TAMPER DETECTION METHOD AND SYSTEM

TECHNICAL FIELD

[0001] The present disclosure relates broadly, but not exclusively, to identity fraud detection, including methods, devices and systems for detecting tampering of an identity card.

BACKGROUND

[0002] Identity fraud can generally be considered to be the use by one person of another person's personal information, without authorization, to commit a crime or to deceive or defraud that other person or a third person. One example is identity card (ID) fraud where the identity fraud is committed using fake or tampered identity card.

[0003] For example, in electronic know your customer (eKYC), which usually requires an image, e.g. a photo or a scan, of an ID card to be submitted, it has been observed from some such images that the ID card may feature a replacement face photo while details such as name, identification number, date of birth, and biometric data, etc. are original. The replacement face photo can match a scan of a live face during the eKYC process, but the personal information clearly belongs to another person. Such frauds can pose significant financial risks to businesses that rely on eKYC, as well as potential legal and security issues.

[0004] It may be desirable to provide methods, devices and system that can detect such tampering to prevent identity card fraud.

SUMMARY

[0005] An embodiment provides a tamper detection method comprising extracting image data of a face area of an identity (ID) card from an image of the ID card, applying a trained neural network to the extracted image data to calculate a

confidence score, and comparing the confidence score with a threshold to determine whether the ID card is tampered. Training the neural network comprises inputting each training sample of a plurality of training samples into a binary classifier for a classification based on a face photo of said training sample, and iteratively optimizing parameters of the neural network based on classification results of the plurality of training samples.

[0006] Another embodiment provides a system for tamper detection comprising at least one processor, and a computer-readable memory coupled to the processor and having instructions stored thereon. The instructions are executable by the processor to extract image data of a face area of an identity (ID) card from an image of the ID card, apply a trained neural network to the extracted image data to calculate a confidence score, and determine whether the ID card is tampered based on a comparison of the confidence score with a threshold. The trained the neural network comprises a binary classifier trained to classify each training sample of a plurality of training samples based on a face photo of said training sample, and neural network parameters iteratively optimized based on classification results of the plurality of training samples.

[0007] Another embodiment provides an apparatus comprising a receiving device for receiving an image of an identity (ID) card, and a processing device for extracting image data of a face area of the ID card from the received image. The processing device provides the extracted image data to a trained neural network to calculate a confidence score and determines whether the ID card is tampered based on a comparison of the confidence score with a threshold. The trained neural network comprises a binary classifier trained to classify each training sample of a plurality of training samples based on a face photo of said training sample, and neural network parameters iteratively optimized based on classification results of the plurality of training samples.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] Embodiments of the disclosure will be better understood and readily apparent to one of ordinary skill in the art from the following written description, by way of example only, and in conjunction with the drawings, in which:

[0009] Figure 1 shows a flow chart illustrating a method of detecting tampering of an identity card according to an embodiment.

[0010] Figure 2 shows a flow chart illustrating a detailed implementation of the method of Figure 1.

[0011] Figure 3 shows a flow chart illustrating a detailed implementation of an automatic image generation process for preparing negative training samples.

[0012] Figure 4 shows a schematic diagram of an apparatus suitable for implementing the method of Figure 1.

[0013] Figure 5 shows a schematic diagram illustrating a computer system suitable for implementing the method of Figure 1.

[0014] 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

[0015] The present disclosure provides a method to detect tampering, particularly photo replacement, in identity card fraud. In the present disclosure, tamper detection is treated as a classification task, where a classification model of two classes, i.e. real (i.e. untampered) card and fake (i.e. tampered) card with photo replacement. Moreover, an attention area (e.g. a face area) of the identity card, instead of the whole captured image, is used to train the classification model. In other words, by focusing more attention on the area of the identity card where the tampering may be performed, the present tamper detection method can achieve high accuracy and speed.

[0016] 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.

[0017] 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.

[0018] Unless specifically stated otherwise, and as apparent from the following, it will be appreciated that throughout the present specification, discussions utilizing terms such as “scanning”, “calculating”, “determining”, “applying”, “extracting”, “generating”, “inputting”, “outputting”, “optimizing” 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.

[0019] The present specification also discloses apparatus for performing the operations of the methods. Such apparatus 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 conventional computer will appear from the description below.

[0020] 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 disclosure 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 disclosure.

[0021] 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, GPRS, 3G, 4G or 5G mobile telephone systems, as well as other wireless systems such as Bluetooth, ZigBee, Wi-Fi. The computer program when loaded and executed on such a computer effectively results in an apparatus that implements the steps of the preferred method.

