METHODS AND APPARATUSES FOR IDENTITY DOCUMENT PROCESSING

FIELD OF INVENTION

[001] The present specification relates broadly, but not exclusively, to methods and apparatuses for identity document processing.

BACKGROUND

[002] In some Electronic Know Your Customer (eKYC) processes, a customer is required to provide a picture of a government-recognized identity document (ID) to prove their identity. Nowadays, an ID picture can be provided using mobile devices, such as by capturing a picture of an ID card using a mobile phone camera. Key information from the ID card, such as name, ID number and date of birth, can then be extracted from the ID picture by optical character recognition (OCR). Besides obtaining key information from the ID card, it is also important to check the authenticity of the ID card when processing the ID picture. This is to prevent, for example, fraudsters providing an ID picture that is not captured from the physical ID card in his possession, but from a screenshot of the ID card or from a printed copy of ID on paper.

[003] In existing methods, the customer is required to capture only one ID picture for the eKYC process. For a successful OCR, the customer is likely to place the camera close to the ID card to capture a close and clear image. In the meantime, background details surrounding the ID card are barely captured. However, such background information is useful in checking the authenticity of the ID card (i.e., a physical card rather than a screenshot or paper print). While a close and clear image is helpful for extracting correct information by OCR, checking the authenticity of the ID card is compromised due to the presence of little background information. Therefore, when a single ID picture captured from one angle and one distance is used in existing methods, difficulties exist in balancing the accuracies of both OCR function and fraud detection function.

SUMMARY

[004] Described embodiments provide methods and apparatuses for identity document (ID) processing which include extracting texts on the ID and authenticating the ID. In some embodiments, the method requests a user to capture two images of his / her ID card from two different distances for processing, for example, in an Electronic Know Your Customer (eKYC) process. The ID image captured from the shorter distance shows clear image of the key information of the ID, which is used for extracting key information by optical character recognition (OCR). The ID picture captured from the longer distance includes details of background surrounding the ID card, which is advantageously used for checking the authenticity of the ID card. In some implementations, if the ID image is captured from a screenshot of a physical card, the background can include boundary of the screen, which is helpful in authenticating the ID and determining that it is a picture of a non-authentic ID card. In some implementations, authenticating the ID includes adopting a Convolutional Neural Network (CNN) based binary classifier, which can be trained on images of both authentic, physical ID cards and non-authentic ID cards (e.g., a screenshot of an ID card, or a paper print of an ID card).

[005] According to one embodiment, there is provided a method for ID processing. The method includes: capturing, by an image acquisition module, a first image of an ID from a first effective position, such that texts on the ID can be extracted from the first image; capturing, by the image acquisition module, a second image of the ID from a second effective position, such that the second image comprises the ID and a background thereof; extracting the texts on the ID from the first image; and authenticating the ID, by an authentication module, based on at least the background in the second image.

[006] According to other embodiments, one or more of these general and specific embodiments may be implemented using an apparatus including a plurality of modules, a system, a method, or a computer-readable media, or any combination of devices, systems, methods, and computer-readable media. The foregoing and other described embodiments can each, optionally, include some, none or all of the following embodiments.

BRIEF DESCRIPTION OF THE DRAWINGS

[007] Embodiments and implementations are provided by way of example only, and will be better understood and readily apparent to one of ordinary skill in the art from the following written description, read in conjunction with the drawings, in which:

[008] Figure 1 is a flow chart illustrating an example of a method for identity document (ID) processing, according to an embodiment.

[009] Figure 2 is a flow diagram illustrating an example eKYC process using the method for ID processing in Figure 1, according to an embodiment.

[0010] Figure 3 is a schematic of an example of capturing an image of an ID card from two distances, according to an embodiment.

[0011] Figures 4 is a schematic diagram of an example of modules of an apparatus for ID processing, according to an embodiment.

[0012] Figure 5 is a block diagram of an example of a computer system suitable for executing at least some steps of the example methods shown in Figures 1 and 2, according to an embodiment.

[0013] Skilled artisans will appreciate that elements in the figures are illustrated for simplicity and clarity and have not necessarily been depicted to scale. For example, the dimensions of some of the elements in the illustrations, block diagrams or flowcharts may be exaggerated in respect to other elements to help to improve understanding of the present embodiments.

