

IMAGE PROCESSING METHOD AND APPARATUS

Cross-Reference To Related Application

[0001] This present application is based upon and claims priority to Chinese Application
5 No. 201810961701.X, filed on August 22, 2018, which is incorporated herein by reference in its entirety.

Technical Field

[0002] Embodiments of this specification relate to the field of image processing
10 technologies, and in particular, to an image processing method and apparatus for guiding a user to take vehicle damage assessment photos.

Technical Background

[0003] In conventional auto insurance claim settlement scenarios, insurance companies
15 need to send professional survey and damage assessment personnel to an accident site to conduct on-site survey and damage assessment, provide a vehicle repair plan and a compensation amount, take on-site photos, and keep the damage assessment photos on file for verifiers to verify the damage and cost. As the survey and damage assessment need to be conducted manually, insurance companies need a large investment for labor costs and professional knowledge training costs. In
20 terms of the experience of ordinary users, as the users need to wait for a manual surveyor to take photos on site, a damage assessor to assess damage on the repair site, and a damage verifier to verify the damage at the background during the claim settlement process, the claim settlement cycle is relatively long.

[0004] With the development of the Internet, a claim settlement solution has emerged, in
25 which a user takes vehicle damage photos on site and uploads the photos to a server, so that damage assessment and claim settlement are performed based on the vehicle damage photos according to

an algorithm or manually. However, in the solution, there are usually some requirements on the vehicle damage photos taken. For example, the user is usually required to take vehicle damage photos from far to near. In addition, there are also some requirements on the sharpness, brightness and shooting angle of the photos. With respect to such requirements, in the prior art, customer service personnel of an insurance company need to remotely guide a user to take photos for damage assessment by telephone, network or other means.

[0005] Therefore, a more effective solution for guiding users to take photos for vehicle damage assessment is needed.

10 Summary of the Invention

[0006] The embodiments of this specification are aimed at providing a more effective image processing method and apparatus to resolve the shortcomings in the prior art.

[0007] In order to achieve the foregoing objective, in one aspect, an image processing method is provided in this specification, the method being performed at a mobile device that includes a camera and comprising:

[0008] acquiring a video stream of an accident vehicle by the camera according to a user instruction;

[0009] obtaining an image of a current frame in the video stream;

[0010] determining whether the image meets a predetermined criterion by inputting the image into a predetermined classification model, wherein the predetermined classification model includes a convolutional neural network for use in the mobile device;

[0011] adding a target box and/or target segmentation information to the image by inputting the image into a target detection and segmentation model when the image meets the predetermined criterion, wherein the target box and the target segmentation information both correspond to at least one of a vehicle part and vehicle damage of the vehicle, and the target detection and segmentation model includes a convolutional neural network for use in the mobile device; and

[0012] displaying the target box and/or the target segmentation information to the user.

[0013] In an embodiment, the image processing method further comprises prompting the user correspondingly based on a classification result of the model when the image does not meet the predetermined criterion.

5 [0014] In an embodiment, in the image processing method, the predetermined classification model classifies the image based on at least one of the following conditions: whether the image is blurred, whether the image includes vehicle damage, whether the light intensity is sufficient, whether the shooting angle is skewed, and whether the shooting distance is appropriate.

[0015] In an embodiment, the image processing method further comprises presenting a
10 shooting flow to the user before the video stream of the accident vehicle is acquired by the camera.

[0016] In an embodiment, the image processing method further comprises prompting the user correspondingly based on the target box and/or the target segmentation information after the target box and/or the target segmentation information are/is added.

[0017] In an embodiment, in the image processing method, the prompting the user
15 correspondingly based on the target box and/or the target segmentation information includes prompting the user to move forward or backward based on the target box and/or the target segmentation information.

[0018] In an embodiment, in the image processing method, the prompting the user correspondingly based on the target box and/or the target segmentation information includes
20 prompting the user to shoot based on the target box and/or the target segmentation information to obtain a damage assessment photo corresponding to the image of the current frame.