[0022] The present disclosure may also be implemented as hardware elements. More particularly, in the hardware sense, an element is a functional hardware unit designed for use with other components or elements. For example, an element 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 elements.

[0023] According to various embodiments, a “circuit” may be understood as any kind of a logic implementing entity, which may be special purpose circuitry or a processor executing software stored in a memory, firmware, or any combination thereof. Thus, in an embodiment, a “circuit” may be a hard-wired logic circuit or a programmable logic circuit such as a programmable processor, e.g. a microprocessor (e.g. a Complex Instruction Set Computer (CISC) processor or a Reduced Instruction Set Computer (RISC) processor). A “circuit” may also be a processor executing software, e.g. any kind of computer program, e.g. a computer program using a virtual machine code such as e.g. Java. Any other kind of implementation of the respective functions which may be described in more detail herein may also be understood as a “circuit” in accordance with an alternative embodiment.

[0024] Figure 1 shows a flow chart 100 illustrating a method of detecting tampering of an identity (ID) card according to an embodiment. Examples of an ID card

include, but are not limited to, a driving license, a social security card, a passport page, and other officially issued documents that can be admitted as valid photo ID. At step 102, image data of a face area of the ID card is extracted from an image of the ID card. As described in further details below, this may involve a first cropping of the image to substantially contain just the ID card, and a second cropping to retain a portion of the image where the face area is located (e.g. on the bottom left of an ID card).

[0025] At step 104, a trained neural network is applied to the extracted image data to calculate a confidence score. The confidence score is then compared with a threshold at step 106 to determine whether the ID card is tampered. For example, if the confidence score is greater than the threshold, the photo on the ID card is determined to a replacement photo, and the ID card is determined to be a fake (i.e. tampered) ID card. On the other hand, if the confidence score is less than the threshold, the photo on the ID card is determined to be an original photo, and the ID card is determined to be a real (i.e. not tampered) ID card.

[0026] The neural network can be a convolutional neural network (CNN), such as ResNet-18 CNN, suitable for image classification tasks, and the confidence score can be calculated from the trained CNN using softmax function. The threshold can be determined based on a labeled validation dataset. More specifically, the validation dataset includes both normal ID card images and fake ID card images (i.e. with photo replacement), then a receiver operating characteristic (ROC) curve can be generated from the labels (i.e. 0 or 1) of the validation dataset and classification/prediction results (i.e. 0 or 1) from the binary classifier. In one embodiment, from the ROC curve, the threshold can be set when the false accepted rate (FAR) of the binary classifier is equal to 0.01. In other embodiment, the threshold can be set when the FAR of the binary classifier is equal to 0.001. It will be appreciated that different FAR values may be used in alternate embodiments, and the threshold may change accordingly.

[0027] In the present method, the neural network may be trained with a plurality of training samples (e.g. tens of thousands of samples). Training the neural network may include inputting each training sample of the plurality of training samples into a binary classifier for a classification based on the face photo of said training sample, and iteratively optimizing parameters of the neural network based on classification results of the plurality of training samples. During the training phase, both the

quantity and the quality of the training samples can affect the accuracy of the classifier.

[0028] Figure 2 shows a flow chart 200 illustrating a detailed implementation of the method of Figure 1, including a data processing phase 202, a training phase 204, and a testing phase 206.

[0029] In the data processing phase 202, a portion 208 of an ID card 210 that contains the face photo, also referred hereinafter as the face area, is extracted. An input image I1 of the ID card 210 in the data processing phase 202 can be a photo or a scan of the ID card 210. An image alignment step 212 is performed by detecting the four corners of the ID card 210 from the input image I1 to perform a first cropping, and aligning the cropped image I2 based on the four corners. In example embodiments, the resolution of the cropped image I2 is 400 x 300 pixels, and some background surrounding the ID card 210 may be included in the image I2. Then, a face cropping step (i.e. a second cropping) is performed to extract the face area 208 with the size of 200 x 300 pixels from the cropped image I2, followed by resizing the face area 208 to a rectangle of a size of 192 x 256 pixels.

[0030] The data processing phase 202 can be used to prepare training samples for the training phase 204. In addition, the data processing phase can be used to extract the face area of an ID card during testing phase 206 or actual use of the method for tamper detection.

[0031] With reference to the training phase 204, the training samples can include a plurality of positive training samples 214 and a plurality of negative training samples 216. The positive training samples 214 are obtained from actual images of ID cards, each of the ID cards having a respective face photo, using the data processing phase 202 as described above. In one implementation, about 150 thousand such positive training samples are used, but it will be appreciated that the quantity may vary depending on the training outcome (e.g. accuracy achieved).

