

METHOD AND APPARATUS FOR IDENTIFYING DANGEROUS ACTIONS OF PERSONS IN VEHICLE, ELECTRONIC DEVICE AND STORAGE MEDIUM

CROSS-REFERENCE TO RELATED APPLICATION

The present application claims priority to Chinese Patent Application No. CN 201910152525.X, filed on February 28, 2019 and entitled "METHOD AND DEVICE FOR RECOGNIZING DANGEROUS ACTION OF PERSONNEL IN VEHICLE, ELECTRONIC DEVICE, AND STORAGE MEDIUM", which is incorporated herein by reference in its entirety.

TECHNICAL FIELD

The present disclosure relates to computer vision technologies, and in particular, to a method and device for recognizing a dangerous action of personnel in a vehicle, an electronic device, and a storage medium.

BACKGROUND

With the rapid development of in-vehicle personnel and employee intelligence, various artificial intelligence (AI) technologies are implemented, and at present, the demand for driver monitoring in the market is increasingly urgent. The main function modules of driver monitoring may be generally summarized as modules such as in-vehicle personnel face recognition module and fatigue detection module. By monitoring the state of a driver, a danger signal may be found out in time, and a possible danger may be prevented and dealt with in advance so as to improve driving safety.

SUMMARY

Embodiments of the present disclosure provide an in-vehicle personnel dangerous action recognition technology.

According to an aspect of the embodiments of the present disclosure, a method for recognizing a dangerous action is provided, including:

obtaining at least one video stream of personnel in a vehicle through an image

capturing device, each video stream including information about at least one of the personnel in the vehicle;

performing action recognition on the personnel in the vehicle based on the video stream; and

responsive to that a result of the action recognition belongs to a predetermined dangerous action, performing at least one of: sending prompt information, or executing an operation to control the vehicle, where the predetermined dangerous action includes at least one of the following action representations of the personnel in the vehicle: a distraction action, a discomfort state, or a non-standard behavior.

According to another aspect of the embodiments of the present disclosure, a device for recognizing a dangerous action of personnel in a vehicle is provided, including:

a video collection unit, used for obtaining at least one video stream of the personnel in the vehicle through an image capturing device, each video stream including information about at least one of the personnel in the vehicle;

an action recognition unit, used for performing action recognition on the personnel in the vehicle based on the video stream; and

a danger processing unit, used for, responsive to that a result of the action recognition belongs to a predetermined dangerous action, performing at least one of: sending prompt information, or executing an operation to control the vehicle, where the predetermined dangerous action includes at least one of the following action representations of the personnel in the vehicle: a distraction action, a discomfort state, or a non-standard behavior.

According to still another aspect of the embodiments of the present disclosure, an electronic device is provided and includes a processor, where the processor includes the device for recognizing the dangerous action of the personnel in the vehicle according to any one of the foregoing embodiments.

According to yet another aspect of the embodiments of the present disclosure, an electronic device is provided and includes: a memory used for storing executable instructions;

and a processor, used for communicating with the memory to execute the executable instructions so as to complete operations of the method for recognizing the dangerous action of the personnel in the vehicle according to any one of the foregoing embodiments.

According to a further aspect of the embodiments of the present application, a computer readable storage medium is provided and is used for storing computer readable instructions that, when executed, execute the operations of the method for recognizing the

dangerous action of the personnel in the vehicle according to any one of the foregoing embodiments.

According to yet another aspect of the embodiments of the present disclosure, a computer program product is provided and includes computer readable codes that, when being run on a device, cause a processor in the device to execute instructions for implementing the method for recognizing the dangerous action of the personnel in the vehicle according to any one of the foregoing embodiments.

On the basis of the method and device for recognizing the dangerous action of the personnel in the vehicle, the electronic device, and the storage medium provided in the foregoing embodiments of the present disclosure, at least one video stream of the personnel in the vehicle is obtained through the image capturing device, each video stream including information about at least one of the personnel in the vehicle; action recognition is performed on the personnel in the vehicle on the basis of the video stream; and responsive to that a result of the action recognition belongs to a predetermined dangerous action, at least one of the following is performed: the prompt information is sent or the operation is executed to control the vehicle, where the predetermined dangerous action includes at least one of the following action representations of the personnel in the vehicle: the distraction action, the discomfort state, or the non-standard behavior. Whether the personnel in the vehicle make the predetermined dangerous action is determined by means of action recognition, and at least one of a corresponding prompt or operation are made to the predetermined dangerous action so as to control the vehicle, thereby implementing early detection of a vehicle safety condition to reduce the probability of dangerous situations.

The following further describes in detail the technical solutions of the present disclosure with reference to the accompanying drawings and embodiments.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings constituting a part of the specification describe the embodiments of the present disclosure and are intended to explain the principles of the present disclosure together with the descriptions.

According to the following detailed descriptions, the present disclosure may be understood more clearly with reference to the accompanying drawings.

FIG. 1 is a schematic flowchart of a method for recognizing a dangerous action of personnel in a vehicle provided in the embodiments of the present disclosure.

FIG. 2 is a part of a schematic flowchart in an optional example of a method for recognizing a dangerous action of personnel in a vehicle provided in the embodiments of the present disclosure.

FIG. 3a is a part of a schematic flowchart in another optional example of a method for recognizing a dangerous action of personnel in a vehicle provided in the embodiments of the present disclosure.

FIG. 3b is a schematic diagram of an extracted target area in a method for recognizing a dangerous action of personnel in a vehicle provided in the embodiments of the present disclosure.

FIG. 4 is a schematic structural diagram of a device for recognizing a dangerous action of personnel in a vehicle provided in the embodiments of the present disclosure.

FIG. 5 is a schematic structural diagram of an electronic device, which may be a terminal device or a server, suitable for implementing the embodiments of the present disclosure.

DETAILED DESCRIPTION

Exemplary embodiments of the present disclosure are described in detail with reference to the accompany drawings now. It should be noted that, unless otherwise stated specifically, relative arrangement of the components and steps, the numerical expressions, and the values set forth in the embodiments are not intended to limit the scope of the present disclosure.

In addition, it should be understood that, for ease of description, the size of each part shown in the accompanying drawings is not drawn in actual proportion.

The following descriptions of at least one exemplary embodiment are merely illustrative actually, and are not intended to limit the present disclosure and the applications or uses thereof.

Technologies, methods and devices known to a person of ordinary skill in the related art may not be discussed in detail, but such technologies, methods and devices should be considered as a part of the specification in appropriate situations.

It should be noted that similar reference numerals and letters in the following accompanying drawings represent similar items. Therefore, once an item is defined in an accompanying drawing, the item does not need to be further discussed in the subsequent accompanying drawings.

The embodiments of the present disclosure may be applied to a computer system/server, which may operate with numerous other general-purpose or special-purpose computing system environments or configurations. Examples of well-known computing systems, environments, and/or configurations suitable for use together with the computer system/server include, but are not limited to, personal computer systems, server computer systems, thin clients, thick clients, handheld or laptop devices, microprocessor-based systems, set top boxes, programmable consumer electronics, network personal computers, small computer systems, large computer systems, distributed cloud computing environments that include any one of the foregoing systems, and the like.

