

CLAIMS

1. An image processing method, comprising:

acquiring a first image and a first attention map of a first target object in the first image, wherein the first attention map indicates the position of the first target object in the first image;

acquiring a second image and a second mask image of the second image, wherein pixel values of pixels in the second mask image are equal; and

obtaining, according to first input information and second input information, a first prediction result that indicates the position of the first target object in the second image, wherein the first input information is determined according to the first image and the first attention map, and the second input information is determined according to the second image and the second mask image,

wherein the obtaining, according to first input information and second input information, a first prediction result that indicates the position of the first target object in the second image comprises:

performing feature extraction on the first input information to obtain a feature map corresponding to the first input information;

performing feature extraction on the second input information to obtain a feature map corresponding to the second input information; and

obtaining, according to the feature map corresponding to the first input information and the feature map corresponding to the second input information, the first prediction result that indicates the position of the first target object in the second image, and

wherein the obtaining, according to the feature map corresponding to the first input information and the feature map corresponding to the second input information, the first prediction result that indicates the position of the first target object in the second image comprises:

performing feature association on the feature map corresponding to the first input

information and the feature map corresponding to the second input information to obtain a first associated feature map; and

obtaining, based on the first associated feature map, the first prediction result that indicates the position of the first target object in the second image.

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

acquiring a second attention map of the second image according to a detection result for a second target object in the second image, wherein the second attention map indicates the position of the second target object in the second image;

after obtaining the first prediction result, the method further comprises:

obtaining, according to third input information and fourth input information, a first score of association that represents the degree of association between the first target object and the second target object, wherein the third input information is determined according to the second image and the second attention map, and the fourth input information is determined according to the second image and the first prediction result; and

determining a result of association between the first target object and the second target object according to the first score of association.

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

obtaining, according to the first input information and the third input information, a second score of association that represents the degree of association between the first target object and the second target object; and

the determining a result of association between the first target object and the second target object according to the first score of association comprises:

determining the result of association between the first target object and the second target object according to the first score of association and the second score of association.

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

acquiring a second prediction result that indicates the position of the second target object in the first image;

obtaining, according to fifth input information and the first input information, a third score of association that represents the degree of association between the first target object and the second target object, wherein the fifth input information is determined according to the first image and the second prediction result; and

the determining a result of association between the first target object and the second target object according to the first score of association comprises:

determining the result of association between the first target object and the second target object according to the first score of association and the third score of association.

5. The method according to claim 4, wherein the determining the result of association between the first target object and the second target object according to the first score of association and the third score of association comprises:

determining a final score of association between the first target object and the second target object according to the first score of association and the third score of association;

in a case that the final score of association is greater than a score threshold, determining that the result of association indicates the first target object being potentially associated with the second target object;

in a case that the final score of association is less than or equal to the score threshold, determining that the result of association indicates the first target object being not associated with the second target object.

6. The method according to claim 4, wherein the determining the result of association between the first target object and the second target object according to the first score of association and the third score of association comprises:

determining a final score of association between the first target object in the first image and the second target object according to the first score of association and the third score of association; and

matching, according to the final score of association between each first target object in the first image and each second target object in the second image, the each first target object with the each second target object to obtain the result of association between the each

first target object and the each second target object.

7. The method according to claim 6, wherein the matching, according to the final score of of association between each first target object in the first image and each second target object in the second image, the each first target object with the each second target object to obtain the result of association between the each first target object and the each second target object comprises:

determining a sum of scores of association between the each first target object and the each second target object according to the final score of association between the each first target object in the first image and the each second target object in the second image;

in a case that the sum of scores of association is the maximum and greater than a predetermined sum-of-scores threshold, determining that the each first target object in the first image matches the each second target object in the second image respectively, and obtaining the result of association between the each first target object and the each second target object;

in a case that the maximum sum of scores of association in the sums of scores of association is less than or equal to the sum-of-scores threshold, determining that the each first target object in the first image does not match the each second target object in the second image.

