

# FACE LIVENESS DETECTION METHODS AND APPARATUSES

[0001] This application claims priority to Chinese Patent Application No. 202110966357.5, filed with the China National Intellectual Property Administration on August 23, 2021 and entitled "FACE LIVENESS DETECTION METHODS AND APPARATUSES", which is  
5 incorporated herein by reference in its entirety.

## TECHNICAL FIELD

[0002] One or more embodiments of this specification relate to the field of machine learning technologies, and in particular, to face liveness detection methods and apparatuses.

## BACKGROUND

10 [0003] Liveness detection is a technology for determining genuine physiological features of objects in some identity verification scenarios. Face liveness detection is commonly used in facial recognition applications to verify whether a user is operating as the genuine living person. Such practice can defend against attacks such as photos, face swapping, masks, and obstructions, thereby preventing user identity theft, detecting fraudulent behaviors, and safeguarding user  
15 interests.

[0004] However, with widespread applications of the face liveness detection technology, higher requirements are imposed on this application, for example, enhancement of security and user experience. Therefore, a face liveness detection solution is needed to ensure security of detection algorithms while improving user experience in using related services.

## SUMMARY

20 [0005] One or more embodiments of this specification describe face liveness detection methods and apparatuses. In a detection process, interactive actions between a photographing terminal and a user can be hidden, thereby effectively defending against attacks on detection algorithms and effectively improving user experience.

25 [0006] According to a first aspect, a face liveness detection method is provided, including: multiple face images continuously captured for a first user are obtained; two face images are selected from the multiple face images for any target attribute in one or more predetermined

image attributes, where a degree of difference between two attribute values of the target attribute corresponding to the two face images satisfies a predetermined criterion; and both of the two face images are input into a pre-trained face liveness detection model corresponding to the target attribute to obtain a detection result corresponding to the target attribute.

5 [0007] In an embodiment, that multiple face images continuously captured for a first user are obtained includes the following: based on a user terminal of the first user, the first user is prompted to adjust a photographing method to improve face image quality; and during the prompting, the multiple face images are continuously photographed by using the user terminal.

[0008] In a specific embodiment, that the first user is prompted to adjust a photographing  
10 method includes the following: prompting to adjust a position of a face relative to the user terminal; and/or prompting to adjust a light intensity of a photographing environment.

[0009] In an embodiment, that two face images are selected from the multiple face images for any target attribute in one or more predetermined image attributes includes the following: multiple attribute values of the target attribute corresponding to the multiple face images are  
15 determined; and two face images whose degree of difference between corresponding attribute values satisfies the predetermined criterion are selected from the multiple face images based on the multiple attribute values.

[0010] In a specific embodiment, when the target attribute is a face size, that multiple attribute values of the target attribute corresponding to the multiple face images are determined  
20 includes the following: the multiple face images are separately input into a trained face detection model to obtain multiple face regions corresponding to the multiple face images; and multiple face region areas corresponding to the multiple face regions are determined.

[0011] In another specific embodiment, when the target attribute is a facial organ size, that multiple attribute values of the target attribute corresponding to the multiple face images are  
25 determined includes the following: the multiple face images are separately input into a trained facial organ detection model to obtain multiple facial organ regions corresponding to the multiple face images; and multiple organ region areas corresponding to the multiple facial organ regions are determined.

[0012] In still another specific embodiment, when the target attribute is a light intensity,  
30 that multiple attribute values of the target attribute corresponding to the multiple face images are determined includes the following: the multiple face images are separately into a trained light intensity evaluation model to obtain multiple light intensities.

[0013] In another aspect, in a specific embodiment, that two face images whose degree of difference between corresponding attribute values satisfies the predetermined criterion are  
35 selected from the multiple face images based on the multiple attribute values includes the

following: a maximum value and a minimum value in the multiple attribute values are determined to obtain the two corresponding face images.

