(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 March 2020 (12.03.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/051114 A1

(51) International Patent Classification:
G06K 9/46 (2006.01) *G06K 9/00* (2006.01)

(21) International Application Number:
PCT/US2019/049291

(22) International Filing Date:
03 September 2019 (03.09.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
201811025620.5 04 September 2018 (04.09.2018) CN

(71) Applicant: **ALIBABA GROUP HOLDING LIMITED** [—/US]; Fourth Floor, One Capital Place, P.O. Box 847, George Town, Grand Cayman (KY).

(72) Inventor: **ZHENG, Dandan**; Alibaba Group Legal Department 5/f, Building 3, No. 969 West Wen Yi Road, Yu Hang District, Hangzhou, Zhejiang, 311121 (CN).

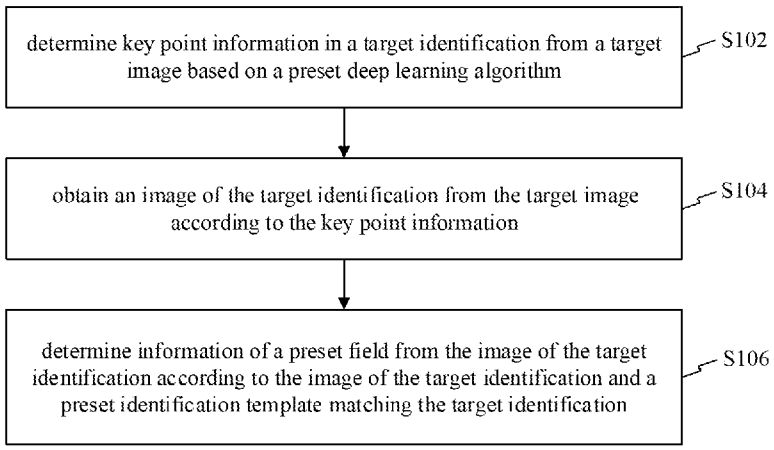
(74) Agent: **XU, Yanbin**; Finnegan, Henderson, Farabow, Garrett & Dunner, LLP, 901 New York Avenue, NW, Washington, DC 20001-4413 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*

(54) Title: INFORMATION DETECTION METHOD, APPARATUS AND DEVICE



(57) Abstract: The embodiments of the present specification disclose an information detection method, an information detection apparatus, and an information detection device. The method comprises: determining key point information in a target identification from a target image based on a preset deep learning algorithm; obtaining an image of the target identification from the target image according to the key point information; and determining information of a preset field from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification.

WO 2020/051114 A1

INFORMATION DETECTION METHOD, APPARATUS AND DEVICE

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This international application is based upon and claims priority to Chinese Application No. 201811025620.5, filed on September 4, 2018, the entire content of which is incorporated herein by reference.

TECHNICAL FIELD

[0002] The present specification relates to the technical field of computers, and in particular, to a method, an apparatus, and a device for information detection.

TECHNICAL BACKGROUND

[0003] With continuous development of an electronic process of real-name authentication, as a weak real-name authentication solution, the electronic process of real-name authentication does not require users to go to a real-name authentication site to verify identifications and personal information, thus greatly saving costs of users and real-name authentication authorities (such as banks). Therefore, it is increasingly accepted by real-name authentication authorities.

[0004] Despite the above advantages, the electronic process of real-name authentication needs to implement text line detection for identifications of many different countries or regions in order to extract names, addresses, identification numbers and other information on the identifications. Moreover, characteristics of the identifications of different countries or regions are also different. Generally, the detection of text lines in an identification is realized by traditional graphic methods, such as connected-component based schemes or text detection methods based

on stroke width conversion. However, the detection methods have low accuracies and complicated rules, and for new types of identifications, it is required to rewrite corresponding detection codes. Therefore, it is desired to establish a simple and universal way to accurately detect information of key fields on an identification.

5

SUMMARY OF THE INVENTION

[0005] The objective of the embodiments of the present specification is to provide a method, an apparatus, and a device for information detection to establish a simple and universal way to accurately detect information of key fields on an identification.

10 [0006] To achieve the above technical solutions, the embodiments of the present specification are implemented as follows:

[0007] An information detection method provided by the present specification, includes:

[0008] determining key point information in a target identification from a target image based on a preset deep learning algorithm;

15 [0009] obtaining an image of the target identification from the target image according to the key point information; and

[0010] determining information of a preset field from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification.

20 [0011] Optionally, the determining the key point information in the target identification from the target image based on the preset deep learning algorithm includes:

[0012] determining, based on data of a preset sample image, a relationship among key points corresponding to four right angles of an identification in the sample image and edges of the

identification in the sample image and an avatar in the identification in the sample image to construct a regression network model; and

[0013] determining the key point information in the target identification from the target image based on the constructed regression network model.

[0014] Optionally, the regression network model includes a first convolutional layer, a first pooling layer, a first dropout layer, a second convolutional layer, a second pooling layer, a second dropout layer, a third convolutional layer, a third dropout layer, a first fully connected layer, and a second fully connected layer, wherein the first convolutional layer is connected to an input layer, the second fully connected layer is connected to an output layer, and layers are connected in an order of: the first convolutional layer, the first pooling layer, the first dropout layer, the second convolutional layer, the second pooling layer, the second dropout layer, the third convolutional layer, the third dropout layer, the first fully connected layer, and the second fully connected layer.

[0015] Optionally, the first dropout layer, the second dropout layer, and the third dropout layer are used to increase preset noise information.

[0016] Optionally, the first fully connected layer includes a plurality of cells for mapping a learned distributed feature representation into a space of the sample image.

[0017] Optionally, the second fully connected layer includes eight cells, and the eight cells respectively correspond to the key points corresponding to the four right angles of the target identification and eight parameters.

[0018] Optionally, the determining the key point information in the target identification from the target image based on the constructed regression network model includes:

[0019] determining, based on the constructed regression network model, a region formed by the four right angles of the target identification from the target image; and

[0020] determining, based on the constructed regression network model, the key point information in the target identification from the determined region formed by the four right angles of the target identification.

[0021] Optionally, before determining the information of the preset field from the image of the target identification according to the image of the target identification and the preset identification template matching the target identification, the method further includes:

[0022] adjusting a size of the image of the target identification based on preset reference size information of the identification to obtain an image of the target identification matching the reference size information.

[0023] An information detection apparatus provided by some embodiments of the present specification, includes:

[0024] a key point determining module, configured to determine key point information in a target identification from a target image based on a preset deep learning algorithm;

[0025] an identification obtaining module, configured to obtain an image of the target identification from the target image according to the key point information; and

[0026] a field information determining module, configured to determine information of a preset field from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification.

[0027] Optionally, the key point determining module includes:

[0028] a model constructing unit, configured to determine, based on data of a preset sample image, a relationship among key points corresponding to four right angles of an identification in the sample image and edges of the identification in the sample image and an avatar in the identification in the sample image to construct a regression network model; and

[0029] a key point determining unit, configured to determine key point information in the target identification from the target image based on the constructed regression network model.

[0030] Optionally, the regression network model includes a first convolutional layer, a first pooling layer, a first dropout layer, a second convolutional layer, a second pooling layer, a second dropout layer, a third convolutional layer, a third dropout layer, a first fully connected layer, and a second fully connected layer, wherein the first convolutional layer is connected to an input layer, the second fully connected layer is connected to an output layer, and layers are connected in an order of: the first convolutional layer, the first pooling layer, the first dropout layer, the second convolutional layer, the second pooling layer, the second dropout layer, the third convolutional layer, the third dropout layer, the first fully connected layer, and the second fully connected layer.

[0031] Optionally, the first dropout layer, the second dropout layer, and the third dropout layer are used to increase preset noise information.

[0032] Optionally, the first fully connected layer includes a plurality of cells for mapping a learned distributed feature representation into a space of the sample image.

[0033] Optionally, the second fully connected layer includes eight cells, and the eight cells respectively correspond to the key points corresponding to the four right angles of the target identification and eight parameters.

[0034] Optionally, the key point determining unit is configured to determine, based on the constructed regression network model, a region formed by the four right angles of the target identification from the target image, and determine, based on the constructed regression network model, key point information in the target identification from the determined region formed by the four right angles of the target identification.

[0035] Optionally, the apparatus further includes:

[0036] an image adjusting module, configured to adjust a size of the image of the target identification based on preset reference size information of the identification to obtain an image of the target identification matching the reference size information.

[0037] An information detection device provided by some embodiments of the present specification, including:

[0038] a processor, and

[0039] a memory arranged to store computer executable instructions that, when executed, cause the processor to:

[0040] determine key point information in a target identification from a target image based on a preset deep learning algorithm;

[0041] obtain an image of the target identification from the target image according to the key point information; and

[0042] determine information of a preset field from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification.

[0043] It can be seen from the technical solutions provided by the embodiments of the present specification that according to the embodiments of the present specification, key point information in a target identification is determined from a target image based on a preset deep learning algorithm; then, an image of the target identification is obtained from the target image according to the key point information; information of a preset field is determined from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification. Therefore, key points of the target identification in the target image are detected by the deep learning algorithm, so that the exact

location of the target identification in the target image may be obtained, and then by matching the detected target identification with the preset identification template, information of any field or a specified field may be accurately detected. For example, by comparison with the preset identification template, the information of a name field and/or the information of an address field in the target identification may be accurately detected, so that the processing of detecting the information of the key field on the identification is simplified. Moreover, the solutions are not limited by the type of the identification and may be applied to more types of identifications.

