

## CLAIMS

1. A method for identifying a target, comprising:

acquiring a first image and a second image, the first image and the second image each comprising a target to be determined;

generating a prediction path based on the first image and the second image, both ends of the prediction path respectively corresponding to the first image and the second image; and

performing validity determination on the prediction path, and determining, according to a determination result, whether the targets to be determined in the first image and the second image are the same target to be determined;

wherein before generating the prediction path based on the first image and the second image, the method further comprises:

determining a preliminary sameness probability value of the targets to be determined respectively contained in the first image and the second image according to temporal information, spatial information, and image feature information of the first image and temporal information, spatial information, and image feature information of the second image;

wherein generating the prediction path based on the first image and the second image comprises:

generating the prediction path based on the first image and the second image if the preliminary sameness probability value is greater than a preset value.

2. The method according to claim 1, wherein determining the preliminary sameness probability value of the targets to be determined respectively contained in the first image and the second image according to the temporal information, the spatial information, and the image feature information of the first image and the temporal information, the spatial information, and the image feature information of the second image comprises:

inputting the first image, the second image, and a difference in temporal information and a difference in spatial information between the first image and the second image into a Siamese Convolutional Neural Network (Siamese-CNN) to obtain

the preliminary sameness probability value of the targets to be determined in the first image and the second image.

3. The method according to claim 1 or 2, wherein generating the prediction path based on the first image and the second image comprises:

generating, through a probability model, the prediction path of the targets to be determined according to the feature information of the first image, the temporal information of the first image, the spatial information of the first image, the feature information of the second image, the temporal information of the second image, and the spatial information of the second image.

4. The method according to claim 3, wherein generating, through the probability model, the prediction path of the targets to be determined comprises:

determining, through a chain Markov Random Field (MRF) model, all images comprising information of the targets to be determined and having a spatiotemporal sequence relationship with the first image and the second image from an acquired image set; and

generating, according to temporal information and spatial information corresponding to all the determined images, the prediction path of the targets to be determined;

wherein generating, according to the temporal information and spatial information corresponding to all the determined images, the prediction path of the targets to be determined comprises:

generating a prediction path with the first image as a head node and the second image as a tail node according to the temporal information and the spatial information corresponding to all the determined images, wherein the prediction path further corresponds to at least one intermediate node in addition to the head node and the tail node.

5. The method according to claim 4, wherein determining, through the chain MRF model, all images comprising the information of the targets to be determined and having the spatiotemporal sequence relationship with the first image and the second image from the acquired image set comprises:

acquiring position information of all camera devices from a start position to an end position by using a position corresponding to the spatial information of the first image as the start position and using a position corresponding to the spatial information of the second image as the end position;

generating, according to the relationships between positions indicated by the position information of all the camera devices, at least one device path by using a camera device corresponding to the start position as a start point and using a camera device corresponding to the end position as an end point, wherein each device path further comprises information of at least one other camera device in addition to the camera device as the start point and the camera device as the end point; and

determining, from images captured by each of the other camera devices on the current path, an image for each device path by using time corresponding to the temporal information of the first image as start time and using time corresponding to the temporal information of the second image as end time, wherein the image comprises the information of the targets to be determined, and has a set temporal sequence relationship with an image which comprises the information of the targets to be determined and is captured by a previous camera device adjacent to the current camera device.

6. The method according to claim 5, wherein generating the prediction path with the first image as the head node and the second image as the tail node according to the temporal information and the spatial information corresponding to all the determined images comprises:

generating, according to the temporal sequence relationship of the determined images, a plurality of connected intermediate nodes having a spatiotemporal sequence relationship for each device path, and generating, according to the head node, the tail node, and the intermediate nodes, an image path having a spatiotemporal sequence relationship and corresponding to the current device path; and

determining, from the image path corresponding to each device path, a maximum probability image path with the first image as the head node and the second image as the tail node as the prediction path of the targets to be determined.

7. The method according to claim 6, wherein determining, from the image path corresponding to each device path, the maximum probability image path with the first

image as the head node and the second image as the tail node as the prediction path of the targets to be determined comprises:

acquiring, for the image path corresponding to each device path, a probability of images of every two adjacent nodes in the image path having information of the same target to be determined;

calculating, according to the probability of the images of every two adjacent nodes in the image path having the information of the same target to be determined, a probability of the image path being a prediction path of the target to be determined; and

determining, according to the probability of each image path being a prediction path of the target to be determined, the maximum probability image path as the prediction path of the target to be determined.

8. The method according to any one of claims 1 to 7, wherein performing the validity determination on the prediction path and determining, according to the determination result, whether the targets to be determined in the first image and the second image are the same target to be determined comprises:

performing, through a neural network, validity determination on the prediction path and determining whether the targets to be determined in the first image and the second image are the same target to be determined according to the determination result, and

the performing, through the neural network, the validity determination on the prediction path and determining whether the targets to be determined in the first image and the second image are the same target to be determined according to the determination result comprises:

acquiring a temporal difference between adjacent images in the prediction path according to temporal information of the adjacent images; acquiring a spatial difference between the adjacent images according to spatial information of the adjacent images; and acquiring a feature difference between the targets to be determined in the adjacent images according to feature information of the targets to be determined in the adjacent images;

inputting the obtained temporal difference, spatial difference, and feature difference between the adjacent images in the prediction path into a Long Short-Term Memory (LSTM) network to obtain an identification probability of the targets to be determined in the prediction path; and

determining, according to the identification probability of the targets to be determined in the prediction path, whether the targets to be determined in the first image and the second image are the same target to be determined.