[0014] In a specific embodiment, that two face images whose degree of difference between corresponding attribute values satisfies the predetermined criterion are selected from the multiple face images based on the multiple attribute values includes the following: two face images whose degree of difference between the corresponding attribute values is greater than a predetermined threshold are selected from the multiple face images.

[0015] In an embodiment, the one or more image attributes are multiple image attributes; and after the obtaining a detection result corresponding to the target attribute, the method further includes the following: a final detection result for the first user is determined based on multiple detection results corresponding respectively to the multiple image attributes.

[0016] In a specific embodiment, that a final detection result for the first user is determined based on multiple detection results corresponding respectively to the multiple image attributes includes the following: when the multiple detection results all indicate that the first user succeeds in face liveness detection, the final detection result is determined as a face liveness detection success.

[0017] According to a second aspect, a face liveness detection apparatus is provided, including: an image acquisition unit, configured to obtain multiple face images continuously captured for a first user; an image selection unit, configured to select two face images from the multiple face images for any target attribute in one or more predetermined image attributes, where a degree of difference between two attribute values of the target attribute corresponding to the two face images satisfies a predetermined criterion; and a liveness detection unit, configured to input both of the two face images into a pre-trained face liveness detection model corresponding to the target attribute to obtain a detection result corresponding to the target attribute.

[0018] According to a third aspect, a computer-readable storage medium is provided, where the computer-readable storage medium stores a computer program, and when the computer program is executed in a computer, the computer is enabled to perform the method according to the first aspect.

[0019] According to a fourth aspect, a computing device is provided, including a memory and a processor. The memory stores executable code, and when executing the executable code, the processor implements the method according to the first aspect.

[0020] By using the methods and the apparatuses provided in the embodiments of this specification, face image-related user interaction information actually used in algorithms can be effectively hidden to improve security of face liveness detection algorithms. In addition,

accuracy and availability of detection results can be effectively improved by performing face liveness detection on face images with a relatively large attribute difference.

## **BRIEF DESCRIPTION OF DRAWINGS**

[0021] To describe the technical solutions of the embodiments of this specification more clearly, the following briefly describes the accompanying drawings needed for describing the embodiments. Clearly, the accompanying drawings in the following description show merely some embodiments of this specification, and a person of ordinary skill in the art can still derive other drawings from these accompanying drawings without creative efforts.

[0022] FIG. 1 is a schematic diagram illustrating an implementation architecture of face liveness detection, according to some embodiments;

[0023] FIG. 2 is a schematic flowchart illustrating a face liveness detection method, according to some embodiments; and

[0024] FIG. 3 is a schematic structural diagram illustrating a face liveness detection apparatus, according to some embodiments.

## **DESCRIPTION OF EMBODIMENTS**

[0025] The solutions provided in this specification are described below with reference to the accompanying drawings.

[0026] Currently, face liveness detection technologies are widely used. In a detection solution, a user is required to perform a combination of actions such as blinking, mouth opening, head shaking, nodding, etc. to verify whether the user operates as the genuine living person by using technologies such as facial landmark localization and face tracking. However, this solution exposes the user's interactive actions in a detection algorithm, and is prone to exploit and attacks by attackers. In another detection solution, interaction with a user is not needed. However, in this solution, it is difficult to collect face images with relatively large changes, and consequently reliability of detection results is low.

[0027] Based on the above-mentioned observation and analysis, the inventor proposes a new face liveness detection method, which can effectively hide facial action images actually used in algorithms, thereby improving security of the detection algorithms, and effectively improving accuracy and availability of detection results. FIG. 1 is a schematic diagram of an implementation architecture of face liveness detection, according to some embodiments. As shown in FIG. 1, multiple face images continuously captured for a first user are obtained, and N (which is a positive integer) face image pairs corresponding to N image attributes are

constructed based on the multiple face images, where each image pair has a large difference for an image attribute corresponding to the image pair; afterwards, the  $i$ th random face image pair is input into the  $i$ th face liveness detection model corresponding to the  $i$ th image attribute to obtain the  $i$ th detection result; and then, a final detection result of performing face liveness detection on the first user can be determined based on  $N$  obtained detection results. As such, an accurate face detection result can be obtained securely and efficiently.

