WHAT IS CLAIMED IS:

1. A method of detecting associated objects, comprising:

detecting at least one matching object group from an image to be detected, wherein each of the at least one matching object group comprises at least two target objects;

acquiring visual information of each of the at least two target objects in each matching object group, and spatial information of the at least two target objects in each matching object group; and

determining whether the at least two target objects in each matching object group are associated, according to the visual information and the spatial information of the at least two target objects in each matching object group,

wherein

the detecting the at least one matching object group from the image to be detected comprises:

detecting each target object and an object category of each target object from the image to be detected; and

combining each target object of each object category with each target object of other object categories to obtain the at least one matching object group, and

the acquiring spatial information of the at least two target objects in each matching object group comprises:

detecting a detection box for each target object from the image to be detected;

generating an auxiliary bounding box for each matching object group,

wherein the auxiliary bounding box covers the detection box for each target object in each matching object group;

determining position feature information of each target object in each matching object group, according to the auxiliary bounding box and the detection box for each target object; and

fusing the position feature information of each target object in a same matching object group to obtain the spatial information of the at least two target objects in the matching object group.

2. The method of claim 1, wherein the acquiring visual information of each of the at least two target objects in each matching object group comprises:

performing visual feature extraction on each target object in each matching object group to obtain the visual information of each target object.

3. The method of claim 1, wherein:

the auxiliary bounding box is one of bounding boxes covering each target object in the matching object group which has a minimum area.

4. The method of any one of claims 1 to 3, wherein the determining whether the at least two target objects in each matching object group are associated comprises:

performing fusion processing on the visual information and the spatial information of the at least two target objects in each matching object group to obtain a fusion feature of each matching object group; and

performing association classification processing on the fusion feature of each

matching object group to determine whether the at least two target objects in each matching object group are associated.

5. The method of claim 4, wherein the performing association classification processing on the fusion feature of each matching object group to determine whether the at least two target objects in each matching object group are associated, comprises:

performing the association classification processing on the fusion feature of each matching object group to obtain an association score between the at least two target objects in each matching object group;

determining the matching object group with a highest association score as a target matching object group, for a plurality of matching object groups to which a same target object belongs; and

determining that the at least two target objects in the target matching object group are associated target objects.

6. The method of claim 1, wherein in the case that the target objects are human body parts, the determining whether the at least two target objects in each matching object group are associated comprises:

determining whether respective human body parts in a same matching object group belong to a same human body.

7. The method of claim 1, further comprising:

acquiring a sample image set, wherein the sample image set comprises at least one sample image, each of the at least one sample image comprises at least one

sample matching object group and label information corresponding to the sample matching object group, each of the at least one sample matching object group comprises at least two sample target objects, the label information represents association results for respective sample target objects in the sample matching object group;

processing the sample image through an association detection network to be trained, to detect the sample matching object group from the sample image;

processing the sample image through an object detection network to be trained, to obtain visual information of each of the at least two sample target objects in each sample matching object group, and processing the sample image through the association detection network to be trained to obtain spatial information of the at least two sample target objects in each sample matching object group;

obtaining an association detection result for each sample matching object group through the association detection network to be trained according to the visual information and the spatial information of the at least two sample target objects in each sample matching object group; and

determining an error between the association detection result for each sample matching object group and respective label information, and adjusting a network parameter of at least one of the association detection network and the object detection network according to the error until the error converges.

8. An apparatus for detecting associated objects, comprising:

a detection module, configured to detect at least one matching object group from an image to be detected, wherein each of the at least one matching object group

comprises at least two target objects;

an acquisition module, configured to acquire visual information of each of the at least two target objects in each matching object group, and spatial information of the at least two target objects in each matching object group; and

a determination module, configured to determine whether the at least two target objects in each matching object group are associated, according to the visual information and the spatial information of the at least two target objects in each matching object group,

wherein

the detection module comprises:

a detection submodule, configured to detect each target object and an object category of each target object from the image to be detected; and

a combination submodule, configured to combine each target object of each object category with each target object of other object categories to obtain the at least one matching object group, and

the acquisition module is further configured to:

detect a detection box for each target object from the image to be detected;

generate an auxiliary bounding box for each matching object group, wherein the auxiliary bounding box covers the detection box for each target object in each matching object group;

determine position feature information of each target object in each matching object group, according to the auxiliary bounding box and the detection box for each target object; and

fuse the position feature information of each target object in a same matching

object group to obtain the spatial information of the at least two target objects in the matching object group.

9. The apparatus of claim 8, wherein the acquisition module is further configured to:

perform visual feature extraction on each target object in each matching object group to obtain the visual information of each target object.

10. The apparatus of any one of claims 8 to 9, wherein the determination module comprises:

a fusion submodule, configured to perform fusion processing on the visual information and the spatial information of the at least two target objects in each matching object group to obtain a fusion feature of each matching object group; and

a determination submodule, configured to perform association classification processing on the fusion feature of each matching object group to determine whether the at least two target objects in each matching object group are associated.

11. The apparatus of claim 10, wherein the determination submodule is further configured to:

perform the association classification processing on the fusion feature of each matching object group to obtain an association score between the at least two target objects in each matching object group;

determine the matching object group with a highest association score as a target matching object group, for a plurality of matching object groups to which a same target object belongs; and

determine that the at least two target objects in the target matching object group are associated target objects.

12. The apparatus of claim 8, wherein in the case that the target objects are human body parts, the determination module is further configured to:

determine whether respective human body parts in a same matching object group belong to a same human body.

13. An electronic device, comprising:

a processor; and

a memory communicatively connected with the processor and storing computer instructions readable by the processor, wherein the computer instructions, when read by the processor, cause the processor to perform the method of any one of claims 1 to 7.

14. A storage medium, storing computer-readable instructions for causing a computer to perform the method of any one of claims 1 to 7.