BRIEF DESCRIPTION OF THE DRAWINGS

[0044] In order to more clearly illustrate the technical solutions in the embodiments of the present specification or the prior art, drawings to be used for description of embodiments and the prior art will be introduced briefly hereinafter. Apparently, drawings referred to in the following description are merely some embodiments set forth in the present specification, and those skilled in the art also can obtain other drawings based on these drawings without creative labor.

[0045] FIG. 1 shows an embodiment of an information detection method in the present specification;

[0046] FIG. 2 is a schematic diagram of a target image with a target identification in the present specification;

[0047] FIG. 3 shows an embodiment of another information detection method in the present specification;

[0048] FIG. 4 is a schematic diagram of a region formed by four right angles of a target identification in the present specification;

[0049] FIG. 5 is a schematic diagram of obtaining key points in a target identification based on the region formed by the four right angles of the target identification in the present specification;

[0050] FIG. 6 is a schematic diagram of a matching result after matching a preset identification template in the present specification;

[0051] FIG. 7 shows an embodiment of an information detection apparatus in the present specification; and

[0052] FIG. 8 shows an embodiment of an information detection device in the present specification.

DETAILED DESCRIPTION

[0053] Embodiments of the present specification provide an information detection method, apparatus and device.

[0054] In order to make those skilled in the art better understand the technical solutions in the present specification, the technical solutions in embodiments of the present specification will be clearly and completely described below in conjunction with the accompanying drawings in the embodiments of the present specification. Apparently, the described embodiments are only a part, not all of the embodiments of the present specification. All other embodiments obtained by a person of ordinary skill in the art based on the embodiments of the present specification without creative efforts shall fall within the scope of the present specification.

[0055] Embodiment 1

[0056] As shown in FIG. 1, the embodiment of the present specification provides an information detection method. The execution entity of the method may be a terminal device or a server. The terminal device may be a device such as a personal computer, or may be a mobile

terminal device, such as a mobile phone and a tablet computer; and the terminal device may be a terminal device used by a user. The server may be a stand-alone server, or a server cluster formed by multiple servers. Moreover, the server may be a background server of a certain service, or a background server of a website (e.g., a network shopping website or a payment application). The method may be used in the process of accurately detecting the information of a key field on an identification in a simple and universal manner. In order to improve the detection efficiency of the information, an example where the execution entity is a server is described in this embodiment to explain the method. The case where the execution entity is a terminal device may be handled according to the following related content and will not be elaborated here. The method may specifically include the following steps.

[0057] In step S102, key point information in a target identification is determined from a target image based on a preset deep learning algorithm.

[0058] The deep learning algorithm may be an algorithm based on representation learning of data in machine learning. There may be many types of deep learning algorithms, such as kNN (k-Nearest Neighbor) algorithm or CNN (Convolutional Neural Network) algorithm. The target image may be any image in which an image of a certain identification is included. For example, a user takes a photo of an ID card placed on a desk with a camera of a mobile phone, and during the shooting process, a part of an image of the desk (e.g., an image of the desktop around the ID card) is also captured in the image, and the captured image may be the target image. The target identification may be a credential and/or document used to prove an identity, experience, etc., and the target identification may include many types of forms. For example, the target identification may be a card of a preset size, or the target identification may be one or more credentials and/or documents of a preset size. The key point information may be information of one or more

prominent or notable positions, and the key point information may include position information, coordinate information, etc. of key points.

[0059] In the implementation, with continuous development of the electronic process of real-name authentication, as a weak real-name authentication solution, the electronic process of real-name authentication does not require users to verify identifications and personal information on site for real-name authentication, thus greatly saving costs of users and real-name authentication authorities (such as banks). Therefore, it is increasingly accepted by real-name authentication authorities. However, despite the above advantages, the electronic process of real-name authentication needs to implement text line detection for identifications of different countries or regions in order to extract names, addresses, identification numbers and other information on the identifications. However, characteristics of the identifications in different countries or regions are different. For example, some countries or regions have multiple versions of identifications, and the names and addresses on the identifications may occupy 1-3 lines according to the length of their contents. For identifications of some countries or regions, there are both English and a local language, and it is thus desired to establish a simple and universal way to accurately detect the information of key fields on the identifications.

[0060] There are two ways to detect text lines in an identification, one of which may be implemented by traditional graphic methods, such as connected-component schemes, text detection methods based on stroke width conversion, or schemes based on pixel variation to find edge characteristics, such as MSER (Maximally Stable Extremal Regions). The purpose of the MSER is to find a region that can have significant changes with respect to the surrounding region. When detecting text lines in an identification, the traditional graphic methods have low accuracies and complicated rules. For newly emerging types of identifications, it is required to rewrite

corresponding detection codes. The other way is text detection based on a deep learning target detection algorithm, such as SSD (Single Shot multibox Detector), Faster RCNN and other algorithms for text detection. The core of the SSD algorithm may be predicting category scores, offset, etc. of a series of default bounding boxes by virtue of a convolutional kernel on a feature map. The process of the Faster RCNN algorithm may be: first inputting an entire picture into a CNN to obtain a Feature Map, and then inputting convolutional features into an RPN (Region Proposal Network) to obtain feature information of candidate boxes; determining, by a classifier, whether features extracted from the candidate boxes belong to a particular class; and for a candidate box belonging to a feature, further adjusting the position of the candidate box by using a regressor. The text detection based on the deep learning target detection algorithm has higher accuracy. However, for texts in the identifications of different countries or regions, a large amount of data is required to train a model. Moreover, it cannot distinguish which field is name and which field is address, etc. Therefore, it may also require assistance of other rules or language models to further determine whether the text line in the identification belongs to a name or an address. Therefore, the embodiment of the present specification provides a solution for detecting a text line in an identification, which may specifically include the following contents.

[0061] Considering that an image (i.e., the target image) may not only have an image of an identification (i.e., the target identification), it may also contain an image other than the identification, and since the image other than the identification does not help the text line detection of the target identification, the image of the target identification may be extracted from the target image. It can be considered that the image of the target identification is determined by a plurality of key points. For example, as shown in FIG. 2, points corresponding to the four corners of the target identification in the target image may be used as key points, and points corresponding to the

four edges of the target identification in the target image may be used as key points, etc. In such a way, a region defined by the key points corresponding to the four corners and the key points corresponding to the four edges is the region where the target identification is located, and for this reason, key point detection of the target identification may be performed on the basis of a multi-Patch key point positioning method. Specifically, a deep learning algorithm may be preset, and the deep learning algorithm may be used to detect key point information of the identification from the image, and then, a certain number of sample images may be selected. Each sample image includes an image of a certain identification, and the preset deep learning algorithm may be trained by the sample images to obtain a deep learning algorithm capable of detecting key point information of the identification contained therein from the sample images.

[0062] After the deep learning algorithm is trained in the above manner, the target image may be input into the trained deep learning algorithm. Through the calculation with the trained deep learning algorithm, the key points corresponding to the four corners of the target identification and the key points corresponding to the four edges may be extracted from the target image. In practical applications, the key points of the target identification are not limited to the above-mentioned multiple key points, and may also include other key points in addition to the above-mentioned key points, which is not limited in the embodiment of the present specification.

[0063] In step S104, an image of the target identification is obtained from the target image according to the key point information.

[0064] In the implementation, after the key point information in the target identification in the target image is obtained by the processing of the step S102, the obtained key point information may be respectively set at corresponding positions in the target image, and then the key point information or a region defined by the key points may be determined, for example, the region

defined by the key points corresponding to the four corners of the target identification and the key points corresponding to the four edges, as shown in FIG. 2. The above determined region may be determined as a region where the target identification is located. Then the image of the above-described region may be extracted or intercepted from the target image as the image of the target identification.

[0065] In step S106, information of a preset field is determined from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification.

[0066] The preset identification template may be an identification having the same size, the same field name of each field, and the same relative position of each field on an identification, but the present identification template does not contain the contents of the fields. The preset field may be any field determined in advance, such as a name field, an address field, a gender field, or the like.

[0067] In the implementation, in order to accurately identify different fields and the corresponding information of the fields, one or more identification templates may be pre-configured, which may match identifications commonly used at present. When the image of the target identification is obtained by the processing of the step S104, in order to be able to intuitively match a preset identification template, the image of the target identification may be scaled to match the size of the preset identification templates. Then a preset identification template having fields that match fields included in the target identification is selected from the preset identification templates. The preset field may be obtained from the selected preset identification template, and information of the preset field may be obtained from a position in the image of the target identification which corresponds to the preset field. For example, a name field may be obtained

from the selected preset identification template, in which the name field is positioned at the top and the leftmost side of the preset identification template. In this example, information positioned at the top and the leftmost side of the identification may be searched in the image of the target identification, from which name information in the target identification may be obtained.

[0068] It should be noted that the process of comparing the image of the target identification and the preset identification template matching the target identification, and the process of obtaining the information of the preset field may be implemented using a preset algorithm. The preset algorithm may be set according to actual conditions, which may correspond to, for example, methods based on connected components or methods based on stroke width conversion, which is not limited in the embodiment of the present specification. The process of a text detection method based on stroke width conversion includes: calculating a Canny edge of an image; then, calculating stroke width conversion of the image according to direction information of the edge; integrating pixels into connected components according to stroke width information; filtering the connected components by geometric inference such as a depth-width ratio of the connected components, a variance, mean, or median of the connected component strokes, etc.; integrating the connected components into a text line; and finally partitioning the text line into individual words.