**[0028]** Specific implementation steps of the above-mentioned method are described below with reference to some specific embodiments. FIG. 2 is a schematic flowchart of a face liveness detection method, according to some embodiments. The method can be executed by any server, apparatus, device cluster, or the like that has computing and processing capabilities. As shown in FIG. 2, the method includes the following steps.

**[0029]** Step S210: Obtain multiple face images continuously captured for a first user. Step S220: Select two face images from the multiple face images for any target attribute in one or more predetermined image attributes, where a degree of difference between two attribute values of the target attribute corresponding to the two face images satisfies a predetermined criterion. Step S230: Input both of the two face images into a pre-trained face liveness detection model corresponding to the target attribute to obtain a detection result corresponding to the target attribute.

**[0030]** The above-mentioned steps are elaborated as follows:

**[0031]** First, in step S210, multiple face images continuously captured for the first user are obtained.

**[0032]** In an implementation, when the face images are recorded, the user can be enabled to perceive adjustments on image quality, thereby capturing face images that have a high degree of difference. Such practice can avoid user experience deterioration and decrease of detection algorithm security that are caused by requiring the user to perform a specific combination of actions such as blinking and mouth opening. In addition, enough face images that have large changes (covering a photographing angle of the user, light of a photographing environment, and the like) can be obtained.

**[0033]** In an embodiment, this step can include the following: prompting, based on a user terminal of the first user from a perspective of improving face image quality, the first user to adjust a photographing method; and during the prompting, continuously photographing the multiple face images by using the user terminal. In a specific embodiment, the prompting the first user to adjust a photographing method can include the following: prompting the first user to adjust a position of the first user's face relative to the first user's user terminal. In an example, the first user is prompted to keep the first user's face at the center of a screen image. In an

example, the user is prompted to move away from or approach the user terminal to control a size of a photographed face or to photograph an entire face of the user. In another specific embodiment, the prompting the first user to adjust a photographing method can include the following: prompting the first user to adjust a light intensity of a photographing environment.

5 For example, when light is intense in a current environment, the first user is prompted to move to an area with low light to perform photographing. For another example, when light is low in a current environment, the first user is prompted to increase ambient brightness. As such, the user makes some adjustments according to prompts. Therefore, photos with a significant difference usually exist in multiple face images continuously captured for an adjustment process.

10 **[0034]** In another aspect, in a specific embodiment, the continuously photographing the multiple face images by using the user terminal can include the following: continuously photographing the multiple face images at a predetermined time interval. In another specific embodiment, the continuously photographing the multiple face images by using the user terminal can include the following: performing photographing for consecutive times until a  
15 predetermined quantity of face images are obtained and used as the multiple face images.

**[0035]** In another implementation, it is feasible that no photographing prompt is provided for the user. Instead, the user terminal quickly and continuously photographs the multiple face images in response to start triggering of a photographing function.

**[0036]** From the description above, multiple face images photographed for the first user can  
20 be obtained. Then, in step S220, the two face images are selected from the multiple face images for the any target attribute in the one or more predetermined image attributes, where the degree of difference between the two attribute values of the target attribute corresponding to the two face images satisfies the predetermined criterion.

**[0037]** It should be understood that, "several" in this specification refers to "one or more".

25 In addition, the one or more image attributes can be predetermined by an operator based on actual experience. For example, the one or more image attributes can include a size or an area of a face region, a size or an area of a facial organ (for example, a nose, eyes, a forehead, or the like), a light intensity, and the like of an image. Each of the one or more image attributes can be individually used as a target attribute.

30 **[0038]** In an implementation, this step can include the following: first determining multiple attribute values of the target attribute corresponding to the multiple face images; and then, selecting, from the multiple face images based on the multiple attribute values, two face images whose degree of difference between corresponding attribute values satisfies the predetermined criterion.

