CLAIMS

1. A method for child state analysis, comprising:

performing face feature extraction on at least one image frame in an obtained video stream;

classifying whether a person in the image is a child and at least one state of the person according to face features to obtain a first classification result of whether the person in the image is a child, and a second classification result of the at least one state of the person; and

outputting the first classification result and the second classification result; and/or outputting prompt information according to the first classification result and the second classification result.

2. The method according to claim 1, further comprising:

in response to the first classification result indicating that the person in the image is a child, determining whether the second classification result satisfies a predetermined condition; and

in response to the second classification result satisfying the predetermined condition, outputting prompt information.

3. The method according to claim 1 or 2, wherein the state of the person comprises at least one of the following: a normal state, an abnormal state, a sleep state, a wake-up state, or a state of leaving a child seat; and

the second classification result comprises at least one of the following: whether the state of the person is the normal state, whether the state of the person is the abnormal state, whether the state of the person is the sleep state, whether the state of the person is the wake-up state, or whether the state of the person is the state of leaving the child seat.

4. The method according to claim 3, wherein the abnormal state comprises at least one of the following: a crying state, irritable state, a vomiting state, a choking state, or a pain state,

wherein the predetermined condition comprises at least one of the following: the state of the person is the abnormal state, the state of the person is the wake-up

state, or the state of the person is the state of leaving the child seat.

5. The method according to claim 4, wherein in response to the second classification result satisfying that the state of the person is the abnormal state, outputting the prompt information comprises:

in response to the person being in the crying state for a first preset duration, and/or the person being in the irritable state for a second preset duration, and/or the person being in the pain state for a third preset duration, outputting the prompt information.

6. The method according to any one of claims 1 to 5, wherein performing face feature extraction on the at least one image frame in the obtained video stream comprises: performing face feature extraction on the at least one image frame in the obtained video stream by using a feature extraction branch of a neural network; and

classifying whether the person in the image is a child and at least one state of the person according to the face features comprises: respectively determining whether the person in the image is a child and classification of the at least one state of the person based on the face features extracted by the feature extraction branch by using at least two classification branches connected to the feature extraction branch in the neural network.

7. The method according to any one of claims 1 to 5, wherein performing face feature extraction on the at least one image frame in the obtained video stream comprises: performing face feature extraction on the at least one image frame in the obtained video stream by using a feature extraction branch of a neural network;

classifying whether the person in the image is a child and at least one state of the person according to the face features comprises: respectively determining whether the person in the image is a child and classification of at least one face state of the person based on the face features extracted by the feature extraction branch by using at least two classification branches connected to the feature extraction branch in the neural network; and

performing statistics on classification results of the at least one face state of the person corresponding to the at least one image frame in the video stream by using a statistical classification branch connected to the at least two classification branches in the neural network, and determining the classification of the at least one state of the

person based on a statistical result.

8. The method according to claim 7, wherein the face state of the person comprises at least one of the following: an eye open state or an eye closed state.

9. The method according to any one of claims 6 to 8, wherein the classification branch for determining whether the person in the image is a child is pre-trained based on annotation data for children of distinguishing genders.

10. The method according to any one of claims 6 to 9, wherein before performing face feature extraction on the at least one image frame in the obtained video stream, the method further comprises:

training the neural network by using a sample image which is provided with child annotation information and state annotation information.

11. The method according to claim 10, wherein the child annotation information is used to indicate whether the person in the sample image is a child; in response to an age or estimated age of the person in the sample image being greater than an age or estimated age of a person in a reference image, the child annotation information of the sample image is used to indicate that the person is not a child; in response to the age or estimated age of the person in the sample image being less than or equal to the age or estimated age of the person in the reference image, the child annotation information of the sample image is used to indicate that the person is a child.

12. The method according to claim 11, wherein the reference image comprises a boy reference image and a girl reference image.

13. The method according to any one of claims 10 to 12, wherein training the neural network by using the sample image comprises:

performing face feature extraction on the sample image by using the feature extraction branch;

classifying whether the person in the sample image is a child and at least one state of the person by sharing face features extracted by the at least two classification branches, to obtain a first predicted classification result of whether the person in the sample image is a child and a second predicted classification result of the at least one state of the person;

obtaining a first loss based on the first predicted classification result and the child annotation information, and obtaining a second loss based on the second predicted classification result and the state annotation information; and

adjusting parameters of the neural network based on the first loss and the second loss.

14. The method according to claim 13, wherein adjusting the parameters of the neural network based on the first loss and the second loss comprises:

performing weighted summation on the first loss and the second loss to obtain a network loss; and

adjusting parameters of the feature extraction branch and the at least two classification branches based on the network loss.

15. The method according to any one of claims 1 to 14, further comprising:

displaying at least one piece of the following information by using a vehicle-mounted device and/or a terminal device: the obtained video stream, the first classification result, the at least one second classification result, or the prompt information.

16. The method according to any one of claims 1 to 15, wherein before performing face feature extraction on the at least one image frame in the obtained video stream, the method further comprises:

obtaining the video stream by using at least one camera provided in a vehicle.

17. An apparatus for child state analysis, comprising:

a feature extraction unit, configured to perform face feature extraction on at least one image frame in an obtained video stream;

a classification unit, configured to classify whether a person in the image is a child and at least one state of the person according to face features to obtain a first classification result of whether the person in the image is a child, and a second classification result of the at least one state of the person; and

a result output unit, configured to output the first classification result and the second classification result; and/or output prompt information according to the first classification result and the second classification result.

18. An electronic device, comprising:

the apparatus for child state analysis according to claim 17; or

the electronic device comprising:

a memory, configured to store executable instructions; and

a processor, configured to communicate with the memory to execute the executable instructions so as to implement the method for child state analysis, the method comprising:

performing face feature extraction on at least one image frame in an obtained video stream;

classifying whether a person in the image is a child and at least one state of the person according to face features to obtain a first classification result of whether the person in the image is a child, and a second classification result of the at least one state of the person; and

outputting the first classification result and the second classification result; and/or outputting prompt information according to the first classification result and the second classification result.

19. A computer storage medium, configured to store computer readable instructions, wherein the instructions, when being executed, enables to implement operations of the method for child state analysis according to any one of claims 1 to 16.

20. A computer program product, comprising a computer readable code, wherein the computer readable code, when running on a device, enables a processor in the device to execute instructions for implementing the method for child state analysis according to any one of claims 1 to 16.