

WHAT IS CLAIMED IS:

1. A method for detecting correlated objects involved in an image, comprising:
 - detecting a face object, a hand object and a preset body part object involved in a
 - 5 target image, wherein the preset body part object represents a preset connection part
 - between a face and a hand;
 - respectively predicting correlation between the detected face object and the
 - detected preset body part object, and correlation between the detected preset body part
 - object and the detected hand object, to obtain a first correlation prediction result
 - 10 between the face object and the preset body part object, and a second correlation
 - prediction result between the preset body part object and the hand object;
 - predicting correlation between the detected face object and the detected hand
 - object to obtain a third correlation prediction result,
 - adjusting the third correlation prediction result based on the first correlation
 - 15 prediction result and the second correlation prediction result; and
 - determining correlated objects involved in the target image based on the adjusted
 - third correlation prediction result.
2. The method of claim 1, wherein
 - 20 the target image involves a plurality of face objects and a plurality of hand objects;
 - and
 - predicting correlation between the detected face object and the detected hand
 - object to obtain a third correlation prediction result comprises:
 - combining each of the detected face objects with each of the detected hand
 - 25 objects to form a plurality of first combinations;

obtaining a third correlation prediction result between the face object and the hand object in each of the first combinations by, for each of the first combinations, predicting correlation between the face object and the hand object in the first combination based on visual features and position features of the face object and the hand object in the first combination.

3. The method of claim 1 or 2, wherein the target image involves a plurality of preset body part objects; and respectively predicting correlation between the detected face object and the detected preset body part object, and correlation between the preset body part object and the hand object, to obtain a first correlation prediction result between the face object and the preset body part object, and a second correlation prediction result between the preset body part object and the hand object comprises:

combining each of the detected face objects with each of the preset body part objects to form a plurality of second combinations;

obtaining a first correlation prediction result between the face object and the preset body part object in each of the second combinations by, for each of the second combinations, predicting correlation on the face object and the preset body part object in the second combination based on visual features and position features of the face object and the preset body part object in the second combination;

combining each of the preset body part objects with each of the hand objects to form a plurality of third combinations; and

obtaining a second correlation prediction result between the preset body part object and the hand object in each of the third combinations by, for each of the third combinations, predicting correlation between the preset body part object and the hand

object in the third combination based on visual features and position features of the preset body part object and the hand object in the third combination.

4. The method of claim 1, wherein adjusting the third correlation prediction result
5 based on the first correlation prediction result and the second correlation prediction result comprises:

determining a target face object of which a first correlation prediction score in the first correlation prediction result with respect to the preset body part object is highest;

determining a target hand object of which a second correlation prediction score in
10 the second correlation predication result with respect to the preset body part object is highest; and

based on the first correlation prediction score of the target face object with respect to the preset body part object, and the second correlation prediction score of the target hand object with respect to the preset body part object, adjusting a third
15 correlation prediction score in the third correlation prediction result between the target face object and the target hand object.

5. The method of claim 4, wherein

determining a target face object of which a first correlation prediction score in the
20 first correlation prediction result with respect to the preset body part object is highest comprises:

determining candidate face objects, wherein each of the candidate face objects has a first correlation prediction score with respect to the target body part object greater than a preset threshold; and

25 selecting one from the candidate face objects of which the first correlation

prediction score with respect to the preset body part object is highest as the target face object; and/or

determining a target hand object of which a second correlation prediction score in the second correlation predication result with respect to the preset body part object is
5 highest comprises:

determining candidate hand objects, wherein each of the candidate hand objects has a second correlation prediction score with respect to the target body part object greater than a preset threshold; and

selecting one from the candidate hand objects of which the second
10 correlation prediction score with respect to the preset body part object is highest as the target hand object.

6. The method of claim 4 or 5, wherein based on the first correlation prediction score of the target face object with respect to the preset body part object, and the
15 second correlation prediction score of the target hand object with respect to the preset body part object, adjusting the third correlation prediction score in the third correlation prediction result between the target face object and the target hand object comprises:

determining an average value of the first correlation prediction score of the target
20 face object with respect to the preset body part object, and the second correlation prediction score of the target hand object with respect to the preset body part object; and

obtaining the adjusted third correlation prediction score by adding the average value to the third correlation prediction score between the target face object and the
25 target hand object.