35 **[0039]** In an embodiment, when the target attribute is a face size in an image, the

determining corresponding multiple attribute values can include the following: separately inputting the multiple face images into a trained face detection model to obtain multiple face regions corresponding to the multiple face images; and determining multiple face region areas corresponding to the multiple face regions as multiple attribute values corresponding to the face size attribute. In a specific embodiment, the face detection model can be implemented based on a convolutional neural network (CNN) or a deep neural network (DNN). In a specific embodiment, the face detection model can be implemented based on a target detection algorithm. Specifically, a target detection model is trained by using faces as target objects, and is used as the face detection model. As such, a regression box for selecting a face position in any face image can be obtained by inputting the face image into the trained face detection model, and further a region enclosed by the regression box can be used as a corresponding face region.

**[0040]** In another embodiment, when the target attribute is a facial organ size in an image, the determining corresponding multiple attribute values can include the following: separately inputting the multiple face images into a trained facial organ detection model to obtain multiple facial organ regions corresponding to the multiple face images; and determining multiple organ region areas corresponding to the multiple facial organ regions as multiple attribute values corresponding to the facial organ size attribute. In a specific embodiment, facial organs can include eyes, a nose, a forehead, a mouth, ears, and the like. The above-mentioned facial organ size can be a size of a certain facial organ, or can be a total size of one or more facial organs. In a specific embodiment, the facial organ detection model can be implemented based on a CNN network or a DNN network.

**[0041]** In a specific embodiment, the facial organ detection model can be implemented based on a target detection algorithm. Specifically, a target detection model is trained by using facial organs as target objects, and is used as the facial organ detection model. As such, a regression box for selecting a facial organ position in any face image can be obtained by inputting the face image into the trained facial organ detection model, and further a region enclosed by the regression box can be used as a corresponding facial organ region. In an example, the above-mentioned facial organ size corresponds to a total size of multiple facial organs. Correspondingly, the facial organ detection model can include multiple sub-models, which are configured to correspondingly label multiple sub-regions of multiple facial organs in any face image, so as to determine a sum of areas of the multiple sub-regions as a facial organ size corresponding to the any face image.

**[0042]** In still another embodiment, when the target attribute is an ambient light intensity in an image, the determining corresponding multiple attribute values can include the following: separately inputting the multiple face images into a trained light intensity evaluation model to

obtain multiple light intensities. In a specific embodiment, the light intensity evaluation model can be implemented based on a CNN network or a DNN network.

[0043] In the description above, determining of multiple attribute values of the target attribute corresponding to the multiple face images is described by using examples in which the target attribute is the face size, the facial organ size, and the ambient light intensity in an image. Actually, the target attribute can alternatively be another image attribute, such as an image color temperature or image saturation.

[0044] Further, two face images whose degree of difference between attribute values corresponding to the target attribute satisfies the predetermined criterion are selected from the multiple face images based on the multiple determined attribute values. In an embodiment, a maximum value and a minimum value in the multiple attribute values can be determined, and further a face image corresponding to the maximum value and a face image corresponding to the minimum value are used as the two face images corresponding to the target attribute. In another embodiment, two face images whose degree of difference between corresponding attribute values is greater than a predetermined threshold (denoted as  $k$ ) can be selected from the multiple face images. In a specific embodiment, one attribute value (denoted as  $x_a$ ) can be randomly selected from the multiple attribute values, and then one attribute value  $x_b$  greater than  $x_a+k$  can be selected from the other attribute values. As such, a face image corresponding to the attribute value  $x_a$  and a face image corresponding to the attribute value  $x_b$  can be used as the two face images corresponding to the target attribute. In another specific embodiment, a difference between any two of the multiple attribute values can be first calculated to obtain multiple differences. Then, one difference can be randomly selected from one or more differences that are in the multiple differences and that are greater than the predetermined threshold, and further, two face images corresponding to the difference can be used as the two face images corresponding to the target attribute.