[0032] The negative training samples 216 may be similarly obtained from actual images of ID cards, each with its face photo replaced. However, it is recognized that the process of physically/manually preparing such actual images requires repeatedly printing and trimming a physical face photo, covering the face photo on the face area of an ID card, and taking a picture, and can therefore be time-consuming and costly when large volumes of negative training samples (e.g. 150

thousands) are required. In a preferred embodiment, to efficiently prepare large quantities of negative training samples, the negative training samples are obtained from composite images using the data processing phase 202 above, with each composite image being generated by overlaying a face photo on an ID card photo based on an automatic image generation process, as described in further details below with reference to Figure 3.

[0033] In one implementation, the positive training samples 214 are labelled as 0, while the negative training samples 216 are labelled as 1. Each of the positive training samples 214 and negative training samples 216 has a size of 192 x 256 pixels, corresponding to a rectangle, instead of the size of 224 x 224 pixels normally used in CNNs such as ResNet-18 CNN. The size of 192 x 256 pixels can better match the face area of an ID card, since the shape of the face area is usually a rectangle, and not a square. Each training sample is provided to a neural network-based classifier, e.g. CNN classifier 218, for a binary classification based on the face photo of said training sample. In other words, face area images are used to train the neural network-based classifier, instead of the full ID card images. Therefore, in use, the trained classifier can pay more attention to the face area of ID card image and ignore the unnecessary regions of ID card image, and the accuracy of detecting a replaced photo can be significantly improved. The classification result can be a 0-value or a 1-value. The training phase 202 also includes iteratively optimizing parameters of the neural network based on classification results of the plurality of training samples.

[0034] In the testing phase 206 or actual use, a submitted image of an ID card first undergoes the data processing phase 202 as described above to obtain a cropped image 220 of the face area of the ID card. Then the cropped image 220 is provided to the trained neural network-based classifier 218 to obtain a confidence score. If the confidence score is more than a defined threshold, the submitted image is considered to be that of an ID card with the face photo replaced, i.e. the ID card has been tampered. On the other hand, if the confidence score is less than the defined threshold, the submitted image is considered to be that of an original ID card, i.e. the ID card has not been tampered.

[0035] The confidence score can be calculated using softmax function. In one example, the threshold is determined from a labeled validation dataset such that a false acceptance rate of the classifier 218 based on the threshold is equal to 0.01. In another embodiment, the threshold is determined from a labeled validation

dataset such that a false acceptance rate of the classifier 218 based on the threshold is equal to 0.001. Different false acceptance rates may be used in alternate embodiments.

[0036] Figure 3 shows a flow chart 300 illustrating a detailed implementation of an automatic image generation process for preparing negative training samples. First, a face image library including only face images from ID cards is created, and one face image (face_image) is randomly selected at step 302. An ID card image library including images of ID cards is also created, and one ID card image (card_image) is randomly selected at step 304. Next, a face detection is performed on the selected face image at step 306 to obtain a first rectangle frame (face_roi). Similarly, a face detection is performed on the selected card image at step 308 to obtain a second rectangle frame (card_roi) corresponding to a face photo. Then, at step 310, the first rectangle frame (face_roi) is cropped to extract the face photo (face_crop). In embodiments, the cropping can be performed by one of: (a) randomly moving the four endpoints of face_roi within a small range to obtain a random quadrilateral and crop the quadrilateral; (b) randomly generating n endpoints around face_roi, connecting two by two to obtain a random n-polygon, and cutting the n-polygon; and (c) performing a face segmentation algorithm on face_roi to obtain the avatar edge and trimming the avatar edge.

[0037] The cropped face photo (face_crop) then undergoes further processing. For example, the average brightness of the second rectangle frame (card_roi) is calculated, and the brightness of face_crop is adjusted to the average brightness of card_roi at step 312. In addition, shadow may be added to face_crop at a random angle. Then, at step 314, the size of face_crop is scaled to match that of card_roi, and face_crop is overlaid onto card_roi to generate a composite image. Further, Gaussian noise and Gaussian blur are randomly added to the composite image at step 316 to obtain the final fitted photo (tamper_image).

[0038] Steps 302 to 316 can be repeated in the automatic image generation process to rapidly generate negative training samples with realistic masking traces. The high quality of the generated training samples can improve the accuracy of the neural network-based classifier.

