**CLAIMS**

1. A face anti-spoofing method, the method comprising:

receiving a plurality of images of a user's face from a user device, each of the plurality of images corresponding to a facial expression performed by the user in response to each of a plurality of interaction challenges randomly selected from an interaction challenge library;

concatenating the plurality of images into a concatenated image;

using a convolutional neural network (CNN) classifier to obtain a confidence score for the concatenated image; and

determining liveness of the user's face based on a comparison of the confidence score and a threshold value.

2. The method according to claim 1, further comprising:

building up the interaction challenge library, wherein the interaction challenge library includes interaction challenges of eye blink, mouth open, mouth close, head pitch down, head pitch up, head roll left, head roll right, head yaw left, and/or head yaw right.

3. The method according to claim 1 or 2, further comprising:

randomly selecting three interaction challenges from the interaction challenge library to form the plurality of interaction challenges.

4. The method according to claim 3, further comprising:

training the CNN classifier with a training dataset having a size of N×9×H×W, wherein N is a number of images to be trained in one iteration, H is a height of each of the images, and W is a width of each of the images.

5. A face anti-spoofing method, the method comprising:

receiving a request for a plurality of interaction challenges from a server, wherein the plurality of interaction challenges are randomly selected from an interaction challenge library;

capturing a plurality of images of a user's face, each of the plurality of images corresponding to a facial expression performed by the user in response to each of the plurality of

interaction challenges, wherein each of the plurality of interaction challenges requires at least a start expression and an end expression of the face; and

transmitting the captured plurality of images to the server to obtain a confidence score so as to determine liveness of the face based on a comparison of the confidence score and a threshold value.

6. The method according to claim 5, wherein the interaction challenge library includes interaction challenges of eye blink, mouth open, mouth close, head pitch down, head pitch up, head roll left, head roll right, head yaw left, and/or head yaw right.

7. The method according to claim 5 or 6, wherein the plurality of interaction challenges include three interaction challenges randomly selected from the interaction challenge library.

8. The method according to any one of claims 5 to 7, wherein the facial expression is the end expression of the face.

9. A face anti-spoofing server, the server comprising:

at least one processor; and

a memory coupled to the at least one processor and storing programming instructions for execution by the at least one processor, the programming instructions instruct the server to:

receive a plurality of images of a user's face from a user device, each of the plurality of images corresponding to a facial expression performed by the user in response to each of a plurality of interaction challenges randomly selected from an interaction challenge library;

concatenate the plurality of images into a concatenated image;

use a convolutional neural network (CNN) classifier to obtain a confidence score for the concatenated image; and

determine liveness of the user's face based on a comparison of the confidence score and a threshold value.

10. The server according to claim 9, wherein the server is further instructed to:

build up the interaction challenge library, wherein the interaction challenge library includes interaction challenges of eye blink, mouth open, mouth close, head pitch down, head pitch up, head roll left, head roll right, head yaw left, and/or head yaw right.

11. The server according to claim 9 or 10, wherein the server is further instructed to:

randomly select three interaction challenges from the interaction challenge library to form the plurality of interaction challenges.

12. The server according to claim 11, wherein the server is further instructed to:

train the CNN classifier with a training dataset having a size of N×9×H×W, wherein N is a number of images to be trained in one iteration, H is a height of each of the images, and W is a width of each of the images.

13. A device for face anti-spoofing, the device comprising:

at least one processor; and

a memory coupled to the at least one processor and storing programming instructions for execution by the at least one processor, the programming instructions instruct the device to:

receive a request for a plurality of interaction challenges from a server, wherein the plurality of interaction challenges are randomly selected from an interaction challenge library;

capture a plurality of images of a user's face, each of the plurality of images corresponding to a facial expression performed by the user in response to each of the plurality of interaction challenges, wherein each of the plurality of interaction challenges requires at least a start expression and an end expression of the face; and

transmit the captured plurality of images to the server to obtain a confidence score so as to determine liveness of the face based on a comparison of the confidence score and a threshold value.

14. The device according to claim 13, wherein the interaction challenge library includes interaction challenges of eye blink, mouth open, mouth close, head pitch down, head pitch up, head roll left, head roll right, head yaw left, and/or head yaw right.

15. The device according to claim 13 or 14, wherein the plurality of interaction challenges include three interaction challenges randomly selected from the interaction challenge library.

16. The device according to any one of claims 13 to 15, wherein the facial expression is the end expression of the face.