DETAILED DESCRIPTION

[0014] Embodiments will be described, by way of example only, with reference to the drawings. Like reference numerals and characters in the drawings refer to like elements or equivalents.

[0015] Some portions of the description which follows are explicitly or implicitly presented in terms of algorithms and functional or symbolic representations of operations on data within a computer memory. These algorithmic descriptions

and functional or symbolic representations are the means used by those skilled in the data processing arts to convey most effectively the substance of their work to others skilled in the art. An algorithm is here, and generally, conceived to be a self-consistent sequence of steps leading to a desired result. The steps are those requiring physical manipulations of physical quantities, such as electrical, magnetic or optical signals capable of being stored, transferred, combined, compared, and otherwise manipulated.

[0016] Unless specifically stated otherwise, and as apparent from the following, it will be appreciated that throughout the present specification, discussions utilizing terms such as “processing”, “comparing”, “requesting”, “indicating”, “capturing”, “authenticating”, “classifying”, “receiving”, “obtaining”, “training”, “determining”, “calculating”, “generating”, “detecting”, “indicating”, “changing”, “performing”, “comparing”, “updating”, “extracting”, ”representing”, “checking”, “outputting”, or the like, refer to the action and processes of a computer system, or similar electronic device, that manipulates and transforms data represented as physical quantities within the computer system into other data similarly represented as physical quantities within the computer system or other information storage, transmission or display devices.

[0017] The present specification also discloses apparatuses for performing the operations of the methods. Such apparatuses may be specially constructed for the required purposes, or may comprise a computer or other device selectively activated or reconfigured by a computer program stored in the computer. The algorithms and displays presented herein are not inherently related to any particular computer or other apparatus. Various machines may be used with programs in accordance with the teachings herein. Alternatively, the construction of more specialized apparatus to perform the required method steps may be appropriate. The structure of a computer suitable for executing the various methods / processes described herein will appear from the description below.

[0018] In addition, the present specification also implicitly discloses a computer program, in that it would be apparent to the person skilled in the art that the individual steps of the method described herein may be put into effect by computer code. The computer program is not intended to be limited to any particular programming language and implementation thereof. It will be appreciated that a variety of programming languages and coding thereof may be

used to implement the teachings of the specification contained herein. Moreover, the computer program is not intended to be limited to any particular control flow. There are many other variants of the computer program, which can use different control flows without departing from the scope of the specification.

[0019] Furthermore, one or more of the steps of the computer program may be performed in parallel rather than sequentially. Such a computer program may be stored on any computer readable medium. The computer readable medium may include storage devices such as magnetic or optical disks, memory chips, or other storage devices suitable for interfacing with a computer. The computer readable medium may also include a hard-wired medium such as exemplified in the Internet system, or wireless medium such as exemplified in the GSM mobile telephone system. The computer program when loaded and executed on such a computer effectively results in an apparatus that implements the steps of the preferred method.

[0020] The present specification may also be implemented as hardware modules. More particularly, in the hardware sense, a module is a functional hardware unit designed for use with other components or modules. For example, a module may be implemented using discrete electronic components, or it can form a portion of an entire electronic circuit such as an Application Specific Integrated Circuit (ASIC) or Field Programmable Gate Array (FPGA). Numerous other possibilities exist. Those skilled in the art will appreciate that the system can also be implemented as a combination of hardware and software modules.

[0021] ID authentication is the act or process of verifying the authenticity of an identity and can be considered as a form of fraud detection or fake identity detection, in which legitimacy of users are verified and potential fraudsters may be detected before fraudulent acts are carried out. Effective ID authentication can enhance data security of systems by permitting only authenticated users with valid identities to access its protected resources. Embodiments seek to provide methods and apparatuses for ID processing with an authentication function, which can detect ID pictures of non-authentic IDs provided by fraudsters in an attempt to complete the eKYC process. Advantageously, financial risks such as money laundering and fraud can be effectively reduced or eliminated.

