

CLAIMS

1. A target object identification method, comprising:

performing classification on a to-be-identified target object in a target image to determine a prediction category of the to-be-identified target object;

5 determining whether the prediction category is correct according to a hidden layer feature for the to-be-identified target object; and

outputting prompt information in response to the prediction category being incorrect;

10 wherein determining whether the prediction category is correct according to the hidden layer feature of the to-be-identified target object comprises:

inputting the hidden layer feature for the to-be-identified target object into an authenticity identification model corresponding to the prediction category, such that the authenticity identification model outputs a probability value, wherein the authenticity identification model corresponding to the prediction category reflects
15 distribution of hidden layer features for target objects belonging to the prediction category, and the probability value represents a probability that a final category of the to-be-identified target object is the prediction category;

determining that the prediction category is incorrect when the probability value is less than a probability threshold; and

20 determining that the prediction category is correct when the probability value is greater than or equal to the probability threshold.

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

in response to the prediction category being correct, determining the prediction

category as a final category of the to-be-identified target object; and
outputting the final category of the to-be-identified target object.

3. The method according to claim 1 or 2, wherein the target image comprises multiple
5 stacked to-be-identified target objects;

performing classification on the to-be-identified target object in the target image
to determine the prediction category of the to-be-identified target object comprises:

adjusting a height of the target image to a preset height by scaling, wherein the
target image is obtained by cropping, according to a bounding box of the multiple
10 stacked to-be-identified target objects in an acquired image, the acquired image, and
a height direction of the target image is a stacking direction of the multiple stacked
to-be-identified target objects; and

performing classification on the to-be-identified target object in the adjusted
target image to determine the prediction category of the to-be-identified target object.
15

4. The method according to claim 3, wherein adjusting the height of the target image
to the preset height comprises:

scaling the height and a width of the target image in equal proportions, until the
width of the target image reaches a preset width; and
20 when the width of the scaled target image reaches the preset width, and the height
of the scaled target image is greater than the preset height, reducing the height and the
width of the scaled target image in equal proportions, until the height of the reduced
target image is equal to the preset height.

5. The method according to claim 3, wherein adjusting the height of the target image to the preset height comprises:
- scaling the height and a width of the target image in equal proportions, until the width of the target image reaches a preset width; and
 - 5 when the width of the scaled target image reaches the preset width, and the height of the scaled target image is less than the preset height, filling the scaled target image with a first pixel, such that the height of the filled target image is equal to the preset height.
- 10 6. The method according to claim 3, wherein the preset height is set according to a maximum number of to-be-identified target objects;
- performing classification on the to-be-identified target object in the adjusted target image to determine the prediction category of the to-be-identified target object comprises:
 - 15 performing feature extraction on the adjusted target image to obtain a feature map, wherein a height dimension of the feature map corresponds to the height direction of the target image;
 - performing average pooling on the feature map in a width dimension of the feature map to obtain a pooled feature map;
 - 20 segmenting, according to the maximum number, the pooled feature map in the height dimension to obtain a preset number of features, wherein the preset number of features and the multiple stacked to-be-identified target objects have a one-to-one correspondence; and
 - determining the prediction category of each of the multiple stacked to-be-

identified target objects according to each of the features.

7. The method according to claim 6, wherein performing classification on the to-be-identified target object in the adjusted target image to determine the prediction
5 category of the to-be-identified target object is executed by a neural network which comprises a classification network; wherein the classification network comprises K classifiers, K is the number of known categories when classifying, and K is a positive integer; and

determining the prediction category of each of the multiple stacked to-be-
10 identified target objects according to each of the features comprises:

respectively calculating cosine similarities between each of the features and
a weight vector of each of the K classifier; and

determining the prediction category of each of the multiple stacked to-be-
identified target objects according to the calculated cosine similarities.

15

8. The method according to claim 6, wherein performing classification on the to-be-identified target object in the adjusted target image to determine the prediction
category of the to-be-identified target object is executed by a neural network which
comprises a feature extraction network, wherein the feature extraction network
20 comprises multiple convolutional layers, respective stride of the last N convolutional
layers of the multiple convolutional layers in the feature extraction network is 1 in the
height dimension of the feature map, and N is a positive integer.