[0039] The automatic image generation process further includes hard samples mining. The purpose is to find out those negative samples that the current neural network-based classifier cannot recognize properly. The process collects “hard”

negative samples, and then sends them to further optimize the next generation/iteration of the neural network-based classifier. Specifically, at step 318, tamper_image is provided to the neural network-based classifier to determine if the image is one with a replacement photo. If the classifier correctly determines that tamper_image is indeed one with a replacement photo, which is a normal outcome, the sample is considered an easy/regular sample and is ignored (step 320). On the other hand, if the classifier wrongly determines that tamper_image is one with an original photo (i.e. no replacement), which is a false negative outcome, the sample is considered a “hard” sample and is saved (step 322). The collected hard samples can then be used to retrain the neural network-based classifier by updating and optimizing the neural network parameters to substantially reduce or minimise instances of false negatives.

[0040] Figure 4 shows a schematic diagram of an apparatus 400 suitable for implementing the method of Figure 1. The apparatus 400 includes a receiving device 402 connected to a processing device 404. The receiving device 402 can receive an image 406 of an identity (ID) card. The processing device 404 can extract image data of a face area of the ID card from the received image 406, provide the extracted image data to a trained neural network 408 to calculate a confidence score 410, and determine whether the ID card is tampered based on a comparison of the confidence score 410 with a threshold 412. The trained neural network 408 includes a binary classifier 414 trained to classify each training sample of a plurality of training samples based on a face photo of said training sample, and neural network parameters 416 iteratively optimized based on classification results of the plurality of training samples.

[0041] As described, the present method, system and apparatus can detect identity card fraud by detecting tampering of the face photo based on an image of the identity card. The tamper detection is performed by a classifier that is trained to focus on an attention area which generally corresponds to the face photo area of the identity card in shape, size and position, such that the classification has increased speed and accuracy. In some embodiments, large quantities of samples for training the classifier are generated by a rigorous automatic process a high quality. Furthermore, the classifier may be trained on selected hard samples that normally give false negatives, such that classification accuracy is improved.

[0042] Figure 5 depicts an exemplary computing device 500, hereinafter interchangeably referred to as a computer system 500, where one or more such computing devices 500 may be used for apparatus 400. The following description of the computing device 500 is provided by way of example only and is not intended to be limiting.

[0043] As shown in Figure 5, the example computing device 500 includes a processor 504 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 504 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.

[0044] The computing device 500 further includes a main memory 508, such as a random access memory (RAM), and a secondary memory 510. The secondary memory 510 may include, for example, a hard disk drive 512 and/or a removable storage drive 514, which may include a floppy disk drive, a magnetic tape drive, an optical disk drive, or the like. The removable storage drive 514 reads from and/or writes to a removable storage unit 518 in a well-known manner. The removable storage unit 518 may include a floppy disk, magnetic tape, optical disk, 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 unit 518 includes a computer readable storage medium having stored therein computer executable program code instructions and/or data.

[0045] 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, 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.

[0046] 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 disclosure, 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, 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.

[0047] As shown in Figure 5, the computing device 500 further includes a display interface 502 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.

[0048] As used herein, the term "computer program product" may refer, in part, to removable storage unit 518, removable storage unit 522, a hard disk installed in hard disk 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 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 floppy disks, magnetic tape, CD-ROM, DVD, Blu-ray™ Disc, a hard disk drive, a ROM or integrated circuit, USB memory, 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.

[0049] The computer programs (also called computer program code) are stored in main memory 508 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 504 to perform features of the above-described embodiments. Accordingly, such computer programs represent controllers of the computer system 500.

[0050] Software may be stored in a computer program product and loaded into the computing device 500 using the removable storage drive 514, the hard disk drive 512, or the interface 520. Alternatively, the computer program product may be downloaded to the computer system 500 over the communications path 526. The software, when executed by the processor 504, causes the computing device 500 to perform functions of embodiments described herein.

[0051] It is to be understood that the embodiment of Figure 5 is presented merely by way of example. 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.

[0052] It will be appreciated that the elements illustrated in Figure 5 function to provide means for performing the various functions and operations of the servers as described in the above embodiments.

[0053] In an implementation, a server may be generally described as a physical device comprising at least one processor and at least one memory including computer program code. The at least one memory and the computer program code are configured to, with the at least one processor, cause the physical device to perform the requisite operations.

[0054] It will be appreciated by a person skilled in the art that numerous variations and/or modifications may be made to the present disclosure as shown in the specific embodiments without departing from the scope of the disclosure as broadly described. For example, the value of the threshold may be adjusted based on practical requirements. The present embodiments are, therefore, to be considered in all respects to be illustrative and not restrictive.

CLAIMS

1. A tamper detection method comprising:

extracting image data of a face area of an identity (ID) card from an image of the ID card;

applying a trained neural network to the extracted image data to calculate a confidence score, wherein training the neural network comprises:

inputting each training sample of a plurality of training samples into a binary classifier for a classification based on a face photo of said training sample; and

iteratively optimizing parameters of the neural network based on classification results of the plurality of training samples; and

comparing the confidence score with a threshold to determine whether the ID card is tampered.