The computer system/server may be described in the general context of computer system executable instructions (for example, program modules) executed by the computer system. Generally, the program modules may include routines, programs, target programs, components, logics, data structures, and the like for performing specific tasks or implementing specific abstract data types. The computer system/server may be practiced in the distributed cloud computing environments in which tasks are performed by remote processing devices that are linked through a communications network. In the distributed cloud computing environments, the program modules may be located in local or remote computing system storage media including storage devices.

Dangerous action recognition has a wide application prospect in in-vehicle security monitoring fields. First, a dangerous action recognition system may give a prompt when a driver makes a dangerous action, thereby performing early-warning and avoiding possible accidents; furthermore, the system may monitor some behaviors that are substandard or may cause discomfort of an in-vehicle passenger to give a prompt and restrain the behaviors; meanwhile, the monitoring of the dangerous action reflects some habits and hobbies of a driver, which facilitates the system to establish a user portrait and perform big data analysis; and an emotional state, a fatigue state, and a behavior habit of the driver is monitored by means of dangerous action recognition.

FIG. 1 is a schematic flowchart of a method for recognizing a dangerous action of

personnel in a vehicle (i.e., in-vehicle personnel) provided in the embodiments of the present disclosure. The method is performed by any electronic device, such as a terminal device, a server, a mobile device, and a vehicle-mounted device. As shown in FIG. 1, the method of this embodiment includes the following steps.

At step 110, at least one video stream of in-vehicle personnel is obtained through an image capturing device (for example, a camera or the like).

Each video stream includes information about at least one of the personnel in the vehicle; in the embodiments of the present disclosure, an image of the in-vehicle personnel is collected by means of a photographing apparatus (also called the image capturing device, for example, one or more cameras provided in the vehicle and used for performing photographing on a vehicle seat position), so as to obtain the video stream; optionally, the video streams of multiple in-vehicle personnel (for example, all of in-vehicle personnel) in the whole vehicle are collected on the basis of one photographing apparatus, or one photographing apparatus facing one or more back row areas for image collection is provided in the vehicle, or one photographing apparatus is respectively provided in the front of each seat, so as to respectively collect the video stream of at least one in-vehicle personnel (for example, each in-vehicle personnel); and action recognition is respectively performed on the in-vehicle personnel by performing processing on the collected video stream.

In practical applications, a situation that video stream collection is performed on the in-vehicle personnel on the basis of an out-vehicle photographing apparatus (for example, a camera provided on a road) may further exist.

Optionally, the photographing apparatus includes, but is not limited to, at least one of the followings: a visible light camera, an infrared camera, or a near-infrared camera. The visible light camera is used for collecting a visible light image, the infrared camera is used for collecting an infrared image, and the near-infrared camera is used for collecting a near-infrared image.

In an optional example, step 110 is executed by a processor by invoking a corresponding instruction stored in a memory, or is executed by a video collection unit 41 run by the processor.

At step 120, action recognition is performed on the in-vehicle personnel on the basis of the video stream.

In a driving process of the vehicle, if one or more of in-vehicle personnel make a dangerous action, the danger of the vehicle is caused, and in particular, if the driver makes some dangerous actions, the whole vehicle may be in danger, thereby causing danger to the vehicle and the in-vehicle personnel. Therefore, it is necessary to perform recognition on the actions of the in-vehicle personnel, so as to ensure the safety of the vehicle. Some actions are determined on the basis of a single image frame in the video stream, while some actions require continuous multiple frames to be recognized. Therefore, in the embodiments of the present disclosure, recognition is performed on the actions by means of the video stream to reduce misjudgment, thereby improving the accuracy of action recognition.

Optionally, an action category is divided into a dangerous action and a normal action, and the dangerous action needs to be processed to exclude a possible danger, where the dangerous action includes, but is not limited to, at least one of the followings: a distraction action, a discomfort state, or a non-standard behavior or the like. The dangerous action may have same or different requirements for an ordinary non-driver and driver, for example, the requirements for the driver are relatively stricter, and meanwhile, the independence and safety of the driver need to be protected, for example, the predetermined dangerous action is divided into the dangerous action of the driver and the dangerous action of the non-driver. The embodiments of the present disclosure do not limit a specific mode for recognizing the action category.

In an optional example, step 120 is performed by the processor by invoking the corresponding instruction stored in the memory, or is performed by an action recognition unit 42 run by the processor.

At step 130, prompt information is sent and/or an operation is executed to control the vehicle in response to that an action recognition result (i.e., a result of the action recognition) belongs to a predetermined dangerous action.

The dangerous action in the embodiments of the present disclosure is a behavior that causes potential safety hazards to the in-vehicle personnel or others. Optionally, the predetermined dangerous action in the embodiments of the present disclosure includes, but is not limited to, at least one of the following action representations of the in-vehicle personnel: a distraction action, a discomfort state, or a non-standard behavior and the like. Optionally, the distraction action mainly aims at the driver; in a process of driving a vehicle, the driver needs to concentrate; when the distraction action (for example, actions such as eating food and smoking) occurs, the attention of the driver is influenced, and the vehicle is prone to be in

danger; the discomfort state may aim at all of the in-vehicle personnel; when the discomfort state occurs in the in-vehicle personnel, on the basis of human safety considerations, some dangerous situations need to be processed in time, for example, situations such as frequent yawning of the driver or sweat wiping of a passenger; and the non-standard behavior is a behavior that do not comply with safe driving regulations, and may further be a behavior that is dangerous to the driver or other people in the vehicle, etc. In order to overcome an adverse effect of the predetermined dangerous action, in the embodiments of the present disclosure, the probability of danger is reduced by sending the prompt information or executing the operation to control the vehicle, and the safety and/or comfort level of the in-vehicle personnel are improved.

Optionally, a representation form of the prompt information may include, but is not limited to, at least one of the followings: voice prompt information, vibration prompt information, light prompt information, or smell prompt information and the like; for example, when the in-vehicle personnel smokes, the voice prompt information is sent to prompt that smoking in the vehicle is not allowed, so as to reduce the danger of smoking to other in-vehicle personnel; for another example: when the in-vehicle personnel wipes sweat, it means that a temperature in the vehicle is too high, and an in-vehicle air conditioning temperature is reduced by means of intelligent control, so as to solve a problem of the discomfort of the in-vehicle personnel.

Dangerous action recognition has an important status and high application value in driver monitoring. At present, in the driving process of the driver, many dangerous actions commonly exist, and the actions often cause the driver to be distracted, so that certain potential safety hazards exist.

In an optional example, step 130 is performed by the processor by invoking the corresponding instruction stored in the memory, or is performed by a danger processing unit 43 run by the processor.

On the basis of the method for recognizing the dangerous action of personnel in the vehicle provided in the forgoing embodiments of the present disclosure, at least one video stream of the in-vehicle personnel is obtained by using the photographing apparatus, each video stream including information about at least one in-vehicle personnel; action recognition is performed on the in-vehicle personnel on the basis of the video stream; and the prompt information is sent and/or the operation is executed to control the vehicle in

response to that the action recognition result belongs to the predetermined dangerous action, where the predetermined dangerous action includes at least one of the following action representations of the in-vehicle personnel: the distraction action, the discomfort state, or the non-standard behavior. Whether the in-vehicle personnel make the predetermined dangerous action is determined by means of action recognition, and the corresponding prompt and/or operation are made to the predetermined dangerous action so as to control the vehicle, thereby implementing early detection of the vehicle safety condition to reduce the probability of dangerous situations.

