WHAT IS CLAIMED IS

1. A method for detecting a signal light, the method being implemented on a computing device and comprising:

obtaining a first image and a second image previous to the first image in time sequence, wherein both the first image and the second image include a same target signal light;

determining a first state of the target signal light in the first image by a trained first machine learning model that receives the first image as input;

characterized in that;

determining whether the first state is a preset state, the preset state being a state where false detection is prone to occur, the preset state being a red light state or a yellow light state;

in response to a determination that the first state is the preset state,

determining a second state of the target signal light in the first image by a trained second machine learning model that receives the second image and the first image as input, including:

determining a state change of the target signal light from a time point when the second image is captured to a time point when the first image is captured by the trained second machine learning model that receives the second image and the first image as input; and

determining, by analyzing a target state of the target signal light at the time point when the second image is captured and the state change, the second state of the target signal light in the first image; and

determining one of the first state and the second state as a target state of the target signal light at the time point when the first image is captured; and

in response to a determination that the first state is not the preset state, determining the first state as the target state.

2. The method of claim 1, wherein the first machine learning model determines the first

state by receiving position information and color information of the target signal light in the first image as input.

3. The method of claim 1, wherein the determining, by analyzing a target state of the target signal light at the time point when the second image is captured and the state change, the second state of the target signal light in the first image includes:

obtaining a first confidence level of the first state, an output of the trained first machine learning model includes the first state and the first confidence level of the first state;

obtaining, by processing a confidence level corresponding to the state change, a second confidence level of the second state or an adjusted second confidence level, an output of the trained second machine learning model includes the state change and the confidence level of the state change; and

determining, by comparing the first confidence level and one of the second confidence level and the adjusted second confidence level, the target state.

4. The method of claim 3, wherein the determining, by comparing the first confidence level and one of the second confidence level and the adjusted second confidence level, the target state includes:

obtaining an adjusted first confidence level corresponding to the first state by adjusting the first confidence level; and

determining, by comparing the adjusted first confidence level and one of the second confidence level and the adjusted second confidence level, one of the first state and the second state as the target state of the target signal light at the time point when the first image is captured.

5. The method of claim 4, wherein the obtaining an adjusted first confidence level corresponding to the first state by adjusting the first confidence level includes:

determining a product of the first confidence level and a preset first factor as the adjusted first confidence level.

6. The method of any one of claims 3-5, wherein the obtaining, by processing a confidence level corresponding to the state change, one of a second confidence level of the second state and an adjusted second confidence level includes:

if the state change is that a color of the target signal light does not change from the time point when the second image is acquired to the time point when the first image is acquired,

determining a product of the confidence level of the target state corresponding to the target signal light at the time point when the second image is captured and the confidence level corresponding to the state change as the second confidence level of the second state, and

determining a product of a confidence level of a target state corresponding to the target signal light at the time point when the second image is captured, the confidence level corresponding to the state change, and a preset second factor as the adjusted second confidence level;

if the state change is that the color of the target signal light changes from the time point when the second image is acquired to the time point when the first image is acquired,

determining a sum of the confidence level of the target state corresponding to the target signal light at the time point when the second image is captured and the confidence level corresponding to the state change as the second confidence level of the second state, and

determining an average value of the confidence level corresponding to the state change and the confidence level of the target state corresponding to the target signal light at the time point when the second image is captured as the adjusted second confidence level.

7. The method of any one of claims 3-5, further comprising:

obtaining a first position of the target signal light in the first image;

determining a first comparison result by comparing the first position and a

reference position; and

determining, by comparing the first comparison result and a first threshold, whether the first confidence level needs to be corrected.

8. The method of any one of claims 3-5, further comprising:

obtaining a second position of the target signal light in the first image;

determining a second comparison result by comparing the second position and a reference position; and

determining, by comparing the second comparison result with a second threshold, whether the second confidence level needs to be corrected.

9. The method of claim 7 or 8, further comprising:

in response to a determination that the first confidence level needs to be corrected, correcting the first confidence level by performing operations on an intersection-over-union between the first position and the reference position and the first confidence level; or

in response to a determination that the second confidence level needs to be corrected, correcting the second confidence level by performing operations on an intersection-over-union between the second position and the reference position and the second confidence level.

