

CLAIMS

1. A network training method, comprising:

inputting an unannotated first sample image into a target detection network for processing to obtain a target detection result of the first sample image, the target detection result comprising an image area, feature information and class probability of a target in the first sample image;

determining category confidence of the target according to the class probability of the target;

for a first target with the category confidence greater than or equal to a first threshold value in the target, taking the first sample image where the first target is located as an annotated second sample image, and adding the second sample image to a training set, wherein annotation information of the second sample image comprises the image area of the first target and a category corresponding to the category confidence of the first target, and the training set comprises an annotated third sample image;

for a second target with the category confidence less than the first threshold value in the target, performing feature-relevance mining on the second target according to the feature information of a third target in the third sample image, determining a fourth target and the first sample image where the fourth target is located from the second target through feature-relevance mining, taking the first sample image where the fourth target is located as the fourth sample image, and adding the fourth sample image to the training set; and

training the target detection network according to the annotation information of the fourth sample image, as well as the second sample image, the third sample image and the fourth sample image in the training set.

2. The method according to claim 1, wherein training the target detection network according to the annotation information of the fourth sample image, as well as the second sample image, the third sample image and the fourth sample image in the training set comprises:

respectively determining a first number of samples from a positive sample image of each category according to the category of the target in positive sample images of the training set, wherein the positive sample image is a sample image comprising the target;

sampling the positive sample image of the each category according to the first number of the samples in the positive sample image of the each category to obtain a plurality of fifth sample images;

sampling a negative sample image of the training set to obtain a plurality of sixth sample images, wherein the negative sample image is a sample image that does not comprise the target; and

training the target detection network according to the fifth sample images and the sixth sample images.

3. The method according to claim 1 or 2, wherein performing feature-relevance mining on the second target according to the feature information of the third target in the third sample image, and determining the fourth target and the first sample image where the fourth target is located from the second target through feature-relevance mining comprise:

determining information entropy of the second target according to the class probability of the second target;

selecting a fifth target from the second target according to the category confidence and the information entropy of the second target;

respectively determining a second number of sample images to be mined in each category according to the category of the third target in the third sample image and a total number of sample images to be mined; and

determining the fourth target and the first sample image where the fourth target is located from the fifth target according to the feature information of the third target in the third sample image, feature information of the fifth target and the second number of the sample images to be mined in the each category.

4. The method according to claim 3, wherein selecting the fifth target from the second target according to the category confidence and the information entropy of the second target comprises:

sorting the second target according to the category confidence and the information entropy of the second target respectively , and selecting a third number of sixth targets and a fourth number of seventh targets; and

combining the sixth targets and the seventh targets to obtain the fifth target.

5. The method according to claim 3 or 4, wherein respectively determining the second number of the sample images to be mined in the each category according to the category of the third target in the third sample image and the total number of the sample images to be mined comprises:

determining a proportion of the third target of each category according to the category of the third target in the third sample image;

determining the sampling proportion of each category according to the proportion of the third target of the each category; and

respectively determining the second number of the sample images to be mined in the each category according to the sampling proportion of the each category.

6. The method according to any one of claims 3 to 5, wherein determining the fourth target and the first sample image where the fourth target is located from the fifth target according to the feature information of the third target in the third sample image, the feature information of the fifth target and the second number of the sample images to be mined in the each category comprises:

respectively determining the third target at a smallest distance from each fifth target among the third target of a first category according to the distance between the feature information of the third target of the first category and the feature information of the each fifth target, and taking the third target as an eighth target, wherein the first category is any one of categories of the third target; and

determining a target at a largest distance among the eighth target as the fourth target.

7. The method according to claim 6, wherein determining the fourth target and the first sample image where the fourth target is located from the fifth target according to the feature information of the third target in the third sample image, the feature information of the fifth target and the second number of the sample images to be mined in the each category further comprises:

adding the determined fourth target to the third target of the first category, and removing the determined fourth target from the unannotated fifth target.

8. The method according to any one of claims 1 to 7, further comprising:

inputting the third sample image into the target detection network for processing to obtain the feature information of the third target in the third sample image.

9. The method according to any one of claims 1 to 8, wherein before the step of inputting the unannotated first sample image into the target detection network for processing to obtain the target detection result of the first sample image, the method further comprises:

pre-training the target detection network through the annotated third sample image.

10. The method according to any one of claims 1 to 9, wherein the first sample image comprises a long-tailed image.

11. The method according to claim 2, wherein before the step of respectively determining the first number of samples from the positive sample image of the each category according to the category of the target in the positive sample images of the training set, the method comprises:

sampling the positive sample images and the negative sample image in the training set to obtain a same or similar number of positive sample images and negative sample images.

12. The method according to claim 3, wherein the total number of the sample images to be mined is 5% to 25% of a total number of the first sample images.

13. The method according to claim 4, wherein combining the sixth targets and the seventh targets to obtain the fifth target comprises:

removing a same target as the seventh targets from the sixth targets to obtain a remaining target in the sixth targets different from the seventh targets; and

taking the remaining target and the seventh targets as the fifth target.

14. The method according to claim 6, wherein after the steps of respectively determining the third target at the smallest distance from the each fifth target among the third target of the first category according to the distance between the feature information of the third target of the first category and the feature information of the each fifth target, and taking the third target as an eighth target, the method further comprises:

ending the determination of the eighth target in a case where a number of the first sample image where the fourth target is located reaches the second number of the sample images to be mined of the first category.

