy có thể bao gồm bộ lưu trữ di động 522 và giao diện 520. Ví dụ của bộ lưu trữ di động 522 và giao diện 520 bao gồm hộp chương trình và giao diện hộp (như được tìm thấy trong các thiết bị bàn giao tiếp trò chơi video), chip nhớ di động (như là EPROM hoặc PROM) và ổ cắm kết hợp, và các bộ lưu trữ di động 522 khác và các giao diện 520 cho phép phần mềm và dữ liệu được truyền từ bộ lưu trữ di động 522 đến hệ thống máy tính 500.

Thiết bị điện toán 500 cũng bao gồm ít nhất một giao diện truyền thông 524. Giao diện truyền thông 524 cho phép phần mềm và dữ liệu được truyền giữa thiết bị điện toán 500 và các thiết bị ngoại vi theo đường dẫn truyền thông 526. Theo các phương án khác nhau của sáng chế, giao diện truyền thông 524 cho phép truyền dữ liệu giữa thiết bị điện toán 500 và mạng truyền thông dữ liệu, như là dữ liệu công khai hoặc mạng truyền thông dữ liệu riêng. Giao diện truyền thông 524 có thể được dùng để trao đổi dữ liệu giữa các thiết bị điện toán 500 khác nhau mà những thiết bị điện toán 500 như vậy tạo thành mạng máy tính kết nối với nhau. Ví dụ của giao diện truyền thông 524 có thể bao gồm mô-đem, giao diện mạng (như thẻ Ethernet), cổng truyền thông, ăngten với hệ mạch kết hợp và dạng tương tự. Giao diện truyền thông 524 có thể kết nối có dây hoặc kết nối không dây. Phần mềm và dữ liệu được truyền theo giao diện truyền thông 524 dưới dạng các tín hiệu có thể là điện tử, điện từ, quang hoặc các tín hiệu khác có khả năng nhận được bởi giao diện truyền thông 524. Các tín hiệu này được cung cấp cho giao diện truyền thông theo đường dẫn truyền thông 526.

Như được biểu diễn trên Fig.5, thiết bị điện toán 500 bao gồm thêm giao diện màn hình 502 thực hiện các hoạt động hiển thị các hình ảnh lên màn hình kết hợp 530 và giao diện âm thanh 532 để thực hiện các hoạt động để phát nội dung âm thanh theo (các) loa kết hợp 534.

Như được sử dụng ở đây, thuật ngữ "sản phẩm chương trình máy tính" có thể đề cập một phần đến bộ lưu trữ di động 518, bộ lưu trữ di động 522, đĩa cứng được cài đặt trong ổ đĩa cứng 512, hoặc phần mềm mang sóng mang qua đường dẫn truyền thông 526 (kết hợp không dây hoặc cáp) đến giao diện truyền thông 524. Phương tiện lưu trữ có thể đọc được bằng máy tính đề cập đến phương tiện lưu trữ xác thực không chuyển tiếp bất kỳ cung cấp các lệnh và/hoặc dữ liệu được lưu lại cho thiết bị điện toán 500 để thực thi và/hoặc xử lý. Các ví dụ của các phương tiện lưu trữ như vậy bao gồm các đĩa mềm, băng từ, CD-ROM, DVD, đĩa Blu-ray™, ổ đĩa cứng, ROM hoặc mạch điện tích hợp, bộ nhớ USB, đĩa từ-quang, hoặc thẻ đọc được bằng máy tính như là thẻ PCMCIA và dạng tương tự, liệu các thiết bị như vậy có là thiết bị nội hay ngoại vi của thiết bị điện toán 500 hay không. Các ví dụ của phương tiện truyền có thể đọc được bằng máy tính xác thực hoặc không xác

thực cũng có thể tham gia vào việc cung cấp phần mềm, các lệnh và/hoặc dữ liệu đến thiết bị điện toán 500 bao gồm các kênh truyền phát vô tuyến hoặc hồng ngoại cũng như kết nối mạng đến máy tính hoặc thiết bị được kết nối mạng khác, và mạng Internet hoặc mạng nội bộ bao gồm truyền thư điện tử và thông tin được ghi lên các trang web và dạng tương tự.