9. The method according to claim 1, wherein performing classification on the to-be-

identified target object in the target image is executed by a neural network; the authenticity identification model corresponding to the prediction category is created by using hidden layer features for authenticated target objects belonging to the prediction category; and the authenticated target objects are correctly predicted in a
5 training stage and/or test stage of the neural network.

10. A target object identification apparatus, comprising:

a classification unit, configured to perform classification on a to-be-identified target object in a target image to determine a prediction category of the to-be-
10 identified target object;

a determination unit, configured to determine whether the prediction category is correct according to a hidden layer feature for the to-be-identified target object; and

a prompt unit, configured to output prompt information in response to the prediction category being incorrect;

15 wherein the determination unit is configured to:

input the hidden layer feature for the to-be-identified target object into an authenticity identification model corresponding to the prediction category, such that the authenticity identification model outputs a probability value, wherein the authenticity identification model corresponding to the prediction category reflects
20 distribution of hidden layer features for target objects belonging to the prediction category, and the probability value represents a probability that a final category of the to-be-identified target object is the prediction category;

determine that the prediction category is incorrect when the probability value is less than a probability threshold; and

determine that the prediction category is correct when the probability value is greater than or equal to the probability threshold.

11. The apparatus according to claim 10, further comprising: an output unit configured
5 to:

in response to the prediction category being correct, determine the prediction category as a final category of the to-be-identified target object; and
output the final category of the to-be-identified target object.

10 12. The apparatus according to claim 10 or 11, wherein the target image comprises multiple stacked to-be-identified target objects;

the classification unit is configured to:

adjust a height of the target image to a preset height by scaling, wherein the target image is obtained by cropping, according to a bounding box of the multiple
15 stacked to-be-identified target objects in an acquired image, the acquired image, and a height direction of the target image is a stacking direction of the multiple stacked to-be-identified target objects; and

perform classification on the to-be-identified target object in the adjusted target image to determine the prediction category of the to-be-identified target object.

20

13. The apparatus according to claim 12, wherein the classification unit is configured to:

scale the height and a width of the target image in equal proportions, until the width of the target image reaches a preset width; and

when the width of the scaled target image reaches the preset width, and the height of the scaled target image is greater than the preset height, reduce the height and the width of the scaled target image in equal proportions, until the height of the reduced target image is equal to the preset height.

5

14. The apparatus according to claim 12, wherein the classification unit is configured to:

scale the height and a width of the target image in equal proportions, until the width of the target image reaches a preset width; and

10

when the width of the scaled target image reaches the preset width, and the height of the scaled target image is less than the preset height, fill the scaled target image with a first pixel, such that the height of the filled target image is equal to the preset height.

15

15. The apparatus according to claim 12, wherein the preset height is set according to a maximum number of to-be-identified target objects;
the classification unit is configured to:

perform feature extraction on the adjusted target image to obtain a feature map, wherein a height dimension of the feature map corresponds to the height direction of
the target image;

20

perform average pooling on the feature map in a width dimension of the feature map to obtain a pooled feature map;

segment, according to the maximum number, the pooled feature map in the height dimension to obtain a preset number of features, wherein the preset number of features

and the multiple stacked to-be-identified target objects have a one-to-one correspondence; and

determine the prediction category of each of the multiple stacked to-be-identified target objects according to each of the features.

5

16. The apparatus according to claim 15, wherein performing classification on the to-be-identified target object in the adjusted target image to determine the prediction category of the to-be-identified target object is executed by a neural network which comprises a classification network; wherein the classification network comprises K
10 classifiers, K is the number of known categories when classifying, and K is a positive integer; and

determining the prediction category of each of the multiple stacked to-be-identified target objects according to each of the features comprises:

respectively calculating cosine similarities between each of the features and
15 a weight vector of each of the K classifier; and

determining the prediction category of each of the multiple stacked to-be-identified target objects according to the calculated cosine similarities.

17. An electronic device, comprising:

20 a processor; and

a memory configured to store processor-executable instructions;

wherein the processor is configured to invoke the processor-executable instructions stored in the memory to implement the method according to any one of claims 1 to 9.

18. A computer-readable storage medium, having computer program instructions stored thereon, wherein when the computer program instructions are executed by a processor, the method according to any one of claims 1 to 9 is implemented.

5