10. The method of claim 7 or 8, wherein the reference position is determined by operations including:

obtaining a sample set including positions of the target signal light in a plurality of sample images captured within a preset time period, the preset time period being before the time point when the second image is captured;

obtaining, by the mean a clustering algorithm, a clustering result by clustering the positions of the target signal light in the plurality of sample images in the sample set; and

determining a cluster center in one of one or more cluster sets in the clustering result as the reference position.

11. The method of claim 10, wherein the obtaining, by the mean a clustering algorithm, a clustering result by clustering the positions of the target signal light in the plurality of sample images in the sample set includes:

determining, by sampling in the sample set, an initialization reference position set and a radius, the initialization reference position set including one or more initialization reference positions;

determining positions in the sample sets within the radius of an initialization reference position in the initialization reference position set as an updated reference position set;

determining whether a termination condition is satisfied; and

in response to a determination that the termination condition is satisfied, obtaining the clustering result, wherein the clustering result includes the updated reference position set.

12. The method of claim 10, further comprising:

updating, based on a position of the target signal light in the first image and/or the second image, the sample set and the reference position.

13. The method of any one of claims 1-12, further comprising:

performing a preprocessing operation on at least one of the first image or the second image, wherein the preprocessing operation includes at least one of:

performing color conversion on at least one of the first image or the second image; or

performing region of interest (ROI) extraction on at least one of the first image or the second image, wherein the region of interest includes a light panel region where the target signal light is located.

14. The method of claim 13, wherein the region of interest is larger than the light panel region.

15. A device for detecting a signal light comprising a processor, a storage, and a communication circuit, wherein the processor is respectively coupled to the storage and the communication circuit, program data is stored in the storage, and the processor implements the method for detecting a signal light according to any one of claims 1 to 14 by executing the program data in the storage.

16. A computer-readable storage medium storing computer programs, wherein the computer programs are executed by a processor to implement the method for detecting a signal light according to any one of claims 1 to 14.

17. The method of claim 2, wherein the trained first machine learning model is obtained based on a plurality of training samples and training labels, a training sample including a sample first image and a label including a sample multi-dimensional matrix corresponding to the sample first image, a sample confidence level, and/or a sample first state of a sample signal light in the sample first image.

18. The method of claim 1, wherein the trained second machine learning model is obtained based on a plurality of training samples and training labels, a training sample including a sample first image and a sample second image, a label including a sample multi-dimensional matrix corresponding to the sample first image, a sample confidence level and/or a sample first state of a sample signal light in the sample first image.

19. The method of claim 3, wherein

the output of the trained first machine learning model includes a first multi-dimension matrix, each of elements in the first multi-dimensional matrix corresponding to a pixel in the first image, the method further comprises:

performing, by using a non-maximum suppression algorithm, a filtering operation on confidence levels of different portions of the target signal light in the first multi-dimensional matrix to obtain first filtered confidence levels; and

designate a state corresponding to a maximum value or an average value of the first filtered confidence levels as the first state and the maximum value or the average value of the first filtered confidence levels as the first confidence level of the first state;

the output of the trained second machine learning model includes a second multi-dimension matrix, each of elements in the second multi-dimensional matrix corresponding to a pixel in the first image and a corresponding pixel in the second image, and the method further comprises:

performing, by using the non-maximum suppression algorithm, a filtering operation on confidence levels of different portions of the target signal light in the second multi-dimensional matrix to obtain second filtered confidence levels; and

designate a state corresponding to a maximum value or an average value of the second filtered confidence levels as the state change and the maximum value or the average value of the second filtered confidence levels as the confidence level of the state change.

20. The method of claim 13, wherein the performing region of interest (ROI) extraction on at least one of the first image or the second image includes:

extracting a plurality of ROIs from each of the first image or the second image;

performing clustering processing on the plurality of ROIs through a clustering algorithm;

classifying the ROIs into a plurality of categories; and

setting a detection region corresponding to each category, wherein the detection region includes all or at least part of the ROIs in one same category.