[0069] According to the information detection method provided by the embodiment of the present specification, key point information in a target identification is determined from a target image based on a preset deep learning algorithm. Then, an image of the target identification is obtained from the target image according to the key point information. Information of a preset field is determined from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification. Therefore, key

points of the target identification in the target image are detected by the deep learning algorithm, so that the exact location of the target identification in the target image may be obtained, and then by matching the detected target identification with the preset identification template, information of any field or a specified field may be accurately detected. For example, by comparison with the preset identification template, the information of a name field and/or the information of an address field in the target identification may be accurately detected, so that the processing of detecting the information of the key field on the identification is simplified. Moreover, the solutions are not limited by the type of the identification and may be applied to more types of identifications.

[0070] Embodiment 2

[0071] As shown in FIG. 3, the embodiment of the present specification provides an information detection method. The execution entity of the method may be a terminal device or a server. The terminal device may be a device such as a personal computer, or may be a mobile terminal device, such as a mobile phone and a tablet computer; and the terminal device may be a terminal device used by a user. The server may be a stand-alone server or a server cluster formed by multiple servers. Moreover, the server may be a background server of a certain service, or may be a background server of a website (such as a network shopping website or a payment application). The method may be used in the process of accurately detecting the information of a key field on an identification in a simple and universal manner. In order to improve the detection efficiency of the information, an example where the execution entity is a server is described in this embodiment to explain the method. The case where the execution entity is a terminal device may be handled according to the following related content and will not be elaborated here. The method may specifically include the following steps:

[0072] In step S302, based on data of a preset sample image, a relationship among key points corresponding to four right angles of an identification in the sample image and edges of the identification in the sample image, and an avatar in the identification in the sample image is determined to construct a regression network model.

[0073] The data of the sample image may be pixel data of the sample image, for example, 32×32 or 100×100. The regression network model may be a network model constructed based on a preset deep learning algorithm, and the regression network model may include various types, such as a convolutional neural network model or a deep confidence network model.

[0074] In the implementation, a certain number of sample images may be selected to train the preset deep learning algorithm. The deep learning algorithm is essentially an input-to-output mapping, which can learn a large number of mapping relationships between input and output without any precise mathematical expression between input and output; as long as a corresponding regression network is trained with known patterns, the regression network has the ability to map between input and output pairs. Sample data for the regression network may be formed by vector pairs in the form of (input vector, ideal output vector). All weights may be initialized with a few different small random numbers before starting the training. The small random numbers may be used to ensure that the network does not become saturated due to excessive weights, which leads to a training failure. The difference between the small random numbers can be used to ensure that the network can learn normally. The training of the deep learning algorithm may include two stages. Stage 1 is a forward propagation stage, i.e., selecting a sample data (X, Yp) from a set of sample data, and inputting X into a regression network corresponding to the deep learning algorithm, and then, calculating the corresponding actual output Op. Stage 2 is a backward propagation stage, i.e., calculating the difference between the actual output Op and a corresponding

ideal output Yp, and then carrying out backward propagation according to an error minimizing method to adjust the weight matrix. The trained regression network model may be obtained after performing the two stages.

[0075] Based on the principle of the regression network model corresponding to the deep learning algorithm described above, the preset deep learning algorithm may be used to determine, according to the data of the preset sample image, the relationship among the key points corresponding to the four right angles of the identification in the sample image and edges of the identification in the sample image and an avatar in the identification in the sample image, and a regression network model is then constructed on the basis of the relationship. The constructed regression network model may include five stages. The first stage may include a first convolutional layer, a first pooling layer, and a first dropout layer. The first convolutional layer may perform a convolutional operation, through which the original image features may be enhanced and noise may be reduced. The first convolutional layer may be formed by one or more Feature Maps. Each neuron in the Feature Maps may be connected to the neighborhood of N*M in the input, where N and M are both natural numbers. The first pooling layer may subsample images according to the principle of local correlation of images, which may reduce data to be processed while maintaining useful information. Features may be pre-defined. The first convolutional layer substantially functions to automatically learn the features, and the first pooling layer may reduce the dimensions while keeping the main features. The first dropout layer may be not added to the first convolutional layer. However, because the data of the sample image is limited, for the versatility between the identifications, the first dropout layer may increase the noise and improve the generalization ability.

[0076] The second stage can include a second convolutional layer, a second pooling layer, and a second dropout layer. Multiple convolutional layers may be set for the following reason that the features learned by one convolutional layer are often local, and the higher level the convolutional layer is located, the more global the learned features may be, so that the overall features of the edges of the identification can be learnt step by step.

[0077] The third stage may include a third convolutional layer and a third dropout layer. In the third stage, noise may be increased while the global features may continue to be learned, and no pooling layer is used for pooling process.

[0078] The fourth stage may include a first fully connected layer. The first fully connected layer may include a plurality of cells (the number of cells may be determined based on a design of the output layer, for example, 500 cells, and so on). The first fully connected layer may have a function of mapping the learned "distributed feature representation" to a data space of the sample image. The first fully connected layer may be fully connected to an upper layer and may have a certain number of trainable parameters. Like a classical neural network, the first fully connected layer calculates a dot product between an input vector and a weight vector, then adds a bias to the dot product, and finally passes it to the next layer.

[0079] The fifth stage may include a second fully connected layer. The second fully connected layer may include eight cells, and the eight cells respectively correspond to four points and eight parameters of the identification.

[0080] Based on the above description, the regression network model may include a first convolutional layer, a first pooling layer, a first dropout layer, a second convolutional layer, a second pooling layer, a second dropout layer, a third convolutional layer, a third dropout layer, a first fully connected layer, and a second fully connected layer. The first convolutional layer is

connected to an input layer, the second fully connected layer is connected to an output layer, and all the layers are connected in the following order: the first convolutional layer, the first pooling layer, the first dropout layer, the second convolutional layer, the second pooling layer, the second dropout layer, the third convolutional layer, the third dropout layer, the first fully connected layer, and the second fully connected layer.

[0081] The first dropout layer, the second dropout layer, and the third dropout layer may be used to increase preset noise information. The first fully connected layer may include a plurality of cells for mapping a learned distributed feature representation into a space of the sample image. The second fully connected layer may include eight cells, and the eight cells respectively correspond to the key points corresponding to the four right angles of the identification in the image and eight parameters.

[0082] It should be noted that in the process of training and subsequent prediction of the regression network model corresponding to the deep learning algorithm, the image may be normalized to avoid a gradient explosion.

[0083] In step S304, key point information in the target identification is determined from the target image based on the constructed regression network model.

[0084] In the implementation, after constructing the regression network model by the processing of the step S302, the data of the target image (i.e., pixel data of the target image) may be input into the constructed regression network model for calculation, from which the relationship among the key points corresponding to the four right angles of the target identification in the target image and edges of the target identification and an avatar in the target identification may be determined, and the key point information in the target identification may also be determined.

[0085] The specific processing manner of the step S304 may be various. In order to improve the accuracy of the target identification detection, key point detection based on small patches may be performed additionally. That is, the regions of the four right angles in the detected target identification are extracted, and the calculation of the key point detection is performed on the regions of the four right angles in the target identification (i.e., the processing of the five stages as described above). Thus, more accurate key point information in the target identification is obtained, so that the accuracy of the key point detection may be greatly improved. To this end, the step S304 may specifically include the following step 1 and step 2.

[0086] In step 1, a region formed by the four right angles of the target identification is determined from the target image based on the constructed regression network model.

[0087] In the implementation, as shown in FIG. 2 and FIG. 4, the data of the target image (i.e., the pixel data of the target image) may be input into the constructed regression network model for calculation. That is, the pixel data of the target image is first input to the first convolutional layer of the first stage of the regression network model for calculation, and then an output result is input to the first pooling layer for pooling processing, and then the pooled data is input to the first dropout layer for processing, thus obtaining a corresponding processing result. The processing result may be subjected to the processing of the second stage, the third stage, the fourth stage, and the fifth stage, respectively, and finally the key point information corresponding to the four right angles of the target identification may be obtained. Then, based on the key point information corresponding to the four right angles of the target identification, a region (i.e., the region as shown in FIG. 4) formed by the four right angles of the target identification is determined.

[0088] In step 2, based on the constructed regression network model, key point information in the target identification is determined from the determined region formed by the four right angles of the target identification.

[0089] In an implementation, the region formed by four right angles of the target identification may be used as an initial image, and input into the regression model again for calculation. That is, the processing of the first stage, the second stage, the third stage, the fourth stage, and the fifth stage is respectively performed on pixel data of an image of the region formed by the four right angles of the target identification, thus finally obtaining key point information in the target identification in the image of the region formed by the four right angles of the target identification, as shown in FIG. 5. Through the processing process, the key point information in the target identification may be more accurate.

[0090] In step S306, an image of the target identification is obtained from the target image according to the key point information.

[0091] After the image of the target identification is obtained from the target image through the processing process, the image of the target identification may be subjected to affine transformation correction which will be detailed in the step S308.

[0092] In the step S308, based on preset reference size information of the identification, a size of the image of the target identification is adjusted to obtain an image of the target identification matching the reference size information.

[0093] The reference size information may be real size (or standard size) information of the target identification in actual application or may be information of a size matching the real size.

[0094] In step S310, information of a preset field is determined from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification.

[0095] The preset identification template may be an identification template set based on any of a reference size, a standard size, or a real size. The preset field may be any field in the target identification, such as a name field or a gender field.

[0096] In the implementation, according to the identification template of the reference size (or the standard size or the real size), a required field and a range of pixel values of its corresponding region may be configured. For example, the following rules may be set: the name field starts from pixel coordinates (100, 50) and ends at pixel coordinates (180, 90), the gender field starts from pixel coordinates (100, 100) and ends at pixel coordinates (140, 140), and the pixel coordinates of the name field are located above the pixel coordinates of the gender field.