Các chương trình máy tính (còn gọi là mã chương trình máy tính) được lưu trữ trong bộ nhớ chính 508 và/hoặc bộ nhớ phụ 510. Các chương trình máy tính cũng có thể nhận được theo giao diện truyền thông 524. Các chương trình máy tính như vậy, khi được thực hiện, cho phép thiết bị điện toán 500 thực hiện một hoặc nhiều tính năng của các phương án được thảo luận ở đây. Theo các phương án khác nhau, các chương trình máy tính, khi được thực hiện, cho phép bộ xử lý 504 thực hiện các tính năng của các phương án được mô tả ở trên. Theo đó, các chương trình máy tính như vậy biểu diễn các bộ điều khiển của hệ thống máy tính 500.

Phần mềm có thể được lưu trữ trong sản phẩm chương trình máy tính và được tải vào thiết bị điện toán 500 sử dụng ổ lưu trữ di động 514, ổ đĩa cứng 512, hoặc giao diện 520. Theo cách khác, sản phẩm chương trình máy tính có thể được tải về hệ thống máy tính 500 qua đường dẫn truyền thông 526. Phần mềm, khi được thực thi bởi bộ xử lý 504, khiến thiết bị điện toán 500 thực hiện các chức năng của các phương án được mô tả trong bản mô tả này.

Cần phải hiểu rằng phương án theo Fig.5 chỉ được trình bày làm ví dụ. Do đó, theo một số phương án, một hoặc nhiều tính năng của thiết bị điện toán 500 có thể bị bỏ qua. Ngoài ra, theo một số phương án, một hoặc nhiều tính năng của thiết bị điện toán 500 có thể được được nối với nhau. Ngoài ra, theo một số phương án, một hoặc nhiều tính năng của thiết bị điện toán 500 có thể được chia ra thành một hoặc nhiều bộ phận cấu thành.

Cần hiểu rằng các thành phần đã được minh họa trên Fig.5 thực hiện chức năng cung cấp các phương tiện để thực hiện các chức năng khác nhau và các hoạt động của các máy chủ như được mô tả theo các phương án ở trên.

Theo phương án thực hiện, máy chủ có thể được mô tả chung như thiết bị vật lý bao gồm ít nhất một bộ xử lý và ít nhất một bộ nhớ bao gồm mã chương trình

máy tính. Với ít nhất một bộ xử lý, có ít nhất một bộ nhớ và mã chương trình máy tính được tạo cấu hình để khiến thiết bị vật lý thực hiện các hoạt động cần thiết.

Cần được hiểu bởi người có hiểu biết trung bình về lĩnh vực kỹ thuật liên quan rằng nhiều cải biến và/hoặc biến đổi có thể được thực hiện cho sáng chế như được thể hiện trong các phương án cụ thể mà không vượt ra khỏi bản chất và phạm vi của sáng chế theo nghĩa rộng. Ví dụ, các phương pháp, hệ thống và thiết bị như được mô tả có thể được điều chỉnh phù hợp cho các dạng phát hiện đối tượng khác nhau, ví dụ phát hiện khuôn mặt hoặc phát hiện xe cộ. Các phương án hiện tại, do đó, được xem xét trong tất cả các khía cạnh được minh họa và không bị giới hạn.

## YÊU CẦU BẢO HỘ

1. Phương pháp đánh giá hiệu quả hoạt động của mô hình phát hiện đối tượng, phương pháp này bao gồm các bước:

tạo ra khung chữ nhật bao quanh được dự đoán biểu diễn đối tượng dựa trên mô hình phát hiện đối tượng, trong đó đối tượng này được định vị gần một hoặc nhiều đối tượng liền kề;

xác định diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng;

xác định diện tích phần hợp sửa đổi giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng, trong đó bước xác định diện tích phần hợp sửa đổi bao gồm các bước:

xác định diện tích phần hợp bình quân gia quyền giữa các khung chữ nhật bao quanh được dự đoán và đúng của đối tượng dựa trên một hoặc nhiều trọng số; và

thêm vào diện tích phần hợp bình quân gia quyền diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và ít nhất một khung chữ nhật bao quanh đúng của một hoặc nhiều đối tượng liền kề; và

xác định điểm số bằng với diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng chia cho diện tích phần hợp sửa đổi, trong đó điểm số này biểu diễn hiệu quả hoạt động của mô hình phát hiện đối tượng.