7. The method of claim 6, wherein determining correlated objects involved in the target image based on the adjusted third correlation prediction result comprises:

selecting each from a plurality of the third correlation prediction scores in an
5 order of the third correlation prediction scores from high to low, and

for a current combination of the face object and the hand object corresponding to the selected third correlation prediction score:

based on determined correlated objects involved in the target image,
determining a number of hand objects that are correlated with the face object in the
10 current combination as a first number, and determining a number of face objects that
are correlated with the hand object in the current combination as a second number;
and

in response to that the first number is lower than a first preset threshold, and
the second number is lower than a second preset threshold, determining the face
15 object and the hand object in the current combination as correlated objects involved in
the target image.

8. The method of any one of claims 1-7, wherein determining correlated objects
involved in the target image based on the first correlation prediction result and the
20 second correlation prediction result comprises:

based on the first correlation prediction result and the second correlation
prediction result, determining a face object and a hand object of which correlations
with respect to a same preset body part object satisfying a preset condition as
correlated objects involved in the target image.

25

9. The method of any one of claims 1-8, further comprising:

outputting a detection result of the correlated objects involved in the target image.

10. The method of any one of claims 1-9, wherein the preset body part comprises
5 at least one of a shoulder object, an elbow object and a wrist object.

11. The method of any one of claims 1-10, wherein

the face object, the hand object, and the preset body part object involved in the
target image are detected from the target image by a target object detecting network;

10 the third correlation prediction result is detected by a first preset network
comprising a face-hand correlation detecting model;

the first correlation prediction result and the second correlation prediction result
are detected by a second preset network comprising a face-preset-body-part
correlation detecting model and a preset-body-part-hand correlation detecting model;

15 and

the target object detecting network, the face-hand correlation detecting model, the
face-preset-body-part correlation detecting model and the preset-body-part-hand
correlation detecting model are trained by:

training the target object detecting network based on a first training sample
20 set which comprises a plurality of training samples with respective first label
information, wherein the first label information contains respective position label
information of face objects, hand objects and preset body part objects; and

jointly training the target object detecting network, the face-hand correlation
detecting model, the face-preset-body-part correlation detecting model, and the
25 preset-body-part-hand correlation detecting model based on a second training sample

set which comprises a plurality of training samples with respective second label information, wherein the second label information contains respective position label information of face objects, hand objects and preset body part objects, and respective label information on correlations between face objects, preset body part objects and
5 hand objects.

12. An apparatus for detecting correlated objects involved in an image, comprising:

a detector configured to detect a face object, a hand object and a preset body part
10 object involved in a target image, wherein the preset body part object represents a preset connection part between a face and a hand;

a first correlation predicting unit configured to respectively predict correlation between the detected face object and the detected preset body part object, and correlation between the detected preset body part object and the detected hand object,
15 to obtain a first correlation prediction result between the face object and the preset body part object, and a second correlation prediction result between the preset body part object and the hand object; and

a second correlation predicting unit configured to predict correlation between the detected face object and the detected hand object to obtain a third correlation
20 prediction result;

a determining unit configured to, determine correlated objects involved in the target image based on the first correlation prediction result and the second correlation prediction result;

a determining unit includes: an adjustment sub-unit configured to adjust the third
25 correlation prediction result based on the first correlation prediction result and the

second correlation prediction result; a determining sub-unit configured to determine correlated objects involved in the target image based on the adjusted third correlation prediction result.

- 5 13. An electronic device comprising:
- a processor;
- a memory for storing executable instructions of the processor;
- wherein the processor is configured to invoke the executable instructions stored
- in the memory to implement the method for detecting correlated objects involved in
- 10 an image according to any one of claims 1 to 11.

14. A non-transitory computer-readable storage medium, wherein the storage
- medium stores a computer program, and the computer program is configured to
- perform the method for detecting correlated objects involved in an image according to
- 15 any one of claims 1 to 11.