[0097] In addition, template correction may further be performed on the above-mentioned preset identification template, and in the identification template of the reference size (or the standard size, or the real size), a text line in the target identification may be identified by a connected-component based method or the like. Moreover, the target identification and the preset identification template matching the target identification may be compared for a difference, and connected components of the preset field may be extracted and corrected by a connected-component based method or the like. For example, information of some fields may be written across lines, so that corresponding rules or the like may be configured directly through the configuration of the template.

[0098] As shown in FIG. 6, a text line, such as the text line of the name field and the text line of the gender field, may be extracted according to the template matching result to obtain corresponding information of the preset field.

[0099] According to the information detection method provided by the embodiment of the present specification, key point information in a target identification is determined from a target image based on a preset deep learning algorithm. Then, an image of the target identification is obtained from the target image according to the key point information. Information of a preset field is determined from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification. Therefore, key points of the target identification in the target image are detected by the deep learning algorithm, so that the exact location of the target identification in the target image may be obtained, and then by matching the detected target identification with the preset identification template, information of any field or a specified field may be accurately detected. For example, by comparison with the preset identification template, the information of a name field and/or the information of an address field in the target identification may be accurately detected, so that the processing of detecting the information of the key field on the identification is simplified. Moreover, the solutions are not limited by the type of the identification and may be applied to more types of identifications.

[0100] Embodiment 3

[0101] The above is the information detection method provided by the embodiments of the present specification. Based on the same concept, the embodiment of the present specification further provides an information detection apparatus, as shown in FIG. 7.

[0102] The information detection apparatus includes: a key point determining module 701, an identification obtaining module 702 and a field information determining module 703.

[0103] The key point determining module 701 is configured to determine key point information in a target identification from a target image based on a preset deep learning algorithm.

[0104] The identification obtaining module 702 is configured to obtain an image of the target identification from the target image according to the key point information.

[0105] The field information determining module 703 is configured to determine information of a preset field from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification.

[0106] In the embodiment of the present specification, the key point determining module 701 includes:

[0107] a model constructing unit, configured to determine, based on data of a preset sample image, a relationship among key points corresponding to four right angles of an identification in the sample image and edges of the identification in the sample image and an avatar in the identification in the sample image to construct a regression network model; and

[0108] a key point determining unit, configured to determine key point information in the target identification from the target image based on the constructed regression network model.

[0109] In the embodiment of the present specification, the regression network model includes a first convolutional layer, a first pooling layer, a first dropout layer, a second convolutional layer, a second pooling layer, a second dropout layer, a third convolutional layer, a third dropout layer, a first fully connected layer, and a second fully connected layer , wherein the first convolutional layer is connected to an input layer, the second fully connected layer is connected to an output layer, and all the layers are connected in the following order: the first convolutional layer, the first pooling layer, the first dropout layer, the second convolutional layer,

the second pooling layer, the second dropout layer, the third convolutional layer, the third dropout layer, the first fully connected layer, and the second fully connected layer.

[0110] In the embodiment of the present specification, the first dropout layer, the second dropout layer, and the third dropout layer are used to increase preset noise information.

[0111] In the embodiment of the present specification, the first fully connected layer includes a plurality of cells for mapping a learned distributed feature representation into a space of the sample image.

[0112] In the embodiment of the present specification, the second fully connected layer includes eight cells, and the eight cells respectively correspond to the key points corresponding to the four right angles of the target identification and eight parameters.

[0113] In the embodiment of the present specification, the key point determining unit is configured to determine, based on the constructed regression network model, a region formed by the four right angles of the target identification from the target image, and determine, based on the constructed regression network model, key point information in the target identification from the determined region formed by the four right angles of the target identification.

[0114] In the embodiment of the present specification, the apparatus further includes:

[0115] an image adjusting module, configured to adjust the size of the image of the target identification based on preset reference size information of the identification to obtain an image of the target identification matching the reference size information.

[0116] According to an information detection apparatus provided by the embodiment of the present specification, key point information in a target identification is determined from a target image based on a preset deep learning algorithm. Then, an image of the target identification is obtained from the target image according to the key point information. Information of a preset

field is determined from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification. Therefore, key points of the target identification in the target image are detected by the deep learning algorithm, so that the exact location of the target identification in the target image may be obtained. Then by matching the detected target identification with the preset identification template, information of any field or a specified field may be accurately detected. For example, by comparison with the preset identification template, the information of a name field and/or the information of an address field in the target identification may be accurately detected, so that the processing of detecting the information of the key field on the identification is simplified. Moreover, the solutions are not limited by the type of the identification and may be applied to more types of identifications.

[0117] Embodiment 4

[0118] The above description is the information detection apparatus provided by the embodiments of the present specification. Based on the same concept, the embodiment of the present specification further provides an information detection device, as shown in FIG. 8.

[0119] The information detection device may be the server or the terminal device provided by the foregoing embodiments.

[0120] The information detection device may vary considerably depending on different configurations or performances, and may include one or more processors 801 and a memory 802. The memory 802 may store one or more stored applications or data. The memory 802 may be a non-permanent memory or a permanent memory. The application stored in the memory 802 may include one or more modules (not shown), each of which may include a series of computer executable instructions for the information detection device. Still further, the processor 801 may be configured to communicate with the memory 802 to execute, on the information detection

device, a series of computer executable instructions in the memory 802. The information detection device may also include one or more power sources 803, one or more wired or wireless network interfaces 804, one or more input/output interfaces 805, and one or more keyboards 806.

[0121] Specifically, in this embodiment, the information detection device includes a memory and one or more programs, wherein the one or more programs are stored in the memory, the one or more programs may include one or more modules, and each module may include a series of computer executable instructions for the information detection device. Through configuration, one or more processors may execute the one or more programs, including performing the following computer executable instructions:

[0122] determining key point information in a target identification from a target image based on a preset deep learning algorithm;

[0123] obtaining an image of the target identification from the target image according to the key point information; and

[0124] determining information of a preset field from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification.

[0125] In the embodiment of the present specification, the determining key point information in a target identification from a target image based on a preset deep learning algorithm includes:

[0126] determining, based on data of a preset sample image, a relationship among key points corresponding to four right angles of an identification in the sample image and edges of the identification in the sample image and an avatar in the identification in the sample image to construct a regression network model; and

[0127] determining key point information in the target identification from the target image based on the constructed regression network model.

[0128] In the embodiment of the present specification, the regression network model includes a first convolutional layer, a first pooling layer, a first dropout layer, a second convolutional layer, a second pooling layer, a second dropout layer, a third convolutional layer, a third dropout layer, a first fully connected layer, and a second fully connected layer , wherein the first convolutional layer is connected to an input layer, the second fully connected layer is connected to an output layer, and all the layers are connected in the following order: the first convolutional layer, the first pooling layer, the first dropout layer, the second convolutional layer, the second pooling layer, the second dropout layer, the third convolutional layer, the third dropout layer, the first fully connected layer, and the second fully connected layer.

[0129] In the embodiment of the present specification, the first dropout layer, the second dropout layer, and the third dropout layer are used to increase preset noise information.

[0130] In the embodiment of the present specification, the first fully connected layer includes a plurality of cells for mapping a learned distributed feature representation into a space of the sample image.

[0131] In the embodiment of the present specification, the second fully connected layer includes eight cells, and the eight cells respectively correspond to the key points corresponding to the four right angles of the target identification and eight parameters.

[0132] In the embodiment of the present specification, the determining key point information in the target identification from the target image based on the constructed regression network model includes:

[0133] determining, based on the constructed regression network model, a region formed by the four right angles of the target identification from the target image; and

[0134] determining, based on the constructed regression network model, key point information in the target identification from the determined region formed by the four right angles of the target identification.

[0135] In the embodiment of the present specification, before determining the information of the preset field from the image of the target identification according to the image of the target identification and the preset identification template matching the target identification, the following step is further included:

[0136] adjusting the size of the image of the target identification based on preset reference size information of the identification to obtain an image of the target identification matching the reference size information.

[0137] According to an information detection device provided by the embodiment of the present specification, key point information in a target identification is determined from a target image based on a preset deep learning algorithm. Then, an image of the target identification is obtained from the target image according to the key point information. Information of a preset field is determined from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification. Therefore, key points of the target identification in the target image are detected by the deep learning algorithm, so that the exact location of the target identification in the target image may be obtained. Then by matching the detected target identification with the preset identification template, information of any field or a specified field may be accurately detected. For example, by comparison with the preset identification template, the information of a name field and/or the information of an address

field in the target identification may be accurately detected, so that the processing of detecting the information of the key field on the identification is simplified. Moreover, the solutions are not limited by the type of the identification and may be applied to more types of identifications.

[0138] The specific embodiments of the specification are described as above. Other embodiments are within the scope of the claims appended. In some cases, actions or steps recited in the claims may be performed in a different order than that in the embodiments and still achieve the desired results. In addition, the processes depicted in the figures are not necessarily in a particular order or in a sequential order to achieve the desired results. In some embodiments, multitasking processing and parallel processing are also possible or may be advantageous.