2. Phương pháp theo điểm 1, trong đó đối tượng bao gồm chuỗi ký tự, và các đối tượng liền kề bao gồm các chuỗi ký tự liền kề.

3. Phương pháp theo điểm 1 hoặc 2, trong đó một hoặc nhiều trọng số này bao gồm:

trọng số thứ nhất định trước được kết hợp với một phần của khung chữ nhật bao quanh được dự đoán không giao với khung chữ nhật bao quanh đúng của đối tượng,

trọng số thứ hai định trước được kết hợp với một phần của khung chữ nhật bao quanh đúng của đối tượng không giao với khung chữ nhật bao quanh được dự đoán; và

trong đó trọng số thứ nhất ít hơn trọng số thứ hai.

4. Phương pháp theo điểm 3, trong đó bước xác định diện tích phần hợp bình quân gia quyền bao gồm gán trọng số thứ nhất và thứ hai thành các phần kết hợp và bước cộng các diện tích bình quân gia quyền của các phần đã nêu với diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng.

5. Phương pháp theo điểm bất kỳ trong các điểm trên, trong đó điểm số gần sát với 1 biểu diễn độ chính xác cao hơn của mô hình phát hiện đối tượng.

6. Phương pháp phát hiện đối tượng bao gồm các bước:

thu dữ liệu thực chưa dán nhãn và các mô hình phát hiện đối tượng, trong đó các mô hình phát hiện đối tượng này được tạo ra bởi mạng nơron dựa trên dữ liệu huấn luyện được dán nhãn;

đánh giá từng mô hình phát hiện đối tượng với dữ liệu hợp lệ sử dụng phương pháp theo điểm bất kỳ trong các điểm nêu trên để nhận biết mô hình phát hiện đối tượng có điểm số cao nhất; và

phát hiện đối tượng trong dữ liệu thực chưa dán nhãn dựa trên mô hình phát hiện đối tượng đã nhận biết.

7. Phương pháp theo điểm 6, trong đó bước thu các mô hình phát hiện đối tượng bao gồm bước thu các siêu tham số được kết hợp với các mô hình tương ứng, và trong đó mô hình phát hiện đối tượng cùng với một hoặc nhiều siêu tham số được kết hợp có điểm số cao nhất được nhận biết.

8. Phương pháp theo điểm 6 hoặc 7, trong đó bước phát hiện đối tượng trong dữ liệu thực chưa dán nhãn bao gồm phát hiện chuỗi ký tự trong dữ liệu ảnh của ảnh quang học của đối tượng.

9. Hệ thống đánh giá mô hình phát hiện đối tượng, hệ thống này bao gồm:

bộ xử lý; và

bộ nhớ đọc được bằng máy tính được nối với bộ xử lý và có các lệnh được lưu trữ trên đó được thực thi bởi bộ xử lý để:

tạo ra khung chữ nhật bao quanh được dự đoán biểu diễn đối tượng dựa trên mô hình phát hiện đối tượng, trong đó đối tượng này được định vị gần một hoặc nhiều đối tượng liên kề;

xác định diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng;

xác định diện tích phần hợp sửa đổi giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng, trong đó diện tích phần hợp sửa đổi bao gồm tổng của:

diện tích phần hợp bình quân gia quyền giữa các khung chữ nhật bao quanh được dự đoán và đúng của đối tượng dựa trên một hoặc nhiều trọng số; và

diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và ít nhất một khung chữ nhật bao quanh đúng của một hoặc nhiều đối tượng liên kề; và

xác định điểm số bằng với diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng chia cho diện tích phần hợp sửa đổi, trong đó điểm số này biểu thị hiệu quả hoạt động của mô hình phát hiện đối tượng.