[0022] The techniques described in this specification produce one or more technical effects. In some implementations, two images of an ID card captured from two different effective distances are used for the eKYC process. A first ID image may be captured from a close distance for extracting key information by OCR, and a second ID image may be captured from a longer distance for authenticating the ID card by examining the background in the second ID image. Advantageously, such techniques are effective in detecting if it is a picture of an authentic, physical ID card, or a picture of a screenshot or a paper printed copy. In some implementations, examining the background in the second ID image and determining the authenticity of the ID card includes adopting a Convolutional Neural Network (CNN) based binary classifier, which is trained on both authentic ID cards images and non-authentic ID cards images, to accurately return the result of the authenticity of the ID card.

[0023] Figure 1 is a flow chart 100 illustrating an example method for ID processing including the following steps:

- 110: capturing, by an image acquisition module, a first image of ID from a first effective position, such that texts on the ID can be extracted from the first image;
- 120: capturing, by the image acquisition module, a second image of the ID from a second effective position, such that the second image comprises the ID and a background thereof;
- 130: extracting the texts on the ID from the first image; and
- 140: authenticating the ID, by an authentication module, based on at least the background in the second image.

[0024] At step 110, a first image of an ID is captured by an image acquisition module from a first effective position. The image acquisition module may be a camera of a mobile phone or a laptop. For an image acquisition module with a zoom lens, the “effective position” is considered to change if a focal length is changed. The first image should be a close and clear image of the ID such that texts on the ID can be extracted.

[0025] At step 120, a second image of the ID is captured by the image acquisition module from a second effective position. The second image should include the ID card and a background of where the ID card is placed. Comparing to the first image, the ID card in the second image is smaller, such that background of the ID card can be present in the image. In some implementations,

if the first effective position corresponds to a first distance from the ID and the second effective position corresponds to a second effective distance from the ID, the second effective distance is longer than the first effective distance. For example, after the first image is captured, the second image can be captured by moving the camera further from the ID.

[0026] At step 130, texts on the ID can be extracted from the first image, such as name, gender, date of birth, ID number, country, address, height, or the like. The texts can include characters, numbers and symbols. In some implementations, extracting the texts of ID is based on an OCR algorithm, such as tesseract.

[0027] At step 140, the ID is authenticated by an authentication module based on at least the background in the second image. Features in the background may advantageously help in determining if the ID in the second image is authentic or non-authentic. For example, if the second image is captured from a color copy of the ID printed on paper, the background may contain the boundary of the paper, and the ID is determined to be non-authentic.

[0028] In some implementations, authenticating the ID may also include comparing the ID in the first image and the second image. This is to verify that same ID is used in the two images. In implementations, an ID classification module can be used to classify the ID in the images (e.g., unified ID from a certain country, driver license, passport), and ensure that they are the same type of ID. The ID classification module can be based on algorithms for classifying electronic document images. In some implementations, classifying the ID can be done at the same time with the OCR step. In some implementations, authenticating the ID requires the ID in the first image and the second image to have the same ID type. Otherwise, the process is terminated.

[0029] In some implementations, comparing the ID in the two images may also include extracting and comparing the faces on the ID in the two images. This is to ensure that the same person’s ID is used. In implementations, a face recognition and face comparison module can be used to compare the faces. The face recognition module can be based on algorithms such as FaceNet, DeepID3, or the like. In some implementations, authenticating the ID requires the faces extracted from the two images to match. Otherwise, the process is terminated.

[0030] In some implementations, a Convolutional Neural Network (CNN)-based binary classifier can be used to authenticate the ID. The CNN-based binary classifier can be trained on historical data of images of both authentic and non-authentic IDs, with backgrounds of the IDs in the images. In some embodiments, the non-authentic ID images can include an image of a screenshot of an ID with a background of a screen or a boundary of the screen. In some embodiments, the non-authentic ID images can include an image of a paper printed ID with a background of paper or a boundary of paper. Different CNN architecture can be adopted, such as resnet18. In some implementations, after authenticating the ID based on the second image, the historical data can be updated by including the second image and a result of the ID's authenticity, such that the accuracy and robustness of the classifier can be improved.