[0139] In the 1990s, improvements to a technology could be distinguished between hardware improvements (e.g., improvements to circuit structures such as diodes, transistors, and switches) and software improvements (improvements to processes). However, with the development of technologies, many process improvements at present may be seen as direct improvements in hardware circuit structures. Designers almost always obtain a corresponding hardware circuit structure by programming an improved process into a hardware circuit. Therefore, it cannot be said that the improvement of a process cannot be implemented by a hardware entity module. For example, a Programmable Logic Device (PLD) (e.g., a Field Programmable Gate Array (FPGA)) is such an integrated circuit, the logic function of which is determined by a user through programming the device. Designers program by themselves to "integrate" a digital system on a PLD without asking a chip manufacturer for designing and fabricating a dedicated integrated circuit chip. Moreover, instead of manually making integrated circuit chips, this programming is now mostly implemented using "logic compiler" software which is similar to a software compiler used in programming development. Moreover, to compile previous original codes, a specific

programming language called Hardware Description Language (HDL) must be used. There is not only one, but many, HDLs, such as ABEL Advanced Boolean Expression Language (ABEL), Altera Hardware Description Language (AHDL), Confluence, Cornell University Programming Language (CUPL), HDCal, Java Hardware Description Language (JHDL), Lava, Lola, MyHDL, PALASM, Ruby Hardware Description Language (RHDL), or the like. Very-High-Speed Integrated Circuit Hardware Description Language (VHDL) and Verilog are most commonly used at present. It should also be apparent to those skilled in the art that a hardware circuit for implementing a logic process may be easily obtained by simply logically programming the process with a few hardware description languages described above and programming it into an integrated circuit.

[0140] A controller can be implemented in any suitable manner. For example, the controller can take the form of, for example, a microprocessor or processor, and a computer readable medium storing a computer readable program code (e.g., software or firmware) executable by the (micro)processor, a logic gate, a switch, an Application Specific Integrated Circuit (ASIC), a programmable logic controller, or an embedded microcontroller. Examples of the controller include, but are not limited to, the following microcontrollers: ARC 625D, Atmel AT91SAM, Microchip PIC18F26K20 and Silicone Labs C8051F320; and a memory controller can also be implemented as a part of the memory's control logic. Those skilled in the art will also appreciate that in addition to implementing the controller by a purely computer readable program code, the controller may achieve the same function in the format of a logic gate, a switch, an ASIC, a programmable logic controller, an embedded microcontroller or the like by logically programming process steps. Such a controller may therefore be considered as a hardware component, and an apparatus included in the controller and configured to implement various

functions may also be considered as a structure within the hardware component. Or, an apparatus for implementing various functions may be considered as both a software module for an implementation method and a structure within a hardware component.

[0141] The systems, apparatuses, modules or units illustrated in the above embodiments may specifically be implemented by a computer chip or an entity, or by a product having a certain function. A typical implementation device is a computer. Specifically, the computer may be, for example, a personal computer, a laptop computer, a cellular phone, a camera phone, a smart phone, a personal digital assistant, a media player, a navigation device, an email device, a game console, a tablet computer, a wearable device, or a combination of any of these devices.

[0142] For the convenience of description, the above devices are described separately as various units divided by function. Of course, the functions of the various units may be implemented in one or more pieces of software and/or hardware in the implementation of one or more embodiments of the present specification.

[0143] Those skilled in the art should appreciate that embodiments of the present specification may be provided as a method, a system, or a computer program product. Accordingly, the embodiments of the present application may take the form of an entirely hardware embodiment, an entirely software embodiment, or an embodiment combining software and hardware. Moreover, one or more embodiments of the present specification may take the form of a computer program product implemented on one or more computer usable storage media (including, but not limited to, disk memories, CD-ROMs and optical memories) comprising computer usable program codes.

[0144] The embodiments of the present specification are described with reference to the flowcharts and/or the block diagrams of a method, a device (system), and a computer program product according to the embodiments of the present specification. It should be understood that

each process and/or block in the flowcharts and/or block diagrams, and combinations of the processes and/or blocks in the flowcharts and/or the block diagrams, may be implemented by computer program instructions. These computer program instructions may be provided to a processor of a general-purpose computer, a special-purpose computer, an embedded processor, or other programmable data processing devices to produce a machine such that instructions are executed by the processor of the computer or other programmable data processing devices to generate an apparatus for implementing the functions specified in one or more processes in the flowcharts and/or one or more blocks in the block diagrams.

[0145] These computer program instructions may also be stored in a computer readable memory that can direct a computer or other programmable data processing devices to function in a particular manner such that the instructions stored in the computer readable memory produce an article of manufacture including an instruction means which implements functions specified in one or more processes in the flowcharts and/or one or more blocks in the block diagrams.

[0146] These computer program instructions may also be loaded onto a computer or other programmable data processing devices to cause a series of operating steps to be performed on the computer or other programmable devices to produce computer-implemented processing, and the instructions executed on the computer or other programmable devices provide steps for implementing the functions specified in one or more processes in the flowcharts and/or one or more blocks in the block diagrams.

[0147] In a typical configuration, a computing device includes one or more processors (CPUs), input/output interfaces, network interfaces, and a memory.

[0148] The memory may be in a form of a non-permanent memory, a random access memory (RAM), and/or a nonvolatile memory, such as a read-only memory (ROM) or a flash RAM, in computer-readable media. The memory is an example of computer-readable media.

[0149] Computer readable media include permanent and non-permanent, removable and non-removable media and may store information by any method or technology. The information may be a computer readable instruction, a data structure, a module of a program, or other data. Examples of computer storage media include, but are not limited to, a phase change memory (PRAM), a static random access memory (SRAM), a dynamic random access memory (DRAM), other types of random access memories (RAMs), a read only memory (ROM), an electrically erasable programmable read-only memory (EEPROM), a flash memory or other memory technologies, a compact disk read-only memory (CD-ROM), a digital versatile disc (DVD) or other optical storage devices, a magnetic tape cassette, a magnetic tape/disk storage or other magnetic storage devices, or any other non-transmission medium, and may be used for storing information accessible by computing devices. As defined herein, the computer readable media do not include transitory media, such as modulated data signals and carriers.

[0150] It should be further noted that the term “include,” “comprise” or any other variations thereof are intended to cover non-exclusive inclusions such that a process, method, article, or device that includes a series of elements not only includes those elements but also includes other elements that are not listed explicitly, or also includes inherent elements of the process, method, article, or device. In the absence of more limitations, an element defined by the sentence "including a/an..." does not exclude that the process, method, article or device including the element further has other identical elements.

[0151] Those skilled in the art should understand that embodiments of the present specification may be provided as a method, a system, or a computer program product. Accordingly, one or more embodiments of the present specification may take the form of an entirely hardware embodiment, an entirely software embodiment, or an embodiment combining software and hardware. Moreover, one or more embodiments of the present specification may take the form of a computer program product implemented on one or more computer usable storage media (including, but not limited to, disk memories, CD-ROMs and optical memories) comprising computer usable program codes.

[0152] One or more embodiments of the present specification may be described in the general context of computer executable instructions executed by a computer, such as program modules. Generally, the program modules include routines, programs, objects, components, data structures, or the like that perform particular tasks or implement particular abstract data types. One or more embodiments of the present specification may also be practiced in distributed computing environments where tasks are performed by remote processing devices that are connected through a communication network. In a distributed computing environment, program modules may be located in both local and remote computer storage media including storage devices.

[0153] The various embodiments in the present specification are described in a progressive manner, and the same or similar parts between the various embodiments may be referenced to each other, and each embodiment focuses on differences from other embodiments. In particular, because the system embodiments are basically similar to the method embodiments, they are described in a brief way, and the description of the method embodiments may be referred to for relevant parts.

[0154] The above descriptions are only embodiments of the present specification and are not intended to limit the present specification. For those skilled in the art, various modifications

and changes of the present specification may be made. Any modifications, equivalent replacements, improvements, etc., made within the spirit and scope of the present specification, are intended to be included within the scope of the claims of the present specification.

CLAIMS

1. An information detection method, comprising:

determining key point information in a target identification from a target image based on a preset deep learning algorithm;

obtaining an image of the target identification from the target image according to the key point information; and

determining information of a preset field from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification.

2. The method according to claim 1, wherein the determining the key point information in the target identification from the target image based on the preset deep learning algorithm comprises:

determining, based on data of a preset sample image, a relationship among key points corresponding to four right angles of an identification in the sample image and edges of the identification in the sample image and an avatar in the identification in the sample image, to construct a regression network model; and

determining the key point information in the target identification from the target image based on the constructed regression network model.

3. The method according to claim 2, wherein the regression network model comprises a first convolutional layer, a first pooling layer, a first dropout layer, a second convolutional layer, a second pooling layer, a second dropout layer, a third convolutional layer, a third dropout layer, a first fully connected layer, and a second fully connected layer, wherein the first convolutional

layer is connected to an input layer, the second fully connected layer is connected to an output layer, and layers are connected in an order of: the first convolutional layer, the first pooling layer, the first dropout layer, the second convolutional layer, the second pooling layer, the second dropout layer, the third convolutional layer, the third dropout layer, the first fully connected layer, and the second fully connected layer.

4. The method according to claim 3, wherein the first dropout layer, the second dropout layer, and the third dropout layer are used to increase preset noise information.

5. The method according to claim 4, wherein the first fully connected layer comprises a plurality of cells for mapping a learned distributed feature representation into a space of the sample image.

6. The method according to claim 5, wherein the second fully connected layer comprises eight cells, and the eight cells respectively correspond to the key points corresponding to the four right angles of the target identification and eight parameters.

7. The method according to claim 2, wherein the determining the key point information in the target identification from the target image based on the constructed regression network model comprises:

determining, based on the constructed regression network model, a region formed by the four right angles of the target identification from the target image; and

determining, based on the constructed regression network model, the key point information in the target identification from the determined region formed by the four right angles of the target identification.