10. Hệ thống theo điểm 9, trong đó đối tượng bao gồm chuỗi ký tự, và trong đó các đối tượng liên kề bao gồm các chuỗi ký tự liên kề.

11. Hệ thống theo điểm 9 hoặc 10, trong đó một hoặc nhiều trọng số này bao gồm:

trọng số thứ nhất định trước được kết hợp với một phần của khung chữ nhật bao quanh được dự đoán không giao với khung chữ nhật bao quanh đúng của đối tượng,

trọng số thứ hai định trước được kết hợp với một phần của khung chữ nhật bao quanh đúng của đối tượng không giao với khung chữ nhật bao quanh được dự đoán; và

trong đó trọng số thứ nhất ít hơn trọng số thứ hai.

12. Hệ thống theo điểm 11, trong đó các lệnh được thực thi bởi bộ xử lý để gán trọng số thứ nhất và thứ hai cho các phần kết hợp và tổng diện tích bình quân gia quyền của các phần đã nêu với diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng để xác định diện tích phần hợp bình quân gia quyền.

13. Hệ thống theo điểm bất kỳ trong số các điểm từ 9 đến 12, trong đó điểm số gần sát với 1 biểu diễn độ chính xác cao hơn của mô hình phát hiện đối tượng.

14. Thiết bị phát hiện đối tượng bao gồm:

môđun phát hiện đối tượng được tạo cấu hình để tạo ra khung chữ nhật bao quanh được dự đoán biểu diễn đối tượng dựa trên mô hình phát hiện đối tượng, trong đó đối tượng được định vị gần một hoặc nhiều đối tượng liền kề; và

môđun đánh giá được tạo cấu hình để:

xác định diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng;

xác định diện tích phần hợp sửa đổi giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng, trong đó diện tích phần hợp sửa đổi bao gồm tổng của:

diện tích phần hợp bình quân gia quyền giữa các khung chữ nhật bao quanh được dự đoán và đúng của đối tượng dựa trên một hoặc nhiều trọng số; và

diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và ít nhất một khung chữ nhật bao quanh đúng của một hoặc nhiều đối tượng liền kề;

xác định điểm số bằng với diện tích phần giao giữa khung chữ nhật bao quanh được dự đoán và khung chữ nhật bao quanh đúng của đối tượng chia cho diện tích phần hợp sửa đổi; và

đánh giá hiệu quả hoạt động của mô hình phát hiện đối tượng dựa trên điểm số.

15. Bộ phát hiện đối tượng bao gồm:

môđun thu được tạo cấu hình để thu dữ liệu thực chưa dán nhãn và các mô hình phát hiện đối tượng, trong đó mô hình phát hiện đối tượng này được tạo ra bởi mạng nơron dựa trên dữ liệu huấn luyện được dán nhãn; và

các thiết bị theo điểm 14 được nối với bộ môđun thu và được tạo cấu hình để đánh giá từng mô hình phát hiện đối tượng với dữ liệu hợp lệ để nhận biết mô hình phát hiện đối tượng có điểm số cao nhất;

trong đó môđun phát hiện đối tượng còn được tạo cấu hình để phát hiện đối tượng trong dữ liệu thực chưa dán nhãn dựa trên mô hình phát hiện đối tượng đã nhận biết.

16. Bộ phát hiện đối tượng theo điểm 15, trong đó môđun thu còn được tạo cấu hình để thu các siêu tham số được kết hợp với các mô hình phát hiện đối tượng tương ứng, và trong đó mô hình phát hiện đối tượng cùng với một hoặc nhiều siêu tham số được kết hợp có điểm số cao nhất được nhận biết.

17. Bộ phát hiện đối tượng theo điểm 15 hoặc 16, trong đó đối tượng này bao gồm chuỗi ký tự, và dữ liệu thực chưa dán nhãn bao gồm dữ liệu ảnh của ảnh quang học của đối tượng.