8. An image processing apparatus, comprising:

a first acquisition module, configured to acquire a first image and a first attention map of a first target object in the first image, wherein the first attention map indicates the position of the first target object in the first image;

a second acquisition module, configured to acquire a second image and a second mask image of the second image, wherein pixel values of pixels in the second mask image are equal; and

a first determination module, configured to obtain, according to first input information and second input information, a first prediction result that indicates the position of the first target object in the second image, wherein the first input information is determined according to the first image and the first attention map, and the second input information is determined according to the second image and the second mask image,

wherein the first determination module comprises:

a first feature extraction sub-module, configured to perform feature extraction on the first

input information to obtain a feature map corresponding to the first input information;

a second feature extraction sub-module, configured to perform feature extraction on the second input information to obtain a feature map corresponding to the second input information; and

a prediction result determination sub-module, configured to obtain, according to the feature map corresponding to the first input information and the feature map corresponding to the second input information, the first prediction result that indicates the position of the first target object in the second image, and

wherein the prediction result determination sub-module comprises:

a feature association unit, configured to perform feature association on the feature map corresponding to the first input information and the feature map corresponding to the second input information to obtain a first associated feature map; and

a prediction result determination unit, configured to obtain, based on the first associated feature map, the first prediction result that indicates the position of the first target object in the second image.

9. The apparatus according to claim 8, further comprising:

a third acquisition module, configured to acquire a second attention map of the second image according to a detection result for a second target object in the second image, wherein the second attention map indicates the position of the second target object in the second image;

a second determination module, configured to obtain, according to third input information and fourth input information, a first score of association that represents the degree of association between the first target object and the second target object, wherein the third input information is determined according to the second image and the second attention map, and the fourth input information is determined according to the second image and the first prediction result; and

a third determination module, configured to determine a result of association between the first target object and the second target object according to the first score of association.

10. The apparatus according to claim 9, further comprising:

a fourth determination module, configured to obtain, according to the first input information and the third input information, a second score of association that represents the degree of association between the first target object and the second target object,

wherein the third determination module is configured to determine the result of association between the first target object and the second target object according to the first score of association and the second score of association.

11. The apparatus according to claim 9 or 10, further comprising:

a fourth acquisition module, configured to acquire a second prediction result that indicates the position of the second target object in the first image; and

a fifth determination module, configured to obtain, according to fifth input information and the first input information, a third score of association that represents the degree of association between the first target object and the second target object, wherein the fifth input information is determined according to the first image and the second prediction result,

wherein the third determination module is configured to determine the result of association between the first target object and the second target object according to the first score of association and the third score of association.

12. The apparatus according to claim 11, wherein the third determination module comprises:

a first determination sub-module, configured to determine a final score of association between the first target object and the second target object according to the first score of association and the third score of association;

a second determination sub-module, configured to, in a case that the final score of association is greater than a score threshold, determine that the result of association indicates the first target object being potentially associated with the second target object; and

a third determination sub-module, configured to, in a case that the final score of association is less than or equal to the score threshold, determine that the result of association indicates the first target object being not associated with the second target object.

13. The apparatus according to claim 11, wherein the third determination module comprises:

a fourth determination sub-module, configured to determine a final score of association between the first target object in the first image and the second target object according to the first score of association and the third score of association; and

a matching sub-module, configured to match, according to the final score of association between each first target object in the first image and each second target object in the second image, the each first target object with the each second target object to obtain the result of association between the each first target object and the each second target object.

14. The apparatus according to claim 13, wherein the matching sub-module comprises:

a first determination unit, configured to determine a sum of scores of association between the each first target object and the each second target object according to the final score of association between the each first target object in the first image and the each second target object in the second image;

a second determination unit, configured to, in a case that the sum of scores of association is the maximum and greater than a predetermined sum-of-scores threshold, determine that the each first target object in the first image matches the each second target object in the second image respectively, and obtain the result of association between the each first target object and the each second target object; and

a third determination unit, configured to, in a case that the maximum sum of scores of association in the sums of scores of association is less than or equal to the sum-of-scores threshold, determine that the each first target object in the first image does not match the each second target object in the second image.

15. An electronic device, comprising:

a processor; and

a memory configured to store processor-executable instructions,

wherein the processor is configured to invoke the instructions stored in the memory, so as to execute the method according to any one of claims 1 to 7.

16. 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 7 is implemented.