[0045] In the description above, the two face images corresponding to the target attribute can be determined based on the multiple attribute values of the target attribute corresponding to the multiple face images.

[0046] In another implementation, an attribute value (denoted as  $x_1$ ) of the target attribute corresponding to a certain face image is first determined based on the multiple face images, and attribute values of the other face images are then determined one by one, until a difference between the attribute value  $x_1$  and an attribute value  $x_j$  determined at the  $j$ th time is greater than the predetermined threshold. In this case, determining of an attribute value of the target attribute corresponding to a remaining face image is stopped, and the face image corresponding to the attribute value  $x_1$  and a face image corresponding to the attribute value  $x_j$  are used as the two

face images corresponding to the target attribute.

[0047] In the description above, two face images corresponding to the target attribute can be selected based on the multiple obtained face images to form a face image pair with a significant difference in case of the target attribute. Then, in step S230, both of the two face images are input into the pre-trained face liveness detection model corresponding to the target attribute to obtain the detection result corresponding to the target attribute.

[0048] In an embodiment, the face liveness detection model can be implemented based on a DNN network, a CNN network, a fast CNN, or a faster CNN. In an embodiment, the face liveness detection model can be implemented as a learning to compare model. In an embodiment, the face liveness detection model and the above-mentioned machine learning model used to determine attribute values of the target attribute (or referred to as an attribute value prediction model for short) can be separately trained. For example, the attribute value prediction model corresponding to the target attribute is first trained, and then the face liveness detection model corresponding to the target attribute is trained based on the trained attribute value prediction model and multiple obtained training image samples. In another embodiment, the face liveness detection model and the attribute value prediction model that correspond to the target attribute can be jointly trained.

[0049] It should be understood that, an obtained detection result corresponding to the target attribute indicates a face liveness detection result obtained based on the differential pictures in case of the target attribute. For example, if the detection succeeds, it can be determined that the photographed face images come from a genuine living user; or if the detection fails, it can be determined that the photographed face images do not come from a genuine living user, and actually may be rephotographed copies of a color-printed face image, etc.

[0050] In the description above, the detection result corresponding to the target attribute can be obtained. Further, in an implementation, if the one or more image attributes are one image attribute, a detection result corresponding to this image attribute can be directly used as a final detection result. In another implementation, if the one or more image attributes are multiple image attributes, correspondingly, the multiple image attributes can be separately used as target attributes to obtain multiple detection results corresponding respectively to the multiple image attributes. Further, after step S230, a final detection result for the first user can be determined based on the multiple detection results. In an embodiment, when a proportion of results indicating a detection success in the multiple detection results reaches a predetermined proportion (for example, 70% or 100%), the final detection result is determined as that the first user succeeds in the face liveness detection, or otherwise determined as that the first user fails the face liveness detection. In another embodiment, the detection results each include a

probability indicating a detection success. Based on this, in a specific embodiment, these probabilities can be averaged. When an average value is greater than a predetermined threshold (for example, 80% or 90%), the final detection result is determined as that the first user succeeds in the face liveness detection, or otherwise determined as that the first user fails the face liveness detection. In another specific embodiment, different weights can be assigned to the multiple image attributes, and then weighted summation can be performed on the multiple probabilities included in the multiple detection results. When a weighted probability is greater than a predetermined threshold (for example, 60% or 70%), the final detection result is determined as that the first user succeeds in the face liveness detection, or otherwise determined as that the first user fails the face liveness detection.

**[0051]** In the description above, the final face liveness detection result of the first user can be determined. According to some embodiments of another aspect, the above-mentioned method can be performed by a system that provides a face liveness detection service. In one execution case, the method can be performed by a server of the system; and in another execution case, the method can be jointly performed by a client and a server in the system. Specifically, the first user triggers a face liveness detection function by using the client of the system, and then the client processes multiple human images photographed by the user terminal, selects a face image pair corresponding to each of the one or more image attributes, and uploads one or more face image pairs corresponding to the one or more image attributes to the server. Afterwards, the server correspondingly processes the one or more face image pairs by using one or more face liveness detection models corresponding to the one or more image attributes to obtain one or more detection results, and further determines a final face liveness detection result for the first user based on the one or more detection results.