Optionally, the in-vehicle personnel may include a driver and/or a non-driver; the number of in-vehicle personnel generally include at least one (for example: merely including the driver); in order to respectively perform recognition on the action of each in-vehicle personnel, after an image or a video stream is acquired, optionally, the image or the video stream is segmented in terms of different in-vehicle personnel according to different positions (for example, a vehicle seat position), so as to implement performing analysis on the image or the video stream corresponding to each in-vehicle personnel. Because evaluations of dangerous actions of the driver and the non-driver are different in the driving process of the vehicle, when recognizing whether the action is the predetermined dangerous action, optionally, whether the in-vehicle personnel is the driver or the non-driver is first determined.

FIG. 2 is a part of a schematic flowchart in an optional example of the method for recognizing the dangerous action of personnel in the vehicle provided in the embodiments of the present disclosure. As shown in FIG. 2, step 120 includes the followings.

At step 202, at least one target area included by the in-vehicle personnel is detected in at least one frame of video image of the video stream.

In a possible implementation, in order to implement action recognition, the target area may include, but is not limited to, at least one of the followings: a face local area, an action interactive object, or a limb area and the like. For example, when the face local area serves as the target area, because a face action is usually related to five sense organs of the face, for example, a smoking or food-eating action is related to a mouth, and a phone call action is related to an ear; in said example, the target area includes, but is not limited to, one of the following parts or any combination thereof: mouth, ear, nose, eye, and eyebrow. Optionally, a target part on the face is determined according to requirements, where the target part may include one or more parts, and the target part on the face is detected by using a face detection

technology.

At step 204, a target image corresponding to the target area is captured from the at least one frame of video image of the video stream according to the target area obtained through detection.

In a possible implementation, the target area is a certain area centering on the target part, for example, at least one part of the face is used as a center on the basis of the face action. An area outside the face in the video stream may include an object related to an action; for example, a smoking action is centered on the mouth, smoke may appear in areas other than the face in a detection image.

In a possible implementation, a position of the target area is determined in at least one frame of video image according to a detection result of the target area, and a capturing size and/or a capturing position of a target image are determined according to the position of the target area in the at least one frame of video image. In the embodiments of the present disclosure, a target image corresponding to the target area is captured according to a set condition, so that the captured target image better meets the requirements of action recognition. For example, the size of the captured target image is determined according to a distance between the target area and a set position of the face; for example, a target image of the mouth of person A is determined by using a distance between the mouth of person A and a face center point of A, and similarly, a target image of the mouth of person B is determined by using a distance between the mouth of person B and a face center point of B. Because the distance between the mouth and the face center point is related to a feature of the face, the captured target image may better meet the feature of the face. According to the target image captured according to the position of the target area in a video image, noise is reduced, and a more complete image area in which the object related to the action is located may further be included.

At step 206, action recognition is performed on the in-vehicle personnel on according to the target image.

In a possible implementation, the feature of the target image is extracted, and whether the in-vehicle personnel executes the predetermined dangerous action is determined according to the extracted feature.

In a possible implementation, the predetermined dangerous action includes, but is not limited to, at least one of the following action representations of the in-vehicle personnel:

the distraction action, the discomfort state, or the non-standard behavior and the like. When the in-vehicle personnel executes the predetermined dangerous action, potential safety hazards may be generated. Applications such as safety analysis are performed on the in-vehicle personnel by using the action recognition result. For example, when the driver makes the smoking action in the video stream, whether the driver smokes is determined by extracting the feature of the target image of the mouth and determining whether a feature of a cigarette exists in the video stream according to the feature; and if the driver has the smoking action, it is considered that the potential safety hazards exist.

In the embodiments, the target area is recognized in the video stream, the target image corresponding to the target area is captured in the video image according to the detection result of the target area, and whether the in-vehicle personnel executes the predetermined dangerous action is recognized according to the target image. The target image captured according to the detection result of the target area may be applicable to human bodies having different areas in different video images. The embodiments of the present disclosure have a wide application range. On the basis of using the target image as a basis for action recognition, the embodiments of the present disclosure are conducive to obtaining feature extraction corresponding to the dangerous action more accurately, may reduce a detection interference brought by an unrelated area, and improve the accuracy of action recognition. For example, when recognizing the smoking action of the driver, the smoking action is greatly related to a mouth area, and the action of the driver is recognized by using the mouth and the vicinity of the mouth as the mouth area, so as to determine whether the driver smokes, thereby improving the accuracy of smoking action recognition.

FIG. 3a is a part of a schematic flowchart in another optional example of the method for recognizing the dangerous action of personnel in the vehicle provided in the embodiments of the present disclosure. As shown in FIG. 3a, in the method provided in the aforementioned embodiments, step 202 includes the following steps.

At step 302, a feature of the in-vehicle personnel included in the at least one frame of video image of the video stream is extracted.

The embodiments of the present disclosure mainly aim at performing recognition on some dangerous actions made by the in-vehicle personnel when said personnel is inside the vehicle. Moreover, the dangerous actions usually are actions related to the limb and face, and recognition on the actions cannot be implemented by means of detection of a human body key

point and estimation of a human body posture. In the embodiments of the present disclosure, the feature is extracted by performing convolution operation on the video image, and action recognition in the video image is implemented according to the extracted feature. For example, the feature of the aforementioned dangerous action is: the limb and/or face local areas, and the action interactive object. Therefore, real-time photographing needs to be performed on the in-vehicle personnel by means of the photographing apparatus, and a video image including the face is obtained. Then, the convolution operation is performed on the video image, and the action feature is extracted.

At step 304, the target area is extracted from the at least one frame of video image based on the feature.

Optionally, the target area in the embodiments is a target area that may include the action.

The feature of the dangerous action is first defined, and then a neural network implements determination on whether the dangerous action exists in the video image according to the defined feature and the extracted video image. The neural network in the embodiments is trained, i. e. , the neural network may extract the feature of a predetermined action in the video image.

If the extracted feature include: the limb area, the face local area, and the action interactive object, the neural network divides a feature area that simultaneously includes the limb and face local areas and the action interactive object, so as to obtain the target area, where the target area may include, but is not limited to, at least one of the followings: the face local area, the action interactive object, or the limb area and the like. Optionally, the face local area includes, but is not limited to, at least one of the followings: the mouth area, the ear area, or the eye area and the like. Optionally, the action interactive object includes, but is not limited to, at least one of the followings: a container, a cigarette, a mobile phone, food, a tool, a beverage bottle, glasses, or a mask and the like. Optionally, the limb area includes, but is not limited to, at least one of the followings: a hand area or a foot area and the like. For example, the dangerous action includes, but is not limited to: drinking water/beverage, smoking, calling, wearing glasses, wearing a mask, makeup, using a tool, eating food, and putting both feet on a steering wheel and the like. Exemplarily, action features of drinking water may include:

the hand area, the face local area, and a cup; action features of smoking may include: the hand area, the face local area, and the cigarette; action features of calling may include: the hand area, the face local area, and a mobile phone; action features of wearing the glasses may include: the hand area, the face local area, and the glasses; action features of wearing the mask may include: the hand area, the face local area, and the mask; and action features of putting both feet on the steering wheel may include: the foot area and the steering wheel.