15. The method according to claim 7, wherein after the steps of respectively determining the third target at the smallest distance from each fifth target among the third target of the first category according to the distance between the feature information of the third target of the first category and the feature information of the each fifth target, and taking the third target as an eighth target, the method further comprises:

ending the determination of the eighth target in a case where the number of the first sample images where the fourth target is located does not reach the second number of the sample images to be mined of the first category and a set storing the feature information of the fifth target is empty.

16. The method according to claim 8, wherein inputting the third sample image into the target detection network for processing to obtain the feature information of the third target in the third sample image comprises:

inputting the third sample image into the target detection network to obtain a feature vector output by a hidden layer of the target detection network; and
determining the feature vector as the feature information of the third target.

17. A target detection method, comprising:

inputting an image to be processed into a target detection network for processing to obtain a target detection result of the image to be processed, wherein the target detection result comprises a position and a category of a target in the image to be processed;
wherein the target detection network is obtained by training according to a network training method,

wherein the network training method comprises:

inputting an unannotated first sample image into a target detection network for processing to obtain a target detection result of the first sample image, the target detection result comprising an image area, feature information and class probability of a target in the first sample image;

determining category confidence of the target according to the class probability of the target;

for a first target with the category confidence greater than or equal to a first threshold value in the target, taking the first sample image where the first target is located as an annotated second sample image, and adding the second sample image to a training set, wherein annotation information of the second sample image comprises the image area of the first target and a category corresponding to the category confidence of the first target, and the training set comprises an annotated third sample image;

for a second target with the category confidence less than the first threshold value in the target, performing feature-relevance mining on the second target according to the feature information of a third target in the third sample image, determining a fourth target and the first sample image where the fourth target is located from the second target through feature-relevance mining, taking the first sample image where the fourth target is located as the fourth sample image, and adding the fourth sample image to the training set; and

training the target detection network according to the annotation information of the fourth sample image, as well as the second sample image, the third sample image and the fourth sample image in the training set.

18. A network training apparatus, comprising:

a target detection part, configured to input an unannotated first sample image into a target detection network for processing to obtain a target detection result of the first sample image, wherein the target detection result comprises an image area, feature information and class probability of a target in the first sample image;

a confidence determination part, configured to determine category confidence of the target according to the class probability of the target;

an annotation part, configured to, for a first target with the category confidence greater than or equal to a first threshold value in the target, take the first sample image where the first target is located as an annotated second sample image and add the annotated second sample image to a training set, wherein annotation information of the second sample image comprises the image area of the first target and a category corresponding to the category confidence of the first target, and the training set includes an annotated third sample image;

a feature mining part, configured to, for a second target with the category confidence less than the first threshold value in the target, perform feature-relevance mining on the

second target according to the feature information of a third target in the third sample image, determine a fourth target and the first sample image where the fourth target is located from the second target through feature-relevance mining, take the first sample image where the fourth target is located as the fourth sample image, and add the fourth sample image to the training set; and

a training part, configured to train the target detection network according to the annotation information of the fourth sample image, as well as the second sample image, the third sample image and the fourth sample image in the training set.

19. An electronic device, comprising:

a processor; and

a memory, configured to store a processor executable instruction;

wherein the processor is configured to call the instruction stored in the memory to:

input an unannotated first sample image into a target detection network for processing to obtain a target detection result of the first sample image, the target detection result comprising an image area, feature information and class probability of a target in the first sample image;

determine category confidence of the target according to the class probability of the target;

for a first target with the category confidence greater than or equal to a first threshold value in the target, take the first sample image where the first target is located as an annotated second sample image, and add the second sample image to a training set, wherein annotation information of the second sample image comprises the image area of the first target and a category corresponding to the category confidence of the first target, and the training set comprises an annotated third sample image;

for a second target with the category confidence less than the first threshold value in the target, perform feature-relevance mining on the second target according to the feature information of a third target in the third sample image, determine a fourth target and the first sample image where the fourth target is located from the second target through feature-relevance mining, take the first sample image where the fourth target is located as the fourth sample image, and add the fourth sample image to the training set; and

train the target detection network according to the annotation information of the fourth sample image, as well as the second sample image, the third sample image and the fourth sample image in the training set.

20. A computer-readable storage medium, having a computer program instruction stored thereon, wherein the computer program instruction is executed by a processor to:

input an unannotated first sample image into a target detection network for processing to obtain a target detection result of the first sample image, the target detection result comprising an image area, feature information and class probability of a target in the first sample image;

determine category confidence of the target according to the class probability of the target;

for a first target with the category confidence greater than or equal to a first threshold value in the target, take the first sample image where the first target is located as an annotated second sample image, and add the second sample image to a training set, wherein annotation information of the second sample image comprises the image area of the first target and a

category corresponding to the category confidence of the first target, and the training set comprises an annotated third sample image;

for a second target with the category confidence less than the first threshold value in the target, perform feature-relevance mining on the second target according to the feature information of a third target in the third sample image, determine a fourth target and the first sample image where the fourth target is located from the second target through feature-relevance mining, take the first sample image where the fourth target is located as the fourth sample image, and add the fourth sample image to the training set; and

train the target detection network according to the annotation information of the fourth sample image, as well as the second sample image, the third sample image and the fourth sample image in the training set.