[0031] Figure 2 is a flow diagram 200 illustrating an example eKYC process using the method for ID processing as described with reference to Figure 1. In the beginning of the process, a user launches the eKYC process. In the next step, the user captures a first ID image "I_near" 210 from a first distance L1 with a camera of a mobile device. The user may adjust the first distance L1 between the ID card and the mobile device such that texts on the ID in "I_near" 210 are clear for extraction. In some embodiments, "I_near" 210 contains only the ID card with minimal background surrounding the ID card, and the first distance L1 is a short distance from the ID card.

[0032] From I_near 210, key ID information such as name (surname, given name, middle name), sex, date of birth, address, card number, country, and type of ID can be extracted. One may appreciate that I_near 210 is suitable for extracting texts from the ID. However, it is difficult to check the authenticity of the ID card since it contains little background surrounding the ID card.

[0033] In the next step, the user captures a second ID image "I_long" 210 from a second distance L2 with the camera. The user may adjust the second distance L2 between the ID card and the mobile device such that background of the ID card can be included in "I_long" 220. In the present example, the second distance L2 is longer than the first distance L1. For a camera with a zoom lens, L1 and L2 are effective distances between the camera and the ID card, taking into consideration of the focal length.

[0034] In the present example, I_long 220 is an image of the ID card displayed on a screen and the background of the ID card shows a boundary 224 of the screen. From I_long 220, based on the features from the background of the ID (e.g. the screen boundary 224), the ID can be determined as non-authentic. One may appreciate that I_long 220 is useful in authenticating the ID. The eKYC process ends after texts are extracted from I_near 210 and the ID is authenticated based on I_long 220.

[0035] In some implementations, the eKYC process may prompt or display instructions to the user, guiding the user on capturing images of the ID card. In some embodiments, the eKYC process may interact with the user by displaying a frame or corners of a frame on the mobile device screen, indicating an area for the ID card to be placed. In the present embodiments, the area for the ID card in I_near 210 is larger than in I_long 220. In some implementations, authenticity of the ID can be determined based on the texts extracted from the first image of the ID, by comparing the information provided by the user with the information extracted from the ID. In some implementations, the ID can be in a form other than card, such as a passport.

[0036] Figure 3 is a schematic 300 of an example of capturing an image of an ID card from two distances L1 and L2. From a shorter distance L1, the ID image captures a smaller area, which can be the ID card only. From a longer distance L2, the ID image captures a larger area, including the ID card and a background thereof. In implementations, L1 and L2 can be distances corresponding to the first effective position and the second effective position for capturing the ID images, respectively. In the present example, the first and second effective positions are from the same angle with respect to the ID card, and only the distances are different. One may appreciate that, the first and second effective positions can also be from different angles, as long as the texts on the ID can be extracted from the first image and a background of the ID is included in the second image.

[0037] Figure 4 is a schematic diagram of an example apparatus 400 including modules for identity proofing. The apparatus 400 at least includes an image acquisition module 410 and an authentication module 420. With reference to Figure 1 and Figure 2, the image acquisition module can be configured to capture an image. The apparatus 400 may include more than one image acquisition module 410, and the first image and the second image can be

captured by the same or different image acquisition modules 410. The authentication module 420 can be configured to authentic the ID based on images captured by the image acquisition module 410. The apparatus 400 may additionally include an OCR module configured to extract texts from images, and a training module configured to train a classifier, such as a CNN-based binary classifier. The apparatus 400 may also include an ID classification module configured to classify an ID type of the ID in the images, and a face comparison module configured to compare faces on the ID in the first image and the second image. The apparatus 400 may also include an output module configured to output the information extracted by the OCR module and an authentication result obtained by the authentication module. One or more or any combination of these modules can be part of an apparatus for ID processing.

[0038] The system, apparatus, module, or unit illustrated in the previous embodiments can be implemented by using a computer chip or an entity, or can be implemented by using a product having a certain function. A typical embodiment device is a computer (and the computer can be a personal computer), a laptop computer, a cellular phone, a camera phone, a smartphone, a personal digital assistant, a media player, a navigation device, an email receiving and sending device, a game console, a tablet computer, a wearable device, or any combination of these devices. The modules described as separate parts may or may not be physically separate, and parts displayed as modules may or may not be physical modules, may be located in one position, or may be distributed on a number of network modules. Some or all of the modules can be selected based on actual demands to achieve the objectives of the solutions of the specification. A person of ordinary skill in the art can understand and implement the embodiments of the present application without creative efforts.