The recognized actions in the embodiments of the present disclosure may further include fine actions related to the face or the limb, and such fine actions at least include two features, i. e. , the face local area and the action interactive object, for example, including two features, i. e. , the face local area and the action interactive object, or including two of three features, i. e. , the face local area, the action interactive object and the limb, and the like. Therefore, the fine actions indicate multiple actions having high similarity, for example, smoking and yawning are both recognized mainly on the basis of the mouth area, and both include actions of opening and closing the mouth, and a difference between smoking and yawning actions is merely in whether the cigarette (the action interactive object) is further included. Therefore, the embodiments of the present disclosure implement recognition on the fine actions by extracting the target area to implement action recognition. For example, for the calling action, the target area includes: the face local area, a mobile phone (i. e. , the action interactive object), and a hand (i. e. , the limb area). For another example, for the smoking action, a target action frame may also include: the mouth area and the cigarette (i. e. , the action interactive object).

FIG. 3b is a schematic diagram of an extracted target area in the method for recognizing the dangerous action of personnel in the vehicle in the embodiments of the present disclosure. The method for recognizing the dangerous action of personnel in the vehicle in the embodiments of the present disclosure is used for performing target area extraction on the video image in the video stream, so as to obtain the target area for performing recognition on the actions. The action of the in-vehicle personnel in the embodiments of the present disclosure is the smoking action. Therefore, the obtained target area is based on the mouth area (the face local area) and the cigarette (the action interactive object); the target area obtained on the basis of the embodiments of the present disclosure may confirm that the in-vehicle personnel in FIG. 3b smokes. In the embodiments of the present disclosure, by obtaining the target area, action recognition is performed on the basis of the target area, noise interference in areas of the entire image that are not related to the action of the in-vehicle

personnel (for example, the smoking action) is removed, and the accuracy of action recognition of the in-vehicle personnel is improved, for example, the accuracy of smoking action recognition in the embodiments.

Optionally, before performing action recognition on the in-vehicle personnel according to the target image, pre-processing is further performed on the target image. For example, pre-processing is performed on the target image by means of a method such as normalization and equalization; and a recognition result obtained by the pre-processed target image is more accurate.

Optionally, the dangerous action includes, but is not limited to, at least one of the followings: the distraction action, the discomfort state, or the non-standard behavior or the like. The distraction action indicates that the driver further makes an action that is not related to driving and influences the driving attention degree while driving a vehicle, for example: the distraction action includes, but is not limited to, at least one of the followings: calling, drinking water, putting on or taking off sunglasses, putting on or taking off the mask, or eating food or the like. The discomfort state indicates the physical discomfort of the in-vehicle personnel caused by the influence of an in-vehicle environment or own reasons of the in-vehicle personnel in the vehicle driving process, for example: the discomfort state includes, but is not limited to, at least one of the followings: sweat wiping, rubbing an eye, or yawning or the like. The non-standard behavior indicates a behavior that is made by the in-vehicle personnel and does not comply with regulations, for example, the non-standard behavior includes, but is not limited to, at least one of the followings: smoking, stretching a hand out of the vehicle, bending over a steering wheel, putting both feet on the steering wheel, leaving both hands away from the steering wheel, holding an instrument with a hand, or disturbing a driver or the like. Because multiple dangerous actions are included, when the action category of the in-vehicle personnel belongs to the dangerous action, it is necessary to first determine to which dangerous action the action category belongs, and different dangerous actions may correspond to different processing modes (for example, sending the prompt information or executing the operation to control the vehicle).

In one or more optional embodiments, step 130 includes the followings.

The fact that the action recognition result belongs to the predetermined dangerous action is responded.

A danger level of the predetermined dangerous action is determined.

Corresponding prompt information is sent according to the danger level, and/or an operation corresponding to the danger level is executed and the vehicle is controlled according to the operation.

Optionally, in the embodiments of the present disclosure, when the action of the in-vehicle personnel is determined to belong to the predetermined dangerous action according to the action recognition result, danger level determination is performed on the predetermined dangerous action; optionally, the danger level of the predetermined dangerous action is determined according to a preset rule and/or correspondence, and then how to operate is determined according to the danger level. For example, operations of different degrees are performed according to the dangerous action level of the in-vehicle personnel. For example, if the dangerous action is caused by fatigue and physical discomfort of the driver, timely prompting is required, so that the driver performs adjustment and has a rest in time; and when the driver feels discomfort due to the in-vehicle environment, an adjustment of a certain degree is performed by controlling a ventilation system or an air conditioning system in the vehicle. Optionally, the danger level is set to include primary level, intermediate level, and high level. In this case, the sending of the corresponding prompt information according to the danger level, and/or the executing of the operation corresponding to the danger level and the controlling of the vehicle according to the operation include the followings.

The prompt information is sent in response to that the danger level is the primary level.

The operation corresponding to the danger level is executed and the vehicle is controlled according to the operation in response to that the danger level is the intermediate level.

The operation corresponding to the danger level is executed and the vehicle is controlled according to the operation while sending the prompt information in response to that the danger level is the high level.

In the embodiments of the present disclosure, the danger level is set as 3 levels. Optionally, the embodiments of the present disclosure may further set the danger level more in details, and more levels are included, for example, the danger level includes a first level, a second level, a third level, and a fourth level, where each level corresponds to different danger degrees. The prompt information is sent according to different danger levels, and/or the operation corresponding to the danger level is executed and the vehicle is controlled according to the operation. By executing different operations for different danger levels, the

sending of the prompt information and the controlling of the operation may be more flexible and adapt to different use requirements.

Optionally, the determining of the danger level of the predetermined dangerous action includes the followings.

The frequency and/or duration of occurrence of the predetermined dangerous action in the video stream are acquired, and the danger level of the predetermined dangerous action is determined on the basis of the frequency and/or duration.

In the embodiments of the present disclosure, by performing further abstract analysis on the dangerous action obtained by action recognition, and according to a lasting degree of the action or a priori probability of the occurrence of a dangerous situation, whether a real intention of a passenger performs a dangerous action is output. Optionally, the embodiments of the present disclosure implement measurement of the action lasting degree by means of the frequency and/or duration of the occurrence of the predetermined dangerous action in the video stream. For example, when the driver just scratches the eye quickly, it is considered as just a quick adjustment, and alarming is not required. However, if the driver rubs the eye for a long time along with the occurrence of the action such as yawning, it is considered that the driver is relatively fatigue and should be prompted. For another example, the alarming strength for smoking may be less than that for actions such as bending over the steering wheel and calling.

In a possible implementation, the action includes a duration of an action, an early-warning condition includes: recognizing that the duration of the action exceeds a duration threshold.

In a possible implementation, the action may include the duration of the action; when the duration of the action exceeds the duration threshold, it is considered that the execution of the action distracts much of the attention of an action execution object; said action is considered as a dangerous action, and early-warning information needs to be sent. For example, if the duration of the smoking action of the driver exceeds 3 seconds, it is considered that the smoking action is the dangerous action and influences a driving action of the driver, and the early-warning information needs to be sent to the driver.