**[0052]** In conclusion, by using the face liveness detection method disclosed in the embodiments of this specification, a user does not have to perform a specific combination of actions, and face images actually used in algorithms can be effectively hidden to improve security of the detection algorithms. In addition, accuracy and availability of detection results can be effectively improved by performing face liveness detection on images with a relatively large attribute difference.

**[0053]** Corresponding to the above-mentioned detection method, some embodiments of this specification further disclose a detection apparatus. FIG. 3 is a schematic structural diagram of a face liveness detection apparatus, according to some embodiments. The apparatus can be implemented as any server, device cluster, or the like that has computing and processing capabilities. As shown in FIG. 3, the apparatus 300 includes the following units: an image acquisition unit 310, configured to obtain multiple face images continuously captured for a first

user; an image selection unit 320, configured to select two face images from the multiple face images for any target attribute in one or more predetermined image attributes, where a degree of difference between two attribute values of the target attribute corresponding to the two face images satisfies a predetermined criterion; and a liveness detection unit 330, configured to input both of the two face images into a pre-trained face liveness detection model corresponding to the target attribute to obtain a detection result corresponding to the target attribute.

[0054] In an embodiment, the image acquisition unit 310 is specifically configured to prompt, based on a user terminal of the first user, the first user to adjust a photographing method to improve face image quality; and during the prompting, continuously photograph the multiple face images by using the user terminal.

[0055] In a specific embodiment, that the first user is prompted to adjust a photographing method includes the following: prompting to adjust a position of a face relative to the user terminal; and/or prompting to adjust a light intensity of a photographing environment.

[0056] In an embodiment, the image selection unit 320 includes the following: an attribute value determining subunit 321, configured to determine multiple attribute values of the target attribute corresponding to the multiple face images; and an image selection subunit 322, configured to select, from the multiple face images based on the multiple attribute values, two face images whose degree of difference between corresponding attribute values satisfies the predetermined criterion.

[0057] In a specific embodiment, when the target attribute is a face size, the attribute value determining subunit 321 is specifically configured to separately input the multiple face images into a trained face detection model to obtain multiple face regions corresponding to the multiple face images; and determine multiple face region areas corresponding to the multiple face regions.

[0058] In another specific embodiment, when the target attribute is a facial organ size, the attribute value determining subunit 321 is specifically configured to separately input the multiple face images into a trained facial organ detection model to obtain multiple facial organ regions corresponding to the multiple face images; and determine multiple organ region areas corresponding to the multiple facial organ regions.

[0059] In still another specific embodiment, when the target attribute is a light intensity, the attribute value determining subunit 321 is specifically configured to separately input the multiple face images into a trained light intensity evaluation model to obtain multiple light intensities.

[0060] In another aspect, in a specific embodiment, the image selection subunit 322 is specifically configured to determine a maximum value and a minimum value in the multiple

attribute values to obtain the two corresponding face images.

[0061] In a specific embodiment, the image selection subunit 322 is specifically configured to select, from the multiple face images, two face images whose degree of difference between the corresponding attribute values is greater than a predetermined threshold.

5 [0062] In a specific embodiment, the one or more image attributes are multiple image attributes; and the apparatus 300 further includes the following: a final-result determining unit 340, configured to determine a final detection result for the first user based on multiple detection results corresponding respectively to the multiple image attributes.

10 [0063] In a more specific embodiment, the final-result determining unit 340 is specifically configured to, when the multiple detection results all indicate that the first user succeeds in face liveness detection, determine the final detection result as a face liveness detection success.

[0064] In conclusion, by using the face liveness detection apparatuses disclosed in the embodiments of this specification, a user does not have to perform a specific combination of actions, and face image-related user interaction information actually used in algorithms can be  
15 effectively hidden to improve security of the detection algorithms. In addition, accuracy and availability of detection results can be effectively improved by performing face liveness detection on images with a relatively large attribute difference.

