WHAT IS CLAIMED IS:

1. A method for predicting correlation between objects, comprising:

detecting a first object, a second object, and a third object involved in a target image, wherein the first object and the second object represent different body parts, and the third object indicates a body object, wherein a result of the detecting comprises a bounding box;

determining a joint bounding box surrounding the first object, the second object, and the third object; and

predicting correlation between the first object and the second object based on a region corresponding to the joint bounding box in the target image;

wherein determining the joint bounding box surrounding the first object, the second object, and the third object comprises:

determining a first surrounding box surrounding the first object and the second object based on respective detection results of the first object and the second object, wherein the first surrounding box indicates one of bounding boxes surrounding the first object and the second object of which area is smallest; and

generating the joint bounding box surrounding the first object, the second object, and the third object based on the first surrounding box and a detection result of the third object, wherein the joint bounding box indicates an externally-connected joint bounding box which is externally connected with the first surrounding box.

2. The method of claim 1, wherein

determining the first surrounding box surrounding the first object and the second

object based on respective detection results of the first object and the second object comprises:

obtaining one or more object pairs by combining each of one or more detected first objects with each of one or more detected second objects; and

for each of the object pairs, determining a first surrounding box surrounding the object pair; and/or,

for each of the object pairs, generating the joint bounding box surrounding the first object, the second object, and the third object based on the first surrounding box and a detection result of the third object comprises:

obtain one or more ternary object sets by combining the object pair with each of one or more detected third objects, wherein the object pairs are obtained by combining each detected first object with each detected second object; and

for each of the ternary object sets, generating a joint bounding box surrounding the ternary object set based on the first surrounding box of the object pair and a detection result of the third object in the ternary object set.

3. The method of claim 1, wherein

generating the joint bounding box surrounding the first object, the second object, and the third object based on the first surrounding box and a detection result of the third object comprises:

selecting one from a bounding box of a respective detected third object, as a target bounding box corresponding to the first surrounding box, in a way that an overlapping area between the selected bounding box and the first surrounding box is largest; and

generating the joint bounding box surrounding the first object, the second object, and the third object based on the first surrounding box and the target bounding box.

4. The method of claim 3, wherein

the joint bounding box indicates an externally-connected joint bounding box which is externally connected with the bounding box of the third object; or

the joint bounding box indicates an externally-connected joint bounding box which is externally connected with the first surrounding box and the bounding box of the third object.

5. The method of any one of claims 1-4, wherein predicting correlation between the first object and the second object based on the region corresponding to the joint bounding box in the target image comprises:

obtaining region features of the joint bounding box and respective region features of the first object and the second object surrounded by the joint bounding box by inputting the joint bounding box and respective bounding boxes of the first object and the second object surrounded by the joint bounding box into a region feature extraction unit;

obtaining a spliced feature of a preset length by splicing the region features of the joint bounding box, the respective region features of the first object and the second object surrounded by the joint bounding box, and respective position features of the first object and the second object surrounded by the joint bounding box; and

predicting correlation between the first object and the second object based on the

spliced feature.

6. The method of any one of claims 1-5, further comprising:

determining correlated objects involved in the target image based on a prediction result of the correlation between the first object and the second object.

7. The method of claim 6, wherein

the prediction result of the correlation comprises a correlation prediction score,

determining correlated objects involved in the target image based on the prediction result of the correlation between the first object and the second object comprises:

selecting each from a plurality of object pairs, as a current object pair, in an order of respective correlation prediction scores of the object pairs from high to low, wherein the plurality of object pairs are obtained by combining each detected first object with each detected second object; and

for the current object pair,

based on determined correlated objects involved in the target image, determining a number of second objects that are correlated with the first object in the current object pair as a first number, and determining a number of first objects that are correlated with the second object in the current object pair 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 first object and the second object in the current object pair as correlated objects involved in the target image.

8. The method of any one of claims 5-7, further comprising:

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

9. The method of any one of claims 1-8, wherein the correlation between the first object and the second object comprises a confidence regarding that the first object and the second object belong to one same body object.

10. The method of any one of claims 1-9, wherein the first object comprises a face object; and the second object comprises a hand object.

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

detecting a first object, a second object, and a third object in a target image comprises:

detecting the first object, the second object, and the third object involved in the target image with a target object detecting model;

predicting the correlation between the first object and the second object based on a region corresponding to the joint bounding box in the target image comprises:

predicting the correlation between the first object and the second object based on the region corresponding to the joint bounding box in the target image with a correlation predicting model;

wherein the target object detecting model and the correlation predicting model are trained by:

training the target object detecting model based on a first training sample set

which comprises a plurality of training samples with respective first label information, wherein the first label information comprises respective bounding boxes of a first object, a second object, and a third object; and

jointly training the target object detecting model and the correlation predicting 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 comprises respective bounding boxes of a first object, a second object, and a third object, and respective label information on correlations between first objects and second objects.

12. An apparatus for predicting correlation between objects, comprising:

an object detecting unit configured to detect a first object, a second object, and a third object involved in a target image, wherein the first object and the second object represent different body parts, and the third object indicates a body object, wherein a result of the detecting comprises a bounding box;

a determining unit configured to determine a joint bounding box surrounding the first object, the second object, and the third object; and

a correlation predicting unit configured to predict correlation between the first object and the second object based on a region corresponding to the joint bounding box in the target image;

wherein the determining unit comprises:

a determining sub-unit configured to determine a first surrounding box surrounding the first object and the second object based on respective detection results of the first object and the second object, wherein the first surrounding box indicates

one of bounding boxes surrounding the first object and the second object of which area is smallest;

a generating unit configured to generate the joint bounding box surrounding the first object, the second object, and the third object based on the first surrounding box and a detection result of the third object, wherein the joint bounding box indicates an externally-connected joint bounding box which is externally connected with the first surrounding box.

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 predicting correlation between objects 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 predicting correlation between objects according to any one of claims 1 to 11.