In the embodiments, according to the duration and the duration threshold of the predetermined dangerous action, a sending condition of the prompt information and/or a control condition of the vehicle are adjusted, so that the sending of the prompt

information and the controlling of the operation are more flexible and better adapt to different use requirements.

In a possible implementation, the action recognition result includes a duration of an action, and a condition of belonging to the predetermined dangerous action includes: recognizing that the duration of the action exceeds the duration threshold. Some actions do not cause potential safety hazards to the in-vehicle personnel and the vehicle in a short time; only when the duration of the action reaches a set duration threshold, the action is confirmed as the predetermined dangerous action, for example, an action of closing eyes of the driver is considered as a normal blink when the duration of closing eyes is short (for example, 0.5 second), and when the duration of closing eyes exceeds the duration threshold (being set according to the requirements, for example, setting to be 3 seconds), said action is considered to belong to the predetermined dangerous action, and the corresponding prompt information is sent.

In a possible implementation, the action recognition result includes the number of times for which an action is performed, and a condition of belonging to the predetermined dangerous action includes: recognizing that the number of times exceeds a number threshold. When the number of times exceeds the number threshold, it is considered that the action of the action execution object is frequent, and much attention is distracted; the action is considered as the dangerous action, and the early-warning information needs to be sent. For example, if the number of the smoking actions of the driver exceeds 5 times, it is considered that the smoking action is the dangerous action and influences the driving action of the driver, and the prompt information needs to be sent to the driver.

In a possible implementation, the action recognition result includes a duration of an action and the number of times for which the action is performed, and the condition of belonging to the predetermined dangerous action includes: recognizing that the duration of the action exceeds the duration threshold, and the number of times exceeds the number threshold.

In a possible implementation, when the duration of the action exceeds the duration threshold, and the number of times exceeds the number threshold, it is considered that the action of the action execution object is frequent and the duration of the action is long, and much attention is distracted; said action is considered as the dangerous action, and the prompt information needs to be sent and/or the vehicle is controlled, so that the vehicle is more flexibly controlled and adapts to different use requirements.

The dangerous actions corresponding to different in-vehicle personnel are different, for example, the driver at a driving seat is required not to be distracted, and the distraction action belongs to the dangerous action; while the distraction action of the in-vehicle personnel at other positions does not belong to the dangerous action. Therefore, in the embodiments of the present disclosure, in order to implement more accurate alarming and intelligent control, prompting or an intelligent operation is performed by combining the action category and a category of the in-vehicle personnel, so that user experience is not influenced because of frequently alarming while improving the driving safety. Optionally, the embodiments of the present disclosure include: sending corresponding first prompt information and/or controlling the vehicle to execute a corresponding first predetermined operation according to the predetermined dangerous action in response to that the in-vehicle personnel is the driver; and/or,

sending corresponding second prompt information and/or executing a corresponding second predetermined operation according to the predetermined dangerous action in response to that the in-vehicle personnel is the non-driver.

Because the driver is responsible for the safety of the whole vehicle, in order to improve the driving safety of the vehicle and the freedom of the passenger, the in-vehicle personnel are divided into two categories, i. e. , the driver and the non-driver. Different dangerous actions are respectively set for the driver and the non-driver, so as to implement flexible alarming and operation. Optionally, the distraction action of the driver may include, but is not limited to at least one of the followings: calling, drinking water, putting on or taking off sunglasses, putting on or taking off the mask, or eating food and the like. The discomfort state of the driver may include, but is not limited to at least one of the followings: sweat wiping, rubbing an eye, or yawning and the like. The non-standard behavior of the driver may include, but is not limited to at least one of the followings: smoking, stretching the hand out of the vehicle, bending over the steering wheel, putting both feet on the steering wheel, or leaving both hands away from the steering wheel and the like.

Optionally, the discomfort state of the non-driver may include, but is not limited to at least one of the followings: sweat wiping or the like. The non-standard behavior of the non-driver may include, but is not limited to at least one of the followings: smoking, stretching the hand out of the vehicle, holding the instrument with the hand, or disturbing

the driver or the like.

In the embodiments of the present disclosure, different prompt information and predetermined operations are further respectively set for the driver and the non-driver, so as to implement flexible safety control of the vehicle, for example, when the driver makes an action of leaving both hands from the steering wheel, automatically driving (for example, the corresponding first predetermined operation) is executed while sending strong prompt information (for example, the corresponding first prompt information), so as to improve the safety of vehicle driving; moreover, for the non-driver, for example, when the non-driver makes an action of sweat wiping, weak prompt information (for example, the corresponding second prompt information) is sent, and/or an operation of adjusting an in-vehicle air conditioning temperature (for example, the corresponding second predetermined operation) is executed.

The in-vehicle personnel of ordinary skill in the art may understand that: all or some steps of implementing the foregoing embodiments of the method may be achieved by a program by instructing related hardware; the foregoing program may be stored in a computer readable storage medium; when the program is executed, steps including the foregoing embodiments of the method are performed; moreover, the foregoing storage medium includes various media capable of storing program codes such as an ROM, an RAM, a magnetic disk, or an optical disk.

FIG. 4 is a schematic structural diagram of a device for recognizing a dangerous action of personnel in a vehicle (i.e., in-vehicle personnel) provided in the embodiments of the present disclosure. The apparatus of this embodiment is configured to implement the foregoing method embodiments of the present disclosure. As shown in FIG. 4, the apparatus of this embodiment includes:

- a video collection unit 41, used for obtaining at least one video stream of the in-vehicle personnel through an image capturing device,

- each video stream including information about at least one in-vehicle personnel;

- an action recognition unit 42, used for performing action recognition on the in-vehicle personnel on the basis of the video stream; and

- a danger processing unit 43, used for sending prompt information and/or executing an operation to control a vehicle in response to that an action recognition result belongs to a

predetermined dangerous action,

where the predetermined dangerous action includes at least one of the following action representations of the in-vehicle personnel: a distraction action, a discomfort state, or a non-standard behavior and the like.

On the basis of the in-vehicle personnel dangerous action recognition apparatus provided in the foregoing embodiments of the present disclosure, whether the predetermined dangerous action is made is determined by means of action recognition, and corresponding prompt/or operation is made to the predetermined dangerous action to control the vehicle, so that early detection of vehicle safety conditions is implemented to reduce the probability of dangerous situations.

In one or more optional embodiments, the action recognition unit 42 is used for detecting at least one target area included by the in-vehicle personnel in at least one frame of video image of the video stream, capturing a target image corresponding to the target area from the at least one frame of video image of the video stream according to the target area obtained through detection, and performing action recognition on the in-vehicle personnel according to the target image.

In the embodiments, the target area is recognized in the video stream, the target image corresponding to the target area is captured in the video image according to the detection result of the target area, and whether the in-vehicle personnel executes the predetermined dangerous action is recognized according to the target image. The target image captured according to the detection result of the target area may be applicable to human bodies having different areas in different video images. The embodiments of the present disclosure have a wide application range.