8. The method according to claim 1, wherein before determining the information of the preset field from the image of the target identification according to the image of the target identification and the preset identification template matching the target identification, the method further comprises:

adjusting a size of the image of the target identification based on preset reference size information of the identification to obtain an image of the target identification matching the reference size information.

9. An information detection apparatus, comprising:

a key point determining module, configured to determine key point information in a target identification from a target image based on a preset deep learning algorithm;

an identification obtaining module, configured to obtain an image of the target identification from the target image according to the key point information; and

a field information determining module, configured to determine information of a preset field from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification.

10. The apparatus according to claim 9, wherein the key point determining module comprises:

a model constructing unit, configured to determine, based on data of a preset sample image, a relationship among key points corresponding to four right angles of an identification in the sample image and edges of the identification in the sample image and an avatar in the identification in the sample image, to construct a regression network model; and

a key point determining unit, configured to determine the key point information in the target identification from the target image based on the constructed regression network model.

11. The apparatus according to claim 10, wherein the regression network model comprises a first convolutional layer, a first pooling layer, a first dropout layer, a second convolutional layer, a second pooling layer, a second dropout layer, a third convolutional layer, a third dropout layer, a first fully connected layer, and a second fully connected layer, wherein the first convolutional layer is connected to an input layer, the second fully connected layer is connected to an output layer, and layers are connected in an order of: the first convolutional layer, the first pooling layer, the first dropout layer, the second convolutional layer, the second pooling layer, the second dropout layer, the third convolutional layer, the third dropout layer, the first fully connected layer, and the second fully connected layer.

12. The apparatus according to claim 11, wherein the first dropout layer, the second dropout layer, and the third dropout layer are used to increase preset noise information.

13. The apparatus according to claim 12, wherein the first fully connected layer comprises a plurality of cells for mapping a learned distributed feature representation into a space of the sample image.

14. The apparatus according to claim 13, wherein the second fully connected layer comprises eight cells, and the eight cells respectively correspond to the key points corresponding to the four right angles of the target identification and eight parameters.

15. The apparatus according to claim 10, wherein the key point determining unit is configured to

determine, based on the constructed regression network model, a region formed by the four right angles of the target identification from the target image; and

determine, based on the constructed regression network model, key point information in the target identification from the determined region formed by the four right angles of the target identification.

16. The apparatus according to claim 9, further comprising:

an image adjusting module, configured to adjust a size of the image of the target identification based on preset reference size information of the identification to obtain an image of the target identification matching the reference size information.

17. An information detection device, comprising:

a processor; and

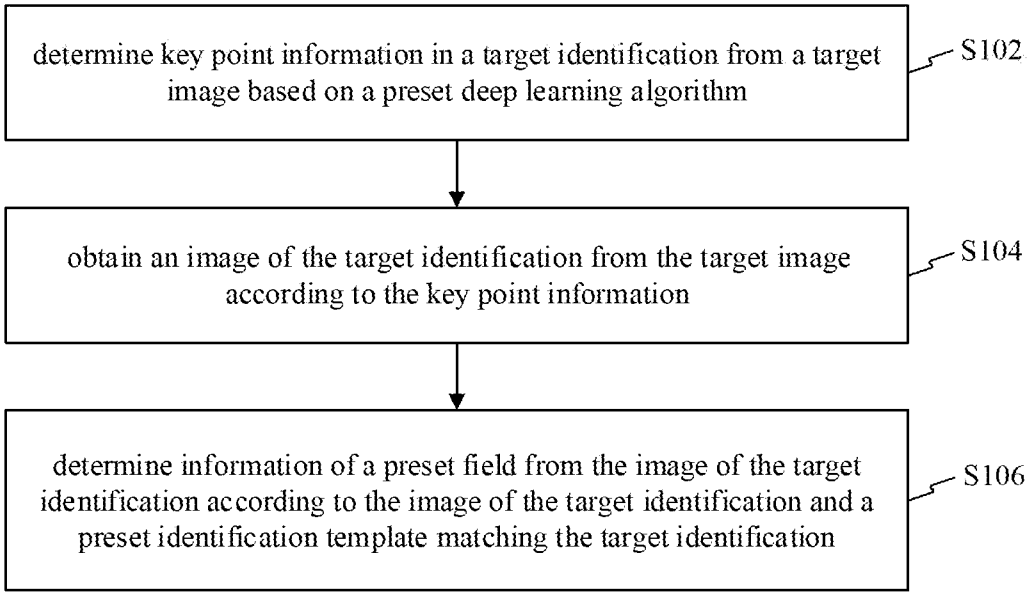
a memory arranged to store computer executable instructions that, when executed, cause the processor to:

determine key point information in a target identification from a target image based on a preset deep learning algorithm;

obtain an image of the target identification from the target image according to the key point information; and

determine information of a preset field from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification.

FIGURES


determine key point information in a target identification from a target image based on a preset deep learning algorithm
S102
obtain an image of the target identification from the target image according to the key point information
S104
determine information of a preset field from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification
S106

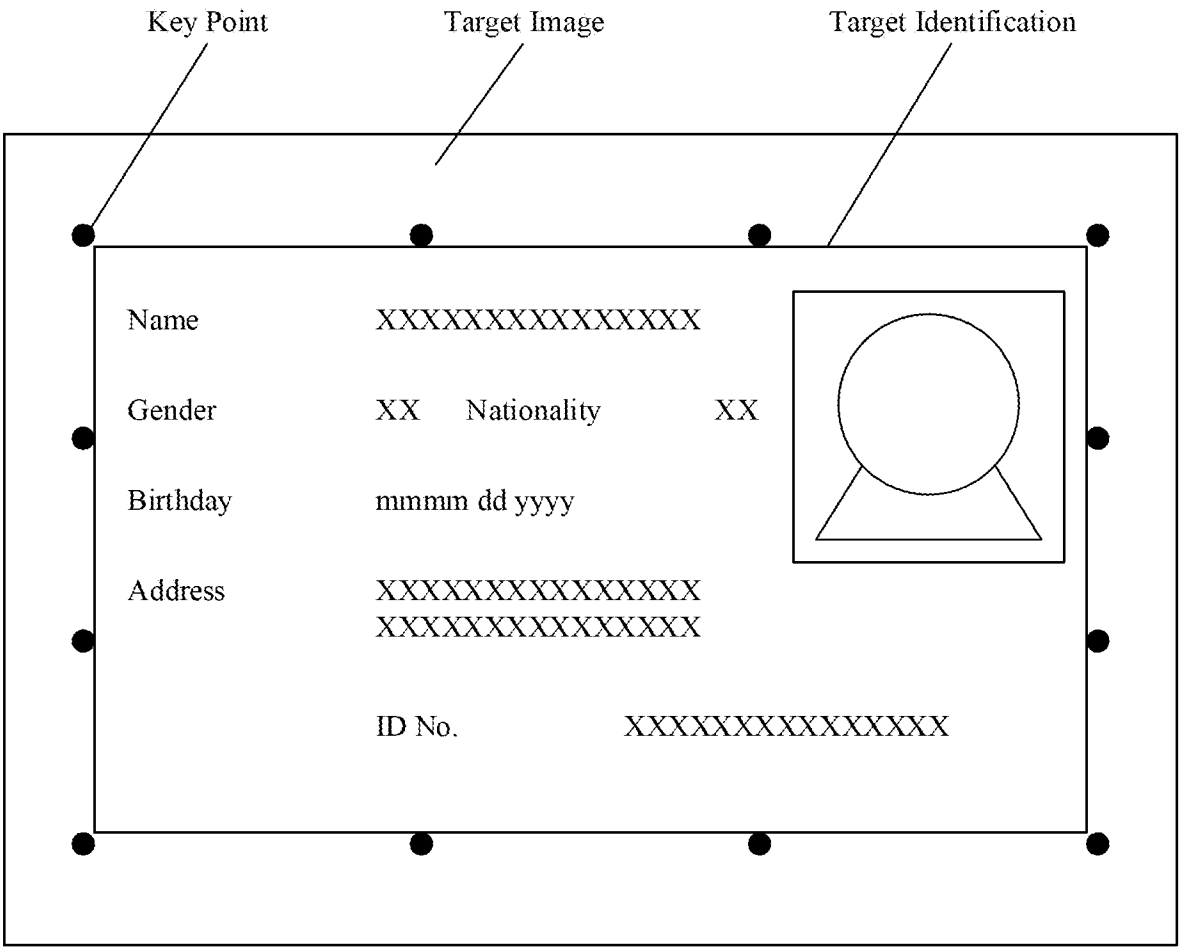

Key Point
Target Image
Target Identification
Name XXXXXXXXXXXXXXX
Gender XX Nationality XX
Birthday mmmm dd yyyy
Address XXXXXXXXXXXXXXX
XXXXXXXXXXXXXXX
ID No. XXXXXXXXXXXXXXX