9. The method according to claim 8, wherein acquiring the feature difference between the targets to be determined in the adjacent images according to the feature information of the targets to be determined in the adjacent images comprises:

separately acquiring feature information of the targets to be determined in the adjacent images through the Siamese-CNN; and

acquiring the feature difference between the targets to be determined in the adjacent images according to the separately acquired feature information.

10. An apparatus for identifying a target, comprising:

an acquisition module configured to acquire a first image and a second image, the first image and the second image each comprising a target to be determined;

a generation module configured to generate a prediction path based on the first image and the second image, both ends of the prediction path respectively corresponding to the first image and the second image;

a first determination module configured to perform validity determination on the prediction path and determine, according to a determination result, whether the targets to be determined in the first image and the second image are the same target to be determined; and

a second determination module configured to determine, according to temporal information, spatial information, and image feature information of the first image and temporal information, spatial information, and image feature information of the second image, a preliminary sameness probability value of the targets to be determined respectively contained in the first image and the second image,

wherein the generation module comprises:

a first generation sub-module configured to generate the prediction path based on the first image and the second image if the preliminary sameness probability value is greater than a preset value.

11. The apparatus according to claim 10, wherein the second determination module comprises:

a first determination sub-module configured to input the first image, the second image, and a difference in temporal information and a difference in spatial information between the first image and the second image into a Siamese Convolutional Neural Network (Siamese-CNN) to obtain a preliminary sameness probability value of the targets to be determined in the first image and the second image.

12. The apparatus according to claim 10 or 11, wherein the generation module comprises:

a second generation sub-module configured to generate the prediction path of the targets to be determined through a probability model according to the feature information of the first image, the temporal information of the first image, the spatial information of the first image, the feature information of the second image, the temporal information of the second image, and the spatial information of the second image.

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

a first determination unit configured to determine, through a chain Markov Random Field (MRF) model, all images comprising information of the targets to be determined and having a spatiotemporal sequence relationship with the first image and the second image from an acquired image set; and

a first generation unit configured to generate, according to temporal information and spatial information corresponding to all the determined images, the prediction path of the targets to be determined;

wherein the first generation unit comprises:

a second generation unit configured to generate a prediction path with the first image as a head node and the second image as a tail node according to the temporal information and the spatial information corresponding to all the determined images, wherein the prediction path further corresponds to at least one intermediate node in addition to the head node and the tail node.

14. The apparatus according to claim 13, wherein the first determination unit is configured to:

acquire position information of all camera devices from a start position to an end position by using a position corresponding to the spatial information of the first image as the start position and using a position corresponding to the spatial information of the second image as the end position;

generate, according to the relationships between positions indicated by the position information of all the camera devices, at least one device path by using a camera device corresponding to the start position as a start point and using a camera device corresponding to the end position as an end point, wherein each device path further comprises information of at least one other camera device in addition to the camera device as the start point and the camera device as the end point; and

determine, from images captured by each of the other camera devices on the current path, an image for each device path by using time corresponding to the temporal information of the first image as start time and using time corresponding to the temporal information of the second image as end time, wherein the image comprises the information of the targets to be determined, and has a set temporal sequence relationship with an image which comprises the information of the targets to be determined and is captured by a previous camera device adjacent to the current camera device.

15. The apparatus according to claim 14, wherein the second generation unit is configured to:

generate, according to the temporal sequence relationship of the determined images, a plurality of connected intermediate nodes having a spatiotemporal sequence relationship for each device path; and generate, according to the head node, the tail node, and the intermediate nodes, an image path having a spatiotemporal sequence relationship and corresponding to the current device path; and

determine, from the image path corresponding to each device path, a maximum probability image path with the first image as the head node and the second image as the tail node as the prediction path of the targets to be determined.

16. The apparatus according to claim 15, wherein the second generation unit is further configured to:

acquire, for the image path corresponding to each device path, a probability of images of every two adjacent nodes in the image path having information of the same target to be determined;

calculate, according to the probability of the images of every two adjacent nodes in the image path having the information of the same target to be determined, a probability of the image path being a prediction path of the target to be determined; and

determine, according to the probability of each image path being a prediction path of the target to be determined, the maximum probability image path as the prediction path of the target to be determined.

17. The apparatus according to any one of claims 11 to 16, wherein the first determination module comprises:

a second determination sub-module configured to perform validity determination on the prediction path through a neural network and determine, according to a determination result, whether the targets to be determined in the first image and the second image are the same target to be determined;

wherein the second determination sub-module comprises:

a first acquisition unit configured to: acquire a temporal difference between adjacent images in the prediction path according to temporal information of the adjacent images; acquire a spatial difference between the adjacent images according to spatial information of the adjacent images; and acquire a feature difference between the targets to be determined in the adjacent images according to feature information of the targets to be determined in the adjacent images;

a second acquisition unit configured to input the obtained temporal difference, spatial difference, and feature difference between the adjacent images in the prediction path into a Long Short-Term Memory (LSTM) network to obtain an identification probability of the targets to be determined in the prediction path; and

a second determination unit configured to determine, according to the identification probability of the targets to be determined in the prediction path, whether the targets to be determined in the first image and the second image are the same target to be determined.

18. The apparatus according to claim 17, wherein the first acquisition unit is configured to:

separately acquire feature information of the targets to be determined in the adjacent images through the Siamese-CNN; and

acquire the feature difference between the targets to be determined in the adjacent images according to the separately acquired feature information.

19. A computer-readable storage medium, having computer program instructions stored thereon, wherein the program instructions, when being executed by a processor, are configured to perform the method according to any one of claims 1 to 11.

20. An electronic device, comprising: a processor, a memory, a communication element, and a communication bus, wherein the processor, the memory, and the communication element communicate with one another via the communication bus; and

the memory is configured to store at least one executable instruction, and the executable instruction causes the processor to execute the method according to any one of claims 1 to 11.