Optionally, the action recognition unit 42 is used for extracting a feature of the in-vehicle personnel included in the at least one frame of video image of the video stream when detecting at least one target area included by the in-vehicle personnel in the at least one frame of video image of the video frame, and extracting a target area from the at least one frame of video image frame on the basis of the feature, where the target area includes, but is not limited to, at least one of the followings: the face local area, the action interactive object, or the limb area or the like.

Optionally, the face local area includes, but is not limited to, at least one of the followings: the mouth area, the ear area, or the eye area or the like.

Optionally, the action interactive object includes, but is not limited to, at least one of the followings: a container, a cigarette, a mobile phone, food, a tool, a beverage bottle, glasses, or a mask or the like.

In one or more optional embodiments, the distraction action includes, but is not limited to, at least one of the followings: calling, drinking water, putting on or taking off sunglasses, putting on or taking off the mask, or eating food or the like; and/or,

the discomfort state includes, but is not limited to, at least one of the followings: wiping sweat, rubbing an eye, or yawning; and/or,

the non-standard behavior includes, but is not limited to, at least one of the followings: smoking, stretching a hand out of the vehicle, bending over a steering wheel, putting both feet on the steering wheel, leaving both hands away from the steering wheel, holding an instrument with a hand, or disturbing a driver or the like.

In one or more optional embodiments, the danger processing unit 43 includes:

a level determination module, used for determining a danger level of the predetermined dangerous action in response to that the action recognition result belongs to the predetermined dangerous action; and

an operation processing module, used for sending corresponding prompt information according to the danger level, and/or executing an operation corresponding to the danger level and controlling the vehicle according to the operation.

Optionally, in the embodiments of the present disclosure, when the action of the in-vehicle personnel is determined to belong to the predetermined dangerous action according to the action recognition result, danger level determination is performed on the predetermined dangerous action; optionally, the danger level of the predetermined dangerous action is determined according to a preset rule and/or correspondence, and then how to operate is determined according to the danger level. For example, operations of different degrees are performed according to the dangerous action level of the in-vehicle personnel. For example, if the dangerous action is caused by fatigue and physical discomfort of the driver, timely prompting is required, so that the driver performs adjustment and has a rest in time; and when the driver feels discomfort due to the in-vehicle environment, an adjustment of a certain degree is performed by controlling a ventilation system or an air conditioning system in the vehicle.

Optionally, the danger level includes a primary level, an intermediate level, and a high level.

The operation processing module is used for sending the prompt information in response to that the danger level is the primary level, executing the operation corresponding to the danger level and controlling the vehicle according to the operation in response to that the danger level is the intermediate level, and executing the operation corresponding to the danger level and controlling the vehicle according to the operation while sending the prompt information in response to that the danger level is the high level.

Optionally, the level determination module is used for acquiring the frequency and/or duration of occurrence of the predetermined dangerous action in the video stream, and determining the danger level of the predetermined dangerous action on the basis of the frequency and/or duration.

In one or more optional embodiments, the action recognition result includes the action duration, and the condition of belonging to the predetermined dangerous action includes: recognizing that the action duration exceeds the duration threshold.

In the embodiments of the present disclosure, by performing further abstract analysis on the dangerous action obtained by action recognition, and according to the lasting degree of the action or the priori probability of the occurrence of the dangerous situation, whether the real intention of the passenger performs the dangerous action is output. Optionally, the embodiments of the present disclosure implement measurement of the action lasting degree by means of the frequency and/or duration of the occurrence of the predetermined dangerous action in the video stream.

Optionally, the action recognition result includes the action duration, and the condition of belonging to the predetermined dangerous action includes: recognizing that the action duration exceeds the duration threshold.

Optionally, the action recognition result includes the number of actions, and the condition of belonging to the predetermined dangerous action includes: recognizing that the number of actions exceeds the number threshold.

Optionally, the action recognition result includes the action duration and the number of actions, and the condition of belonging to the predetermined dangerous action

includes: recognizing that the action duration exceeds the duration threshold, and the number of actions exceeds the number threshold.

Optionally, the in-vehicle personnel includes the driver and/or the non-driver of the vehicle.

Optionally, the danger processing unit 43 is used for sending corresponding first prompt information and/or controlling the vehicle to execute a corresponding first predetermined operation according to the predetermined dangerous action in response to that the in-vehicle personnel is the driver, and/or sending corresponding second prompt information and/or executing a corresponding second predetermined operation according to the predetermined dangerous action in response to that the in-vehicle personnel is the non-driver.

For the working process, the setting mode, and corresponding technical effect of any embodiment of the in-vehicle personnel dangerous action recognition apparatus provided by the embodiments of the present disclosure, reference may be made to the specific descriptions of the corresponding method embodiments of the present disclosure, and details are not described herein again due to space limitation.

According to still another aspect of the embodiments of the present disclosure, an electronic device is provided and includes a processor, where the processor includes the in-vehicle personnel dangerous action recognition apparatus provided according to any one of the foregoing embodiments.

According to yet another aspect of the embodiments of the present disclosure, an electronic device is provided and includes: a memory used for storing executable instructions;

and a processor, used for communicating with the memory to execute the executable instructions so as to complete operations of the method for recognizing the dangerous act of the personnel in the vehicle provided according to any one of the foregoing embodiments.

According to a further aspect of the embodiments of the present application, a computer readable storage medium is provided and is used for storing computer readable instructions, where when the instructions are executed, the operations of the method for recognizing the dangerous act of the personnel in the vehicle provided according to any one of the foregoing embodiments are executed.

According to still another aspect of the embodiments of the present disclosure, a computer program is provided and includes a computer readable code, where if the computer

readable code runs on a device, a processor in the device executes an instruction for implementing the method for recognizing the dangerous act of the personnel in the vehicle provided according to any one of the foregoing embodiments.

The embodiments of the present disclosure further provide an electronic device which, for example, is a mobile terminal, a Personal Computer (PC), a tablet computer, a server, and the like. Referring to FIG. 5 below, a schematic structural diagram of an electronic device 500, which is a terminal device or a server, suitable for implementing the embodiments of the present disclosure is shown. As shown in FIG. 5, the electronic device 500 includes one or more processors, a communication part, and the like. The one or more processors are, for example, one or more Central Processing Units (CPUs) 501 and/or one or more graphic processing unit (an acceleration unit) 513, and may execute appropriate actions and processing according to executable instructions stored in a Read-Only Memory (ROM) 502 or executable instructions loaded from a storage section 508 to a Random Access Memory (RAM) 503. The communication part 512 may include, but is not limited to, a network card. The network card may include, but is not limited to, an Infiniband (IB) network card.

The processor communicates with the ROM 502 and/or the RAM 503 to execute the executable instructions, and is connected to the communication part 512 by means of a bus 504 and communicates with other target devices by means of the communication part 512, so as to complete the operations corresponding to any of the methods provided in the embodiments of the present disclosure, for example, obtaining at least one video stream of in-vehicle personnel by using a photographing apparatus, each video stream including information about at least one in-vehicle personnel; performing action recognition on the in-vehicle personnel on the basis of the video stream; and sending prompt information and/or executing an operation to control a vehicle in response to that an action recognition result belongs to a predetermined dangerous action, where the predetermined dangerous action includes at least one of the following action representations of the in-vehicle personnel: a distraction action, a discomfort state, or a non-standard behavior.