FIG. 2

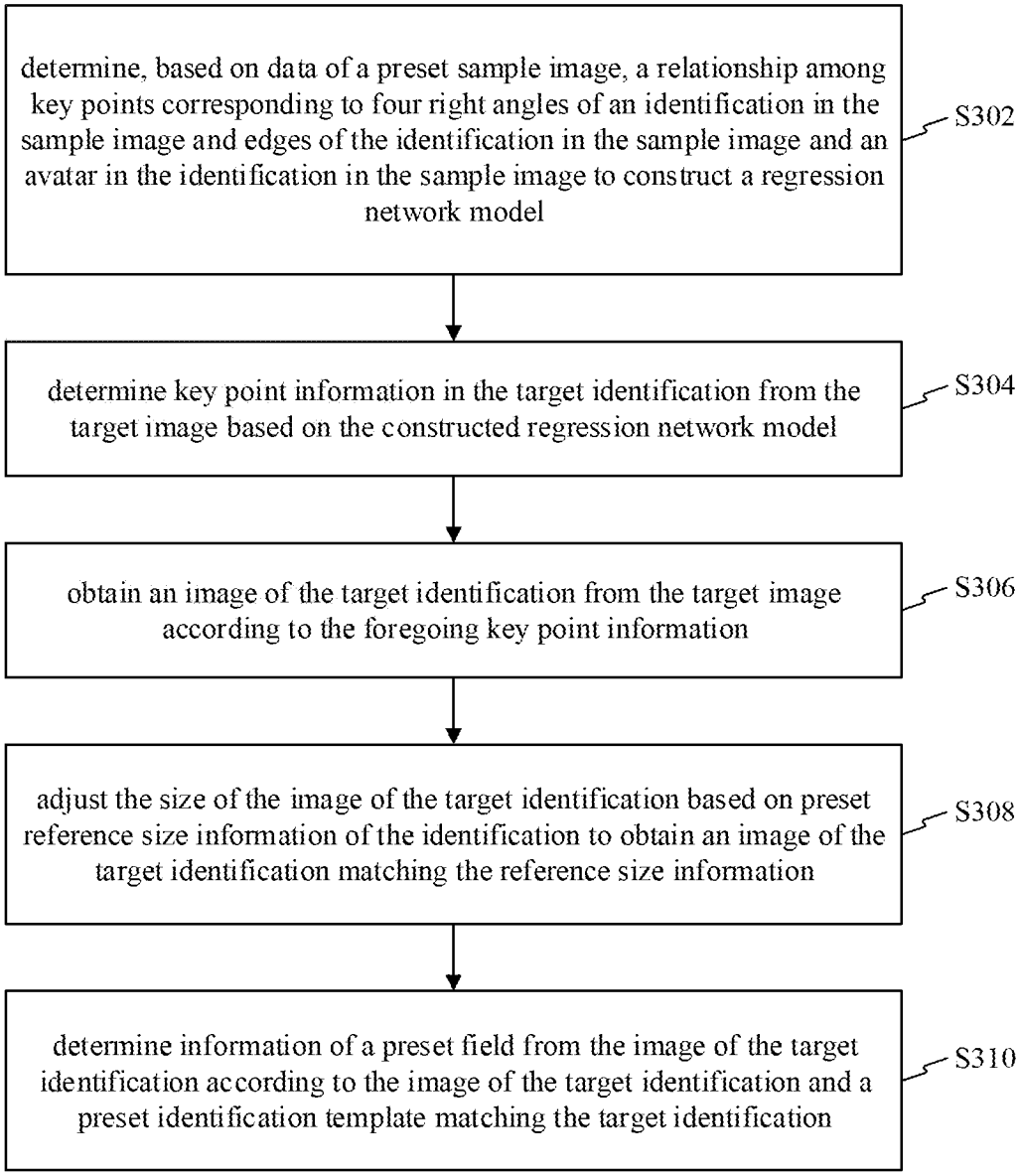

determine, based on data of a preset sample image, a relationship among key points corresponding to four right angles of an identification in the sample image and edges of the identification in the sample image and an avatar in the identification in the sample image to construct a regression network model
S302
determine key point information in the target identification from the target image based on the constructed regression network model
S304
obtain an image of the target identification from the target image according to the foregoing key point information
S306
adjust the size of the image of the target identification based on preset reference size information of the identification to obtain an image of the target identification matching the reference size information
S308
determine information of a preset field from the image of the target identification according to the image of the target identification and a preset identification template matching the target identification
S310

FIG. 3

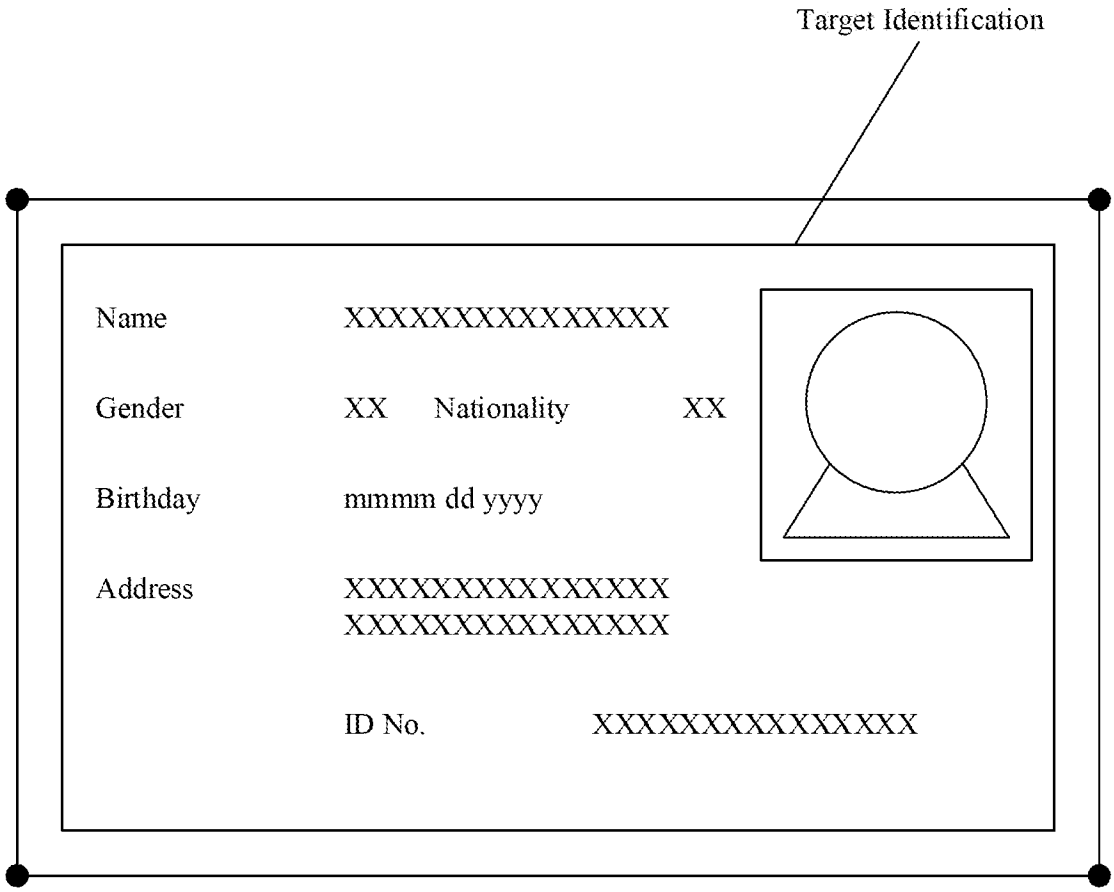

Target Identification
Name XXXXXXXXXXXXXXX
Gender XX Nationality XX
Birthday mmmm dd yyyy
Address XXXXXXXXXXXXXXX
XXXXXXXXXXXXXXX
ID No. XXXXXXXXXXXXXXX

FIG. 4

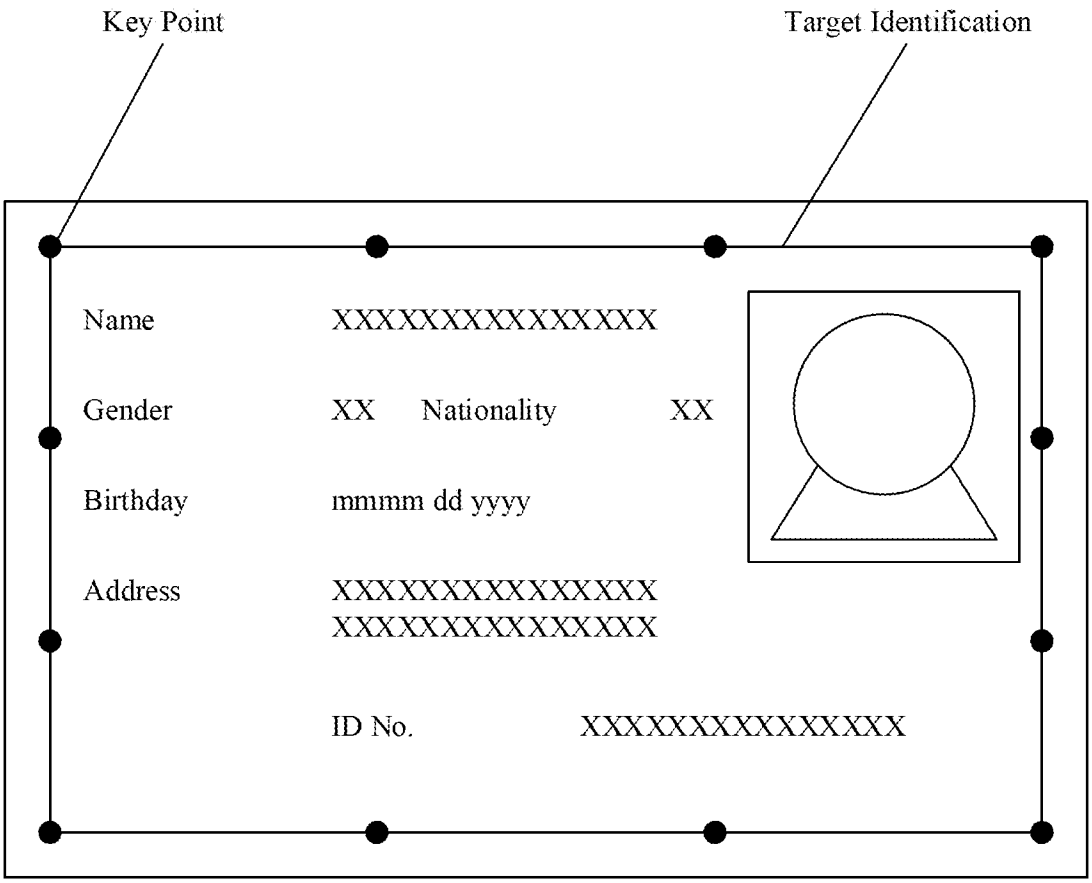

Key Point
Target Identification
Name
XXXXXXXXXXXXXXX
Gender
XX
Nationality
XX
Birthday
mmmm dd yyyy
Address
XXXXXXXXXXXXXXX
XXXXXXXXXXXXXXX
ID No.
XXXXXXXXXXXXXXX