[0039] Figure 5 is a block diagram of an example computer system 500 suitable for executing at least some steps of the example methods shown in Figures 1 and 2. The following description of the computer system / computing device 500 is provided by way of example only and is not intended to be limiting.

[0040] As shown in Figure 5, the example computing device 500 includes a processor 502 for executing software routines. Although a single processor is shown for the sake of clarity, the computing device 500 may also include a multi-processor system. The processor 502 is connected to a communication infrastructure 506 for communication with other components of the computing device 500. The

communication infrastructure 506 may include, for example, a communications bus, cross-bar, or network.

[0041] The computing device 500 further includes a main memory 504, such as a random access memory (RAM), and a secondary memory 510. The secondary memory 510 may include, for example, a storage drive 512, which may be a hard disk drive, a solid state drive or a hybrid drive and/or a removable storage drive 514, which may include a magnetic tape drive, an optical disk drive, a solid state storage drive (such as a USB flash drive, a flash memory device, a solid state drive or a memory card), or the like. The removable storage drive 514 reads from and/or writes to a removable storage medium 518 in a well-known manner. The removable storage medium 518 may include magnetic tape, optical disk, non-volatile memory storage medium, or the like, which is read by and written to by removable storage drive 514. As will be appreciated by persons skilled in the relevant art(s), the removable storage medium 518 includes a computer readable storage medium having stored therein computer executable program code instructions and/or data.

[0042] In an alternative implementation, the secondary memory 510 may additionally or alternatively include other similar means for allowing computer programs or other instructions to be loaded into the computing device 500. Such means can include, for example, a removable storage unit 522 and an interface 520. Examples of a removable storage unit 522 and interface 520 include a program cartridge and cartridge interface (such as that found in video game console devices), a removable memory chip (such as an EPROM or PROM) and associated socket, a removable solid state storage drive (such as a USB flash drive, a flash memory device, a solid state drive or a memory card), and other removable storage units 522 and interfaces 520 which allow software and data to be transferred from the removable storage unit 522 to the computer system 500.

[0043] The computing device 500 also includes at least one communication interface 524. The communication interface 524 allows software and data to be transferred between computing device 500 and external devices via a communication path 526. In various embodiments of the specification, the communication interface 524 permits data to be transferred between the computing device 500 and a data communication network, such as a public data or private data communication network. The communication interface 524 may be used to exchange data between different computing devices 500 which such computing devices 500 form part an

interconnected computer network. Examples of a communication interface 524 can include a modem, a network interface (such as an Ethernet card), a communication port (such as a serial, parallel, printer, GPIB, IEEE 1394, RJ45, USB), an antenna with associated circuitry and the like. The communication interface 524 may be wired or may be wireless. Software and data transferred via the communication interface 524 are in the form of signals which can be electronic, electromagnetic, optical or other signals capable of being received by communication interface 524. These signals are provided to the communication interface via the communication path 526.

[0044] As shown in Figure 5, the computing device 500 further includes a display interface 528 which performs operations for rendering images to an associated display 530 and an audio interface 532 for performing operations for playing audio content via associated speaker(s) 534.

[0045] As used herein, the term "computer program product" may refer, in part, to removable storage medium 518, removable storage unit 522, a hard disk installed in storage drive 512, or a carrier wave carrying software over communication path 526 (wireless link or cable) to communication interface 524. Computer readable storage media refers to any non-transitory, non-volatile tangible storage medium that provides recorded instructions and/or data to the computing device 500 for execution and/or processing. Examples of such storage media include magnetic tape, CD-ROM, DVD, Blu-ray™ Disc, a hard disk drive, a ROM or integrated circuit, a solid state storage drive (such as a USB flash drive, a flash memory device, a solid state drive or a memory card), a hybrid drive, a magneto-optical disk, or a computer readable card such as a PCMCIA card and the like, whether or not such devices are internal or external of the computing device 500. Examples of transitory or non-tangible computer readable transmission media that may also participate in the provision of software, application programs, instructions and/or data to the computing device 500 include radio or infra-red transmission channels as well as a network connection to another computer or networked device, and the Internet or Intranets including e-mail transmissions and information recorded on Websites and the like.