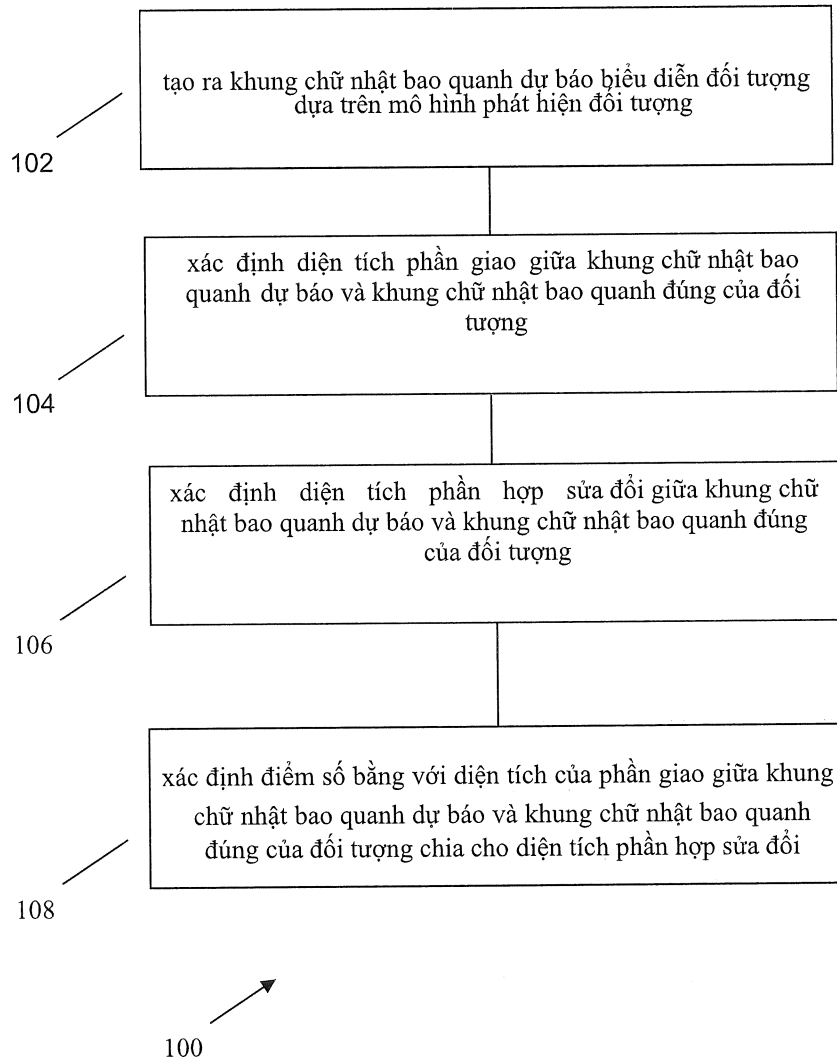


Fig.1

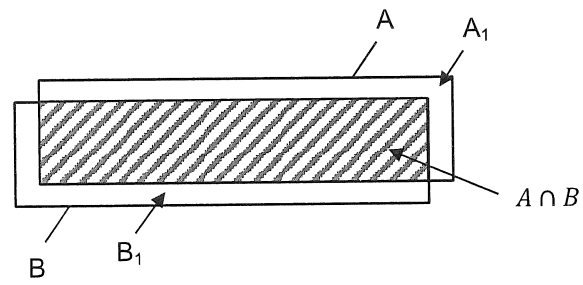


Fig.2a