[0065] According to some embodiments of another aspect, a computer-readable storage medium is further provided. The computer-readable storage medium stores a computer program,  
20 and when the computer program is executed in a computer, the computer is enabled to perform the method described with reference to FIG. 2.

[0066] According to some embodiments of still another aspect, a computing device is further provided, including a memory and a processor. The memory stores executable code, and when  
25 executing the executable code, the processor implements the method described with reference to FIG. 2.

[0067] A person skilled in the art should be aware that in the above-mentioned one or more examples, functions described in this specification can be implemented by hardware, software, firmware, or any combination thereof. When these functions are implemented by software, they  
30 can be stored in a computer-readable medium or transmitted as one or more instructions or code on the computer-readable medium.

[0068] Some specific implementations described above further describe the purposes, technical solutions, and beneficial effects of this specification. It should be understood that the above-mentioned descriptions are merely some specific implementations of this specification and are not intended to limit the protection scope of this specification. Any modification,  
35 equivalent replacement, or improvement made based on the technical solutions of this

specification shall fall within the protection scope of this specification.

## CLAIMS

What is claimed is:

1. A face liveness detection method, comprising:

obtaining multiple face images continuously captured for a first user;

5        selecting two face images from the multiple face images for any target attribute in one or more predetermined image attributes, wherein a degree of difference between two attribute values of the target attribute corresponding to the two face images satisfies a predetermined criterion; and

10        inputting both of the two face images into a pre-trained face liveness detection model corresponding to the target attribute to obtain a detection result corresponding to the target attribute.

2. The method according to claim 1, wherein the obtaining multiple face images continuously captured for a first user comprises the following:

15        prompting, based on a user terminal of the first user, the first user to adjust a photographing method to improve face image quality; and

      during the prompting, continuously photographing the multiple face images by using the user terminal.

3. The method according to claim 2, wherein the prompting to adjust a photographing method comprises the following:

20        prompting to adjust a position of a face relative to the user terminal; and/or

      prompting to adjust a light intensity of a photographing environment.

4. The method according to claim 1, wherein the selecting two face images from the multiple face images for any target attribute in one or more predetermined image attributes comprises the following:

25        determining multiple attribute values of the target attribute corresponding to the multiple face images; and

      selecting, from the multiple face images based on the multiple attribute values, two face images whose degree of difference between corresponding attribute values satisfies the predetermined criterion.

30        5. The method according to claim 4, wherein when the target attribute is a face size, the determining multiple attribute values of the target attribute corresponding to the multiple face images comprises the following:

      separately inputting the multiple face images into a trained face detection model to obtain multiple face regions corresponding to the multiple face images; and

determining multiple face region areas corresponding to the multiple face regions.

6. The method according to claim 4, wherein when the target attribute is a facial organ size, the determining multiple attribute values of the target attribute corresponding to the multiple face images comprises the following:

5 separately inputting the multiple face images into a trained facial organ detection model to obtain multiple facial organ regions corresponding to the multiple face images; and  
determining multiple organ region areas corresponding to the multiple facial organ regions.

7. The method according to claim 4, wherein when the target attribute is a light intensity, the determining multiple attribute values of the target attribute corresponding to the multiple  
10 face images comprises the following:

separately inputting the multiple face images into a trained light intensity evaluation model to obtain multiple light intensities.

8. The method according to claim 4, wherein the selecting, from the multiple face images based on the multiple attribute values, two face images whose degree of difference between  
15 corresponding attribute values satisfies the predetermined criterion comprises the following:

determining a maximum value and a minimum value in the multiple attribute values to obtain the two corresponding face images.

9. The method according to claim 4, wherein the selecting, from the multiple face images based on the multiple attribute values, two face images whose degree of difference between  
20 corresponding attribute values satisfies the predetermined criterion comprises the following:

selecting, from the multiple face images, two face images whose degree of difference between the corresponding attribute values is greater than a predetermined threshold.