[0046] The computer programs (also called computer program code) are stored in main memory 504 and/or secondary memory 510. Computer programs can also be received via the communication interface 524. Such computer programs, when executed, enable the computing device 500 to perform one or more features of embodiments discussed herein. In various embodiments, the computer programs,

when executed, enable the processor 502 to perform features of the above-described embodiments. Accordingly, such computer programs represent controllers of the computer system 500.

[0047] Software may be stored in a computer program product and loaded into the computing device 500 using the removable storage drive 514, the storage drive 512, or the interface 520. The computer program product may be a non-transitory computer readable medium. Alternatively, the computer program product may be downloaded to the computer system 500 over the communication path 526. The software, when executed by the processor 502, causes the computing device 500 to perform the necessary operations to execute the method as shown in Figures 1 and 2.

[0048] It is to be understood that the embodiment of Figure 5 is presented merely by way of example to explain the operation and structure of the system 500. Therefore, in some embodiments one or more features of the computing device 500 may be omitted. Also, in some embodiments, one or more features of the computing device 500 may be combined together. Additionally, in some embodiments, one or more features of the computing device 500 may be split into one or more component parts.

[0049] It will be appreciated that the elements illustrated in Figure 5 function to provide means for performing the various functions and operations of the system as described in the above embodiments.

[0050] It will be appreciated by a person skilled in the art that numerous variations and/or modifications may be made to the present specification as shown in the specific embodiments without departing from the scope of the specification as broadly described. The present embodiments are, therefore, to be considered in all respects to be illustrative and not restrictive.

CLAIMS

What is claimed is:

1. A computer-implemented method for identity document processing, comprising:

capturing, by an image acquisition module, a first image of an identity document (ID) from a first effective position, such that texts on the ID can be extracted from the first image;

capturing, by the image acquisition module, a second image of the ID from a second effective position, such that the second image comprises the ID and a background thereof;

extracting the texts on the ID from the first image; and

authenticating the ID, by an authentication module, based on at least the background in the second image.

2. The method as claimed in claim 1, wherein the first effective position corresponds to a first distance from the image acquisition module to the ID and the second effective position corresponds to a second effective distance from the image acquisition module to the ID, and wherein the second effective distance is longer than the first effective distance.

3. The method as claimed in claim 1 or 2, wherein authenticating the ID comprises performing a binary classification of the ID using a Convolutional Neural Network (CNN)-based classifier.

4. The method as claimed in claim 3, further comprising training the CNN-based classifier based on historical data of authentic and non-authentic ID images, wherein the authentic and non-authentic ID images comprise IDs and backgrounds thereof.

5. The method as claimed in claim 4, wherein the non-authentic ID image comprises an image of a screenshot of an ID, and the background thereof comprises a boundary of a screen.

6. The method as claimed in claim 4, wherein the non-authentic ID image comprises an image of a paper printed ID, and the background thereof comprises a boundary of paper.

7. The method as claimed in claim 4, further comprising:

updating the historical data by including the second image and a result of authenticating the ID in the second image.

8. The method as claimed in any one of the preceding claims, wherein extracting the text on the ID from the first image is based on an optical character recognition (OCR) algorithm.

9. The method as claimed in any one of the preceding claims, wherein authenticating the ID comprises authenticating the ID based on the texts extracted from the first image.

10. The method as claimed in any one of the preceding claims, wherein authenticating the ID comprises authenticating the ID based on a comparison of the ID in the first image and the ID in the second image.

11. The method as claimed in claim 10, wherein the comparison of the ID in the first image and the ID in the second image comprises an ID type classification and/or a face recognition process.

12. A system for identity document processing, the system comprising:

one or more processors; and

one or more computer-readable memories coupled to the one or more processors and having instructions stored thereon that are executable by the one or more processors to perform the method of any one of claims 1 to 11.

13. An apparatus for identity document processing, the apparatus comprising a plurality of modules for performing the method of any one of claims 1 to 11.