2. The method as claimed in claim 1, wherein the confidence score is calculated using a softmax function.

3. The method as claimed in claim 1 or 2, wherein the threshold is determined from a labeled validation dataset such that a false acceptance rate of the binary classifier based on the threshold is equal to 0.01.

4. The method as claimed in claim 1 or 2, wherein the threshold is determined from a labeled validation dataset such that a false acceptance rate of the binary classifier based on the threshold is equal to 0.001.

5. The method as claimed in any one of the preceding claims, wherein the training samples comprise:

a plurality of positive training samples obtained from actual images of ID cards, each of the ID cards having a respective face photo;

a plurality of negative training samples obtained from composite images, wherein each composite image is generated by overlaying a face photo on an ID card photo based on an automatic image generation process.

6. The method as claimed in claim 5, wherein overlaying the face photo on the ID card photo based on the automatic image generation process comprises:

randomly selecting an image having the face photo and an image having the ID card photo from respective image libraries;

extracting the face photo and the ID card photo from the selected images, wherein the ID card photo comprises a region of interest where a face photo is located;

adjusting a size and a brightness of the extracted face photo based on the size and the brightness of the region of interest of the ID card photo; and

overlaying the size- and brightness-adjusted face photo onto the region of interest of the ID card photo.

7. The method as claimed in claim 6, adjusting the size and the brightness of the extracted face photo further comprises adding shadow at a random angle to the extracted face photo.

8. The method as claimed in claim 6 or 7, further comprising adding Gaussian noise and Gaussian blur to the composite image.

9. The method as claimed in any one of the preceding claims, wherein each training sample comprises a rectangular shape, and wherein the face area of the ID card comprises a rectangular shape of the same size as the training sample.

10. The method as claimed in any one of the preceding claims, wherein training the neural network further comprises:

identifying a set of negative training samples that provide false negative classification results;

adjusting parameters of the neural network; and

re-training the neural network based on the identified set of negative training samples and the adjusted parameters.

11. A system for tamper detection, comprising:

at least one processor; and

a computer-readable memory coupled to the processor and having instructions stored thereon that are executable by the processor to:

extract image data of a face area of an identity (ID) card from an image of the ID card;

apply a trained neural network to the extracted image data to calculate a confidence score, wherein the trained the neural network comprises:

a binary classifier trained to classify each training sample of a plurality of training samples based on a face photo of said training sample, and

neural network parameters iteratively optimized based on classification results of the plurality of training samples; and

determine whether the ID card is tampered based on a comparison of the confidence score with a threshold.

12. The system as claimed in claim 11, wherein the instructions are executable by the processor to calculate the confidence score using a softmax function.

13. The system as claimed in claim 11 or 12, wherein the memory further comprises instructions executable by the processor to determine the threshold from a labeled validation dataset such that a false acceptance rate of the binary classifier based on the threshold is equal to 0.01.

14. The system as claimed in claim 11 or 12, wherein the memory further comprises instructions executable by the processor to determine the threshold from a labeled validation dataset such that a false acceptance rate of the binary classifier based on the threshold is equal to 0.001.

15. The system as claimed in any one of claims 11 to 14, wherein the training samples comprise:

a plurality of positive training samples obtained from actual images of ID cards, each of the ID cards having a respective face photo; and

a plurality of negative training samples obtained from composite images, wherein each composite image comprises a face photo overlaid on an ID card photo.

16. The system as claimed in claim 15, wherein the face photo of each composite image comprises a size and a brightness adjusted based on the size and the brightness of a region of interest of the ID card photo where the face photo is located.

17. The system as claimed in claim 16, wherein the face photo of each composite image comprises an added shadow at a random angle.

18. The system as claimed in claim 16 or 17, wherein each composite image comprises added Gaussian noise and Gaussian blur.

19. The system as claimed in any one of claims 11 to 18, wherein each training sample comprises a rectangular shape, and wherein the face area of the ID card comprises a rectangular shape of the same size as the training sample.

20. An apparatus comprising:

a receiving device for receiving an image of an identity (ID) card; and

a processing device for extracting image data of a face area of the ID card from the received image, the processing device providing the extracted image data to a trained neural network to calculate a confidence score and determining whether the ID card is tampered based on a comparison of the confidence score with a threshold, wherein the trained neural network comprises:

a binary classifier trained to classify each training sample of a plurality of training samples based on a face photo of said training sample, and

neural network parameters iteratively optimized based on classification results of the plurality of training samples.