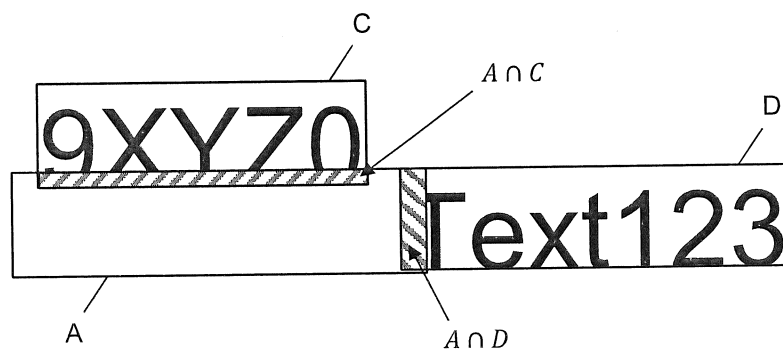


Fig.2b

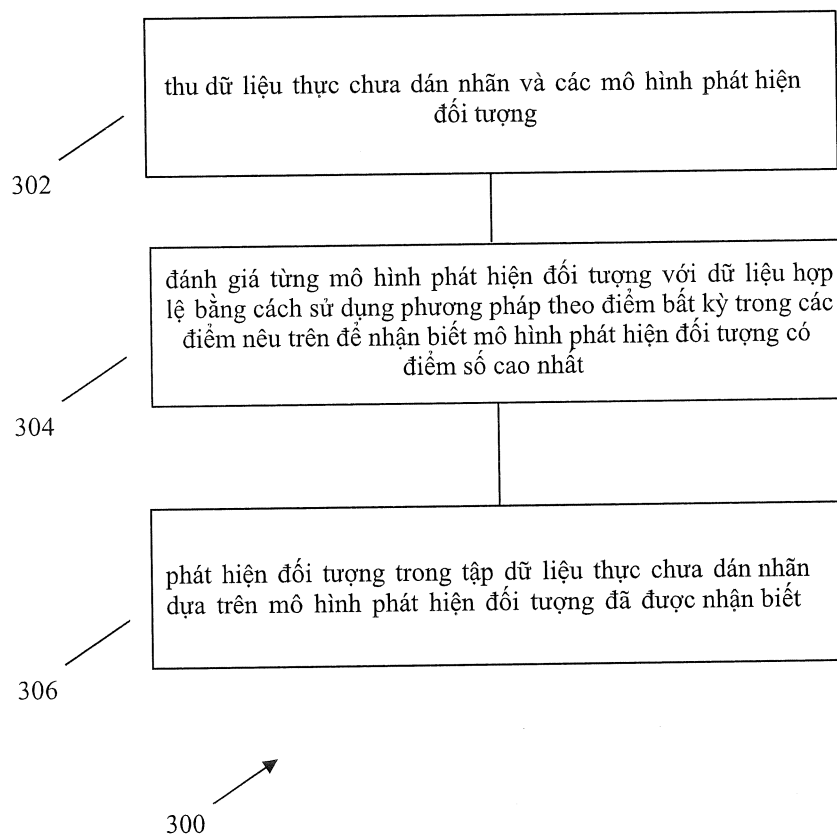


Fig.3