In addition, the RAM 503 may further store various programs and data required for operations of an apparatus. The CPU 501, the ROM 502, and the RAM 503 are connected to each other by means of the bus 504. In the presence of the RAM 503, the ROM 502 is an optional module. The RAM 503 stores executable instructions, or writes the

executable instructions into the ROM 502 during running, where the executable instructions cause the CPU 501 to execute corresponding operations of the foregoing communication method. An Input/Output (I/O) interface 505 is also connected to the bus 504. The communication part 512 is integrated, or is also configured to have multiple sub-modules (for example, multiple IB network cards) linked on the bus.

The following components are connected to the I/O interface 505: an input section 506 including a keyboard, a mouse and the like; an output section 507 including a Cathode-Ray Tube (CRT), a Liquid Crystal Display (LCD), a speaker and the like; the storage section 508 including a hard disk drive and the like; and a communication section 509 of a network interface card including an LAN card, a modem and the like. The communication section 509 performs communication processing via a network such as the Internet. A drive 510 is also connected to the I/O interface 505 according to requirements. A removable medium 511 such as a magnetic disk, an optical disk, a magneto-optical disk, a semiconductor memory or the like is installed on the drive 510 according to requirements, so that a computer program read from the removable medium is installed on the storage section 508 according to requirements.

It should be noted that, the architecture shown in FIG. 5 is merely an optional implementation. During specific practice, the number and types of the components in FIG. 5 may be selected, decreased, increased, or replaced according to actual requirements. Different functional components may be configured separately or integrally or the like. For example, an acceleration unit 513 and the CPU 501 may be configured separately, or the acceleration unit 513 may be integrated on the CPU 501, and the communication part may be configured separately, and may also be configured integrally on the CPU 501 or the acceleration unit 513 or the like. These alternative implementations all fall within the scope of protection of the present disclosure.

Particularly, a process described above with reference to a flowchart according to the embodiments of the present disclosure is implemented as a computer software program. For example, the embodiments of the disclosure include a computer program product. The computer program product includes a computer program tangibly included in a machine-readable medium. The computer program includes a program code for performing a method shown in the flowchart. The program code may include instructions for correspondingly performing steps of the method provided in the embodiments of the present disclosure, for example, obtaining at least one video stream of in-vehicle personnel by using the

photographing apparatus, each video stream including information about at least one in-vehicle personnel; performing action recognition on the in-vehicle personnel on the basis the video stream; and sending the prompt information and/or executing the operation to control the vehicle in response to that the action recognition result belongs to the predetermined dangerous action, where the predetermined dangerous action includes at least one of the following action representations of the in-vehicle personnel: the distraction action, the discomfort state, or the non-standard behavior. In such embodiments, the computer program is downloaded and installed from the network by means of the communication section 509, and/or is installed from the removable medium 511. The computer program, when being executed by the CPU 501, executes the operations of the foregoing functions defined in the method of the present disclosure.

The embodiments in the specification are all described in a progressive manner, for same or similar parts in the embodiments, refer to these embodiments, and each embodiment focuses on a difference from other embodiments. The system embodiments correspond to the method embodiments substantially and therefore are only described briefly, and for the associated part, refer to the descriptions of the method embodiments.

The methods and apparatuses of the present disclosure are implemented in many manners. For example, the methods and apparatuses of the present disclosure are implemented by means of software, hardware, firmware, or any combination of software, hardware, and firmware. Unless otherwise specially stated, the foregoing sequences of steps of the methods are merely for description, and are not intended to limit the steps of the methods of the present disclosure. In addition, in some embodiments, the present disclosure may also be implemented as programs recorded in a recording medium. The programs include machine-readable instructions for implementing the methods according to the present disclosure. Therefore, the present disclosure further covers the recording medium storing the programs for executing the methods according to the present disclosure.

The descriptions of the present disclosure are provided for the purpose of examples and description, and are not intended to be exhaustive or limit the present disclosure to the disclosed form. Many modifications and changes are obvious to a person of ordinary skill in the art. The embodiments are selected and described to better describe a principle and an actual application of the present disclosure, and to make a person of ordinary skill

in the art understand the present disclosure, so as to design various embodiments with various modifications applicable to particular use.

CLAIMS

1. A method for recognizing a dangerous action of personnel in a vehicle, comprising:

obtaining at least one video stream of the personnel in the vehicle through an image capturing device, each video stream comprising information about at least one of the personnel in the vehicle;

performing action recognition on the personnel in the vehicle based on the video stream; and

responsive to that a result of the action recognition belongs to a predetermined dangerous action, performing at least one of: sending prompt information, or executing an operation to control the vehicle, wherein the predetermined dangerous action comprises at least one of the following action representations of the personnel in the vehicle: a distraction action, a discomfort state, or a non-standard behavior.

2. The method according to claim 1, wherein performing the action recognition on the personnel in the vehicle based on the video stream comprises:

detecting at least one target area comprised by the personnel in the vehicle, in at least one frame of video image of the video stream;

capturing a target image corresponding to the target area from the at least one frame of video image of the video stream according to the target area obtained through detection; and

performing action recognition on the personnel in the vehicle according to the target image.

3. The method according to claim 2, wherein detecting the at least one target area comprised by the personnel in the vehicle, in the at least one frame of video image of the video stream comprises:

extracting a feature, comprised in the at least one frame of video image of the video stream, of the personnel in the vehicle; and

extracting a target area from the at least one frame of video image based on the feature, wherein the target area comprises at least one of: a face local area, an action interactive object, or a limb area.

4. The method according to claim 3, wherein the face local area comprises at least one of: a mouth area, an ear area, or an eye area.

5. The method according to claim 3 or 4, wherein the action interactive object comprises at least one of: a container, a cigarette, a mobile phone, food, a tool, a beverage bottle, glasses, or a mask.

6. The method according to any one of claims 1 to 5, wherein the distraction action comprises at least one of: calling, drinking water, putting on or taking off sunglasses, putting on or taking off a mask, or eating food; and/or,

the discomfort state comprises at least one of: wiping sweat, rubbing an eye, or yawning; and/or,

the non-standard behavior comprises at least one of: smoking, stretching a hand out of the vehicle, bending over a steering wheel, putting both feet on the steering wheel, leaving both hands away from the steering wheel, holding an instrument with a hand, or disturbing a driver.

7. The method according to any one of claims 1 to 6, wherein responsive to that the result of the action recognition belongs to the predetermined dangerous action, performing the at least one of: sending the prompt information, or executing the operation to control the vehicle comprises:

responsive to that the result of the action recognition belongs to the predetermined dangerous action;

determining a danger level of the predetermined dangerous action; and

performing at least one of: sending corresponding prompt information according to the danger level, or executing an operation corresponding to the danger level and controlling the vehicle according to the operation.

8. The method according to claim 7, wherein the danger level comprises a primary level, an intermediate level, and a high level;

wherein performing the at least one of: sending the corresponding prompt information according to the danger level, or executing the operation corresponding to the danger level and controlling the vehicle according to the operation comprises:

sending the prompt information responsive to that the danger level is the primary level;

executing the operation corresponding to the danger level and controlling the vehicle according to the operation, responsive to that the danger level is the intermediate level; and

executing the operation corresponding to the danger level and controlling the vehicle according to the operation while sending the prompt information, responsive to that the danger level is the high level.