FIG. 5

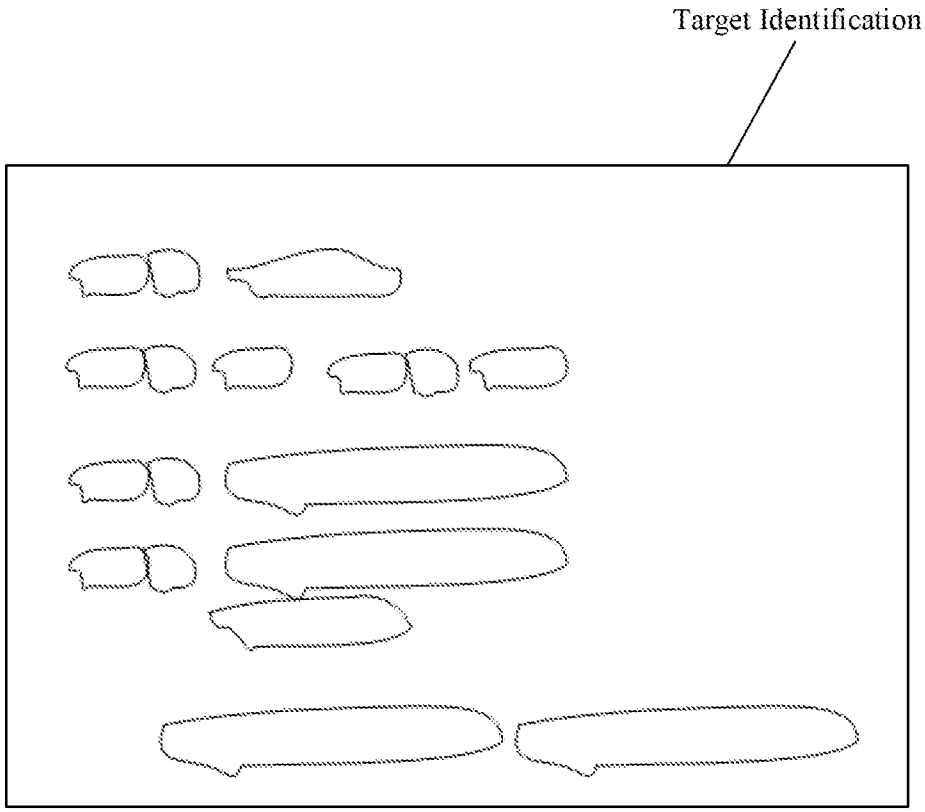

Target Identification

FIG. 6

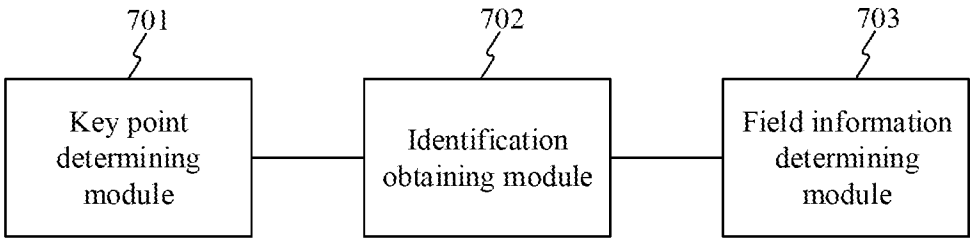

701
Key point determining module
702
Identification obtaining module
703
Field information determining module

FIG. 7

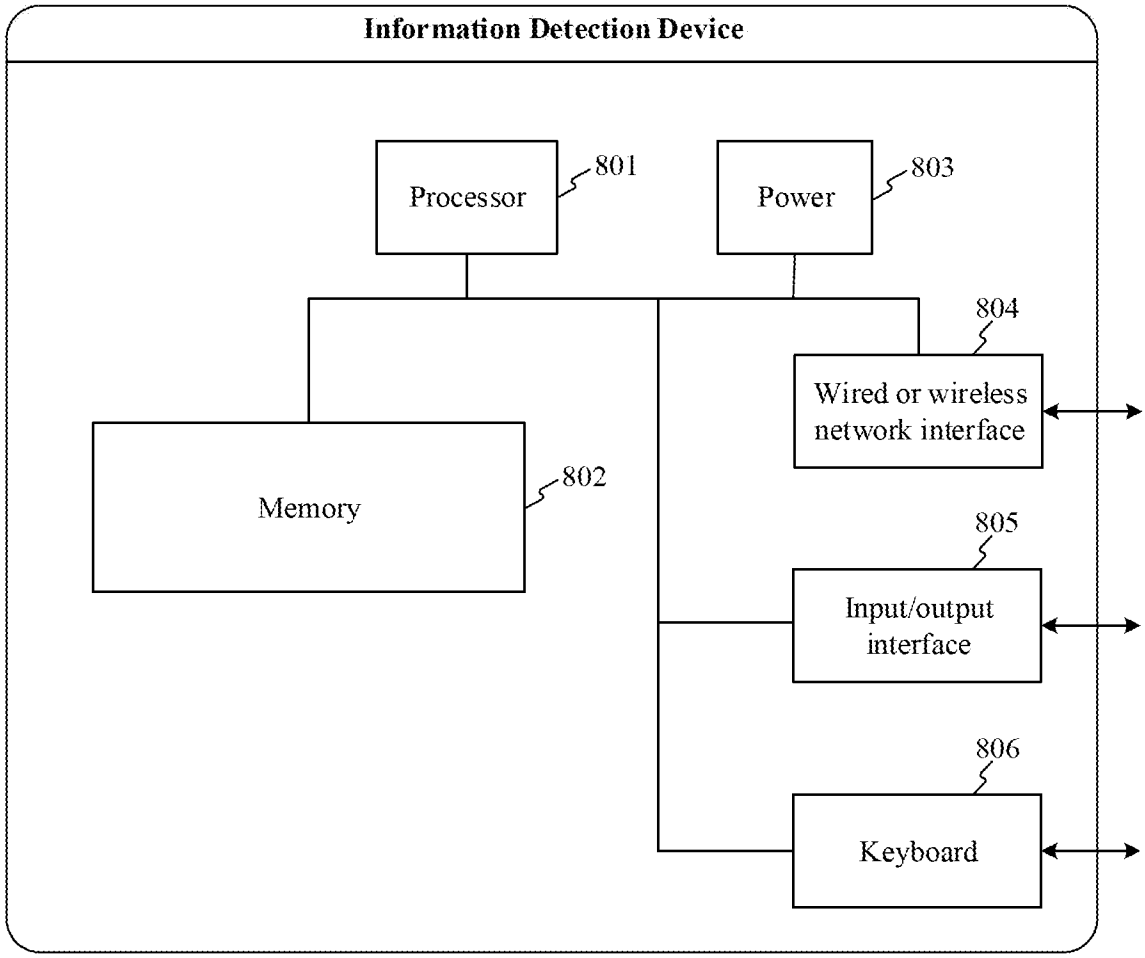

Information Detection Device
Processor
801
Power
803
804
Wired or wireless network interface
Memory
802
805
Input/output interface
806
Keyboard

FIG. 8

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2019/049291

A. CLASSIFICATION OF SUBJECT MATTER
INV. G06K9/46 G06K9/00
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	Sindhu: "Information Extraction from ID Cards - Using YOLOv2", , 4 July 2018 (2018-07-04), pages 1-4, XP055632832, Retrieved from the Internet: URL:https://blog.francium.tech/information-extraction-from-id-cards-using-yolov2-e3c846cb3796 [retrieved on 2019-10-16] page 1, line 1 - page 4, line 3 figures Figures on pages 3-4 ----- -/--	1-17

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
17 October 2019	24/10/2019

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Karwe, Markus

1

Form PCT/ISA/210 (second sheet) (April 2005)

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2019/049291

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2018/108101 A1 (RODRIGUEZ RAPHAEL A [US] ET AL) 19 April 2018 (2018-04-19) abstract paragraph [0003] - paragraph [0019] paragraph [0060] - paragraph [0093] paragraph [0107] - paragraph [0124] -----	1-17
A	Tae Jong Kim ET AL: "Crosscheck of Passport Information for Personal Identification" In: "Graphics Recognition. Ten Years Review and Future Perspectives", 31 December 2006 (2006-12-31), Springer Berlin Heidelberg, Berlin, Heidelberg, XP055632874, ISBN: 978-3-540-34711-8 vol. 3926, pages 162-172, DOI: 10.1007/11767978_15, abstract page 1 - page 2 page 4 - page 5 page 6 -----	1-17

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2019/049291

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2018108101 A1	19-04-2018	US 2018108101 A1	19-04-2018
		US 2019197642 A1	27-06-2019