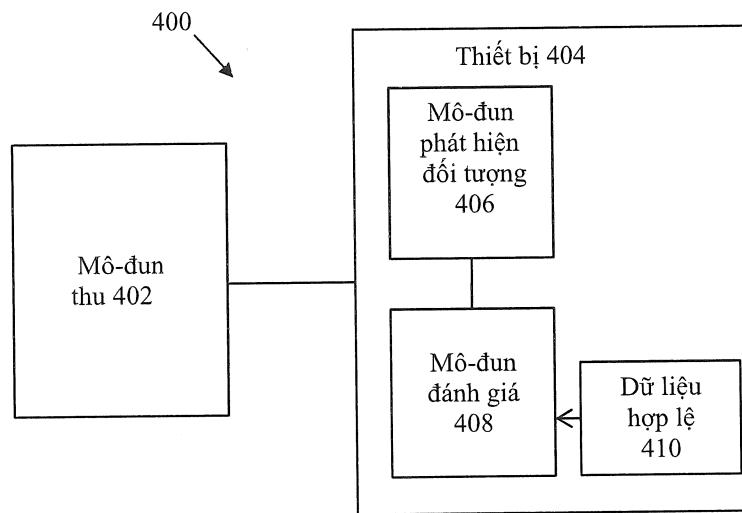


Fig.4

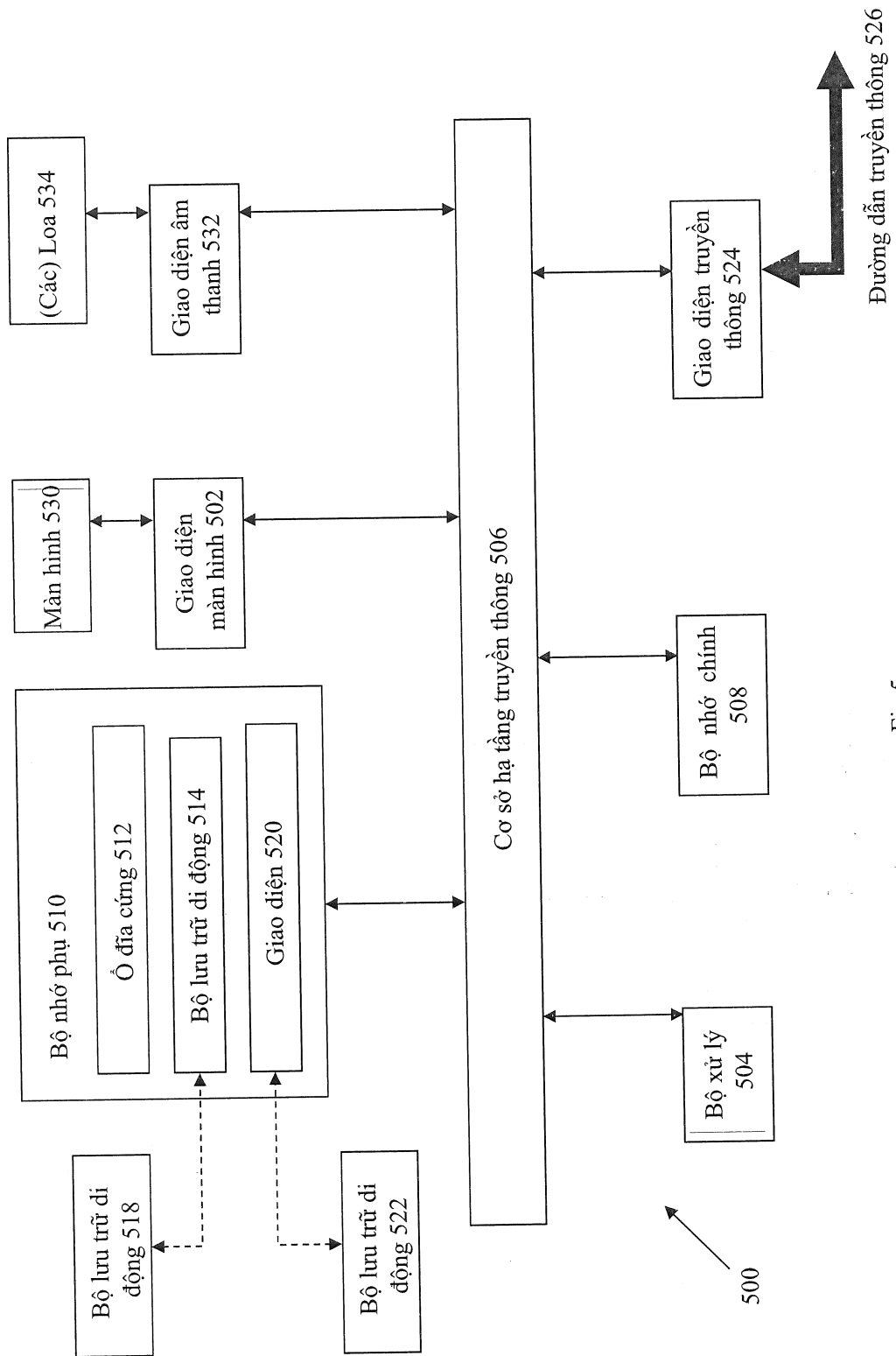


Fig.5