10. The method according to claim 1, wherein the one or more image attributes are multiple image attributes; and after the obtaining a detection result corresponding to the target attribute,  
25 the method further comprises the following:

determining a final detection result for the first user based on multiple detection results corresponding respectively to the multiple image attributes.

11. The method according to claim 10, wherein the determining a final detection result for the first user based on multiple detection results corresponding respectively to the multiple  
30 image attributes comprises the following:

when the multiple detection results all indicate that the first user succeeds in face liveness detection, determining the final detection result as a face liveness detection success.

12. A face liveness detection apparatus, comprising:

an image acquisition unit, configured to obtain multiple face images continuously captured  
35 for a first user;

an image selection unit, configured to select two face images from the multiple face images for any target attribute in one or more predetermined image attributes, wherein a degree of difference between two attribute values of the target attribute corresponding to the two face images satisfies a predetermined criterion; and

5        a liveness detection unit, configured to input both of the two face images into a pre-trained face liveness detection model corresponding to the target attribute to obtain a detection result corresponding to the target attribute.

13. The apparatus according to claim 12, wherein the image acquisition unit is specifically configured to:

10        prompt, based on a user terminal of the first user, the first user to adjust a photographing method to improve face image quality; and

         during the prompting, continuously photograph the multiple face images by using the user terminal.

14. The apparatus according to claim 13, wherein the prompting to adjust a photographing method comprises the following:

         prompting to adjust a position of a face relative to the user terminal; and/or  
         prompting to adjust a light intensity of a photographing environment.

15. The apparatus according to claim 12, wherein the image selection unit comprises the following:

20        an attribute value determining subunit, configured to determine multiple attribute values of the target attribute corresponding to the multiple face images; and

         an image selection subunit, configured to select, from the multiple face images based on the multiple attribute values, two face images whose degree of difference between corresponding attribute values satisfies the predetermined criterion.

25        16. The apparatus according to claim 15, wherein when the target attribute is a face size, the attribute value determining subunit is specifically configured to:

         separately input the multiple face images into a trained face detection model to obtain multiple face regions corresponding to the multiple face images; and

         determine multiple face region areas corresponding to the multiple face regions.

30        17. The apparatus according to claim 15, wherein when the target attribute is a facial organ size, the attribute value determining subunit is specifically configured to:

         separately input the multiple face images into a trained facial organ detection model to obtain multiple facial organ regions corresponding to the multiple face images; and

         determine multiple organ region areas corresponding to the multiple facial organ regions.

35        18. The apparatus according to claim 15, wherein when the target attribute is a light

intensity, the attribute value determining subunit is specifically configured to:

separately input the multiple face images into a trained light intensity evaluation model to obtain multiple light intensities.

5 19. The apparatus according to claim 15, wherein the image selection subunit is specifically configured to:

determine a maximum value and a minimum value in the multiple attribute values to obtain the two corresponding face images.

20. The apparatus according to claim 15, wherein the image selection subunit is specifically configured to:

10 select, from the multiple face images, two face images whose degree of difference between the corresponding attribute values is greater than a predetermined threshold.

21. The apparatus according to claim 12, wherein the one or more image attributes are multiple image attributes; and the apparatus further comprises the following:

15 a final-result determining unit, configured to determine a final detection result for the first user based on multiple detection results corresponding respectively to the multiple image attributes.

22. The apparatus according to claim 21, wherein the final-result determining unit is specifically configured to:

20 when the multiple detection results all indicate that the first user succeeds in face liveness detection, determine the final detection result as a face liveness detection success.

23. A computer-readable storage medium, wherein the computer-readable storage medium stores a computer program, and when the computer program is executed in a computer, the computer is enabled to perform the method according to any one of claims 1 to 11.

25 24. A computing device, comprising a memory and a processor, wherein executable code is stored in the memory, and when the processor executes the executable code, the method according to any one of claims 1 to 11 is implemented.