9. The method according to claim 7 or 8, wherein determining the danger level of the predetermined dangerous action comprises:

acquiring at least one of a frequency or a duration of occurrence of the predetermined dangerous action in the video stream, and determining the danger level of the predetermined dangerous action based on the at least one of the frequency or the duration.

10. The method according to any one of claims 1 to 9, wherein the result of the action recognition comprises a duration of an action, and a condition of belonging to the predetermined dangerous action comprises: recognizing that the duration of the action exceeds a duration threshold.

11. The method according to any one of claims 1 to 9, wherein the result of the action recognition comprises a number of times for which an action is performed, and a condition of

belonging to the predetermined dangerous action comprises: recognizing that the number of times exceeds a number threshold.

12. The method according to any one of claims 1 to 9, wherein the result of the action recognition comprises a duration of an action and a number of times for which the action is performed, and a condition of belonging to the predetermined dangerous action comprises: recognizing that the duration of the action exceeds a duration threshold, and the number of times exceeds a number threshold.

13. The method according to claims 1 to 12, wherein the personnel in the vehicle comprises at least one of a driver or a non-driver of the vehicle.

14. The method according to claim 13, wherein responsive to that the result of the action recognition belongs to the predetermined dangerous action, performing the at least one of: sending the prompt information, or executing the operation to control the vehicle comprises at least one of:

responsive to that the personnel in the vehicle is the driver, performing at least one of: sending corresponding first prompt information according to the predetermined dangerous action, or controlling the vehicle to execute a corresponding first predetermined operation according to the predetermined dangerous action; or

responsive to that the personnel in the vehicle is the non-driver, performing at least one of: sending corresponding second prompt information according to the predetermined dangerous action, or executing a corresponding second predetermined operation according to the predetermined dangerous action.

15. A device for recognizing a dangerous action of personnel in a vehicle, comprising:

a video collection unit, configured to obtain at least one video stream of the personnel in the vehicle through an image capturing device, each video stream comprising information about at least one of the personnel in the vehicle;

an action recognition unit, configured to perform action recognition on the personnel in the vehicle based on the video stream; and

a danger processing unit, configured to, responsive to that a result of the action recognition belongs to a predetermined dangerous action, perform at least one of: sending prompt information, or executing an operation to control the vehicle, wherein the predetermined dangerous action comprises at least one of the following action representations of the personnel in the vehicle: a distraction action, a discomfort state, or a non-standard behavior.

16. The device according to claim 15, wherein the action recognition unit is configured to: detect at least one target area comprised by the personnel in the vehicle in at least one frame of video image of the video stream, capture a target image corresponding to the target area from the at least one frame of video image of the video stream according to the target area obtained through detection, and perform action recognition on the personnel in the vehicle according to the target image.

17. The device according to claim 16, wherein the action recognition unit is configured to: extract a feature, comprised in the at least one frame of video image of the video stream, of the personnel in the vehicle when detecting the at least one target area comprised by the personnel in the vehicle in the at least one frame of video image of the video stream, and extract a target area from the at least one frame of video image based on the feature, wherein the target area comprises at least one of: a face local area, an action interactive object, or a limb area.

18. The device according to claim 17, wherein the face local area comprises at least one of: a mouth area, an ear area, or an eye area.

19. The device according to claim 17 or 18, wherein the action interactive object comprises at least one of: a container, a cigarette, a mobile phone, food, a tool, a beverage bottle, glasses, or a mask.

20. The device according to any one of claims 15 to 19, wherein the distraction action comprises at least one of: calling, drinking water, putting on or taking off sunglasses, putting on or taking off a mask, or eating food; and/or,

the discomfort state comprises at least one of: wiping sweat, rubbing an eye, or yawning; and/or,

the non-standard behavior comprises at least one of: smoking, stretching a hand out of the vehicle, bending over a steering wheel, putting both feet on the steering wheel, leaving both hands away from the steering wheel, holding an instrument with a hand, or disturbing a driver.

21. The device according to any one of claims 15 to 20, wherein the danger processing unit comprises:

a level determination module, configured to determine a danger level of the predetermined dangerous action responsive to that the result of the action recognition belongs to the predetermined dangerous action; and

an operation processing module, configured to perform at least one of: sending corresponding prompt information according to the danger level, or executing an operation corresponding to the danger level and controlling the vehicle according to the operation.

22. The device according to claim 21, wherein the danger level comprises a primary level, an intermediate level, and a high level; and

the operation processing module is configured to:

send prompt information responsive to that the danger level is the primary level;

execute the operation corresponding to the danger level and control the vehicle according to the operation, responsive to that the danger level is the intermediate level; and

execute the operation corresponding to the danger level and control the vehicle according to the operation while sending the prompt information, responsive to that the danger level is the high level.

23. The device according to claim 21 or 22, wherein the level determination module is configured to acquire at least one of a frequency or a duration of occurrence of the predetermined dangerous action in the video stream, and determine the danger level of the predetermined dangerous action based on the at least one of the frequency or the duration.

24. The device according to any one of claims 15 to 23, wherein the result of the action recognition comprises a duration of an action, and a condition of belonging to the predetermined dangerous action comprises: recognizing that the duration of the action exceeds a duration threshold.

25. The device according to any one of claims 15 to 24, wherein the result of the action recognition comprises a number of times for which an action is performed, and a condition of belonging to the predetermined dangerous action comprises: recognizing that the number of times exceeds a number threshold.

26. The device according to any one of claims 15 to 24, wherein the result of the action recognition comprises a duration of an action and a number of times for which the action is performed, and a condition of belonging to the predetermined dangerous action comprises: recognizing that the duration of the action exceeds a duration threshold, and the number of times exceeds a number threshold.

27. The device according to claims 15 to 26, wherein the personnel in the vehicle comprises at least one of a driver or a non-driver of the vehicle.

28. The device according to claim 27, wherein the danger processing unit is configured to perform at least one of:

responsive to that the personnel in the vehicle is the driver, performing at least one of: sending corresponding first prompt information according to the predetermined dangerous action, or controlling the vehicle to execute a corresponding first predetermined operation according to the predetermined dangerous action; or

responsive to that the in-vehicle personnel is the non-driver, performing at least one of: sending corresponding second prompt information according to the predetermined dangerous action, or executing a corresponding second predetermined operation according to the predetermined dangerous action.

29. An electronic device, comprising a processor, wherein the processor comprises the device for recognizing the dangerous action of the personnel in the vehicle according to any one of claims 15 to 28.

30. An 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 complete operations of the method for recognizing the dangerous action of the personnel in the vehicle according to any one of claims 1 to 14.

31. A computer readable storage medium configured to store computer readable instructions that, when executed, perform operations of the method for recognizing the dangerous action of the personnel in the vehicle according to any one of claims 1 to 14.

32. A computer program product, comprising computer readable codes that, when being run on a device, cause a processor in the device to execute instructions for implementing the method for recognizing the dangerous action of the personnel in the vehicle according to any one of claims 1 to 14.