[0019] In an embodiment, the image processing method further comprises uploading the damage assessment photo to a server after the damage assessment photo corresponding to the image of the current frame is obtained.

25 [0020] In an embodiment, the image processing method further comprises obtaining an association between the image of the current frame and a first image based on the video stream

after the damage assessment photo corresponding to the image of the current frame is obtained, wherein the first image is an image, shot by the user, of a frame before the current frame in the video stream.

[0021] In an embodiment, in the image processing method, the association includes at least one of the following relations between the image of the current frame and the first image: an optical flow, a mapping matrix, and a position and angle transformation relation.

[0022] In an embodiment, the image processing method further comprises uploading the association to a server after the association between the image of the current frame and the first image is obtained.

[0023] In an embodiment, in the image processing method, the predetermined classification model and the target detection and segmentation model include a shared convolutional neural network for use in the mobile device.

[0024] In another aspect, an image processing apparatus is provided in this specification, the apparatus being implemented at a mobile device that includes a camera and comprising:

[0025] an acquisition unit configured to acquire a video stream of an accident vehicle by the camera according to a user instruction;

[0026] a first obtaining unit configured to obtain an image of a current frame in the video stream;

[0027] a determination unit configured to determine whether the image meets a predetermined criterion by inputting the image into a predetermined classification model, wherein the predetermined classification model includes a convolutional neural network for use in the mobile device;

[0028] an adding unit configured to add a target box and/or target segmentation information to the image by inputting the image into a target detection and segmentation model when the image meets the predetermined criterion, wherein the target box and the target segmentation information both correspond to at least one of a vehicle part and vehicle damage of the vehicle, and the target

detection and segmentation model includes a convolutional neural network for use in the mobile device; and

[0029] a display unit configured to display the target box and/or the target segmentation information to the user.

5 [0030] In an embodiment, the image processing apparatus further comprises a first prompting unit configured to prompt the user correspondingly based on a classification result of the model when the image does not meet the predetermined criterion.

[0031] In an embodiment, the image processing apparatus further comprises a presentation unit configured to present a shooting flow to the user before the video stream of the accident vehicle
10 is acquired by the camera.

[0032] In an embodiment, the image processing apparatus further comprises a second prompting unit configured to prompt the user correspondingly based on the target box and/or the target segmentation information after the target box and/or the target segmentation information are/is added.

15 [0033] In an embodiment, in the image processing apparatus, the second prompting unit is further configured to prompt the user to move forward or backward based on the target box and/or the target segmentation information.

[0034] In an embodiment, in the image processing apparatus, the second prompting unit is further configured to prompt the user to shoot based on the target box and/or the target
20 segmentation information to obtain a damage assessment photo corresponding to the image of the current frame.

[0035] In an embodiment, the image processing apparatus further comprises a first uploading unit configured to upload the damage assessment photo to a server after the damage assessment photo corresponding to the image of the current frame is obtained.

25 [0036] In an embodiment, the image processing apparatus further comprises a second obtaining unit configured to obtain an association between the image of the current frame and a

first image based on the video stream after the damage assessment photo corresponding to the image of the current frame is obtained, wherein the first image is an image, shot by the user, of a frame before the current frame in the video stream.

[0037] In an embodiment, the image processing apparatus further comprises a second
5 uploading unit configured to upload the association to a server after the association between the image of the current frame and the first image is obtained.

[0038] In another aspect, a computing device is provided in this specification, comprising a memory and a processor, wherein executable codes are stored in the memory, and when the processor executes the executable codes, any image processing method described above is
10 implemented.

[0039] In the image processing solution according to the embodiments of this specification, blurred images, images obtained under poor light intensity, non-vehicle damage pictures and other images in an actual video stream can be effectively filtered out through an image classification algorithm with low computational load, thus helping users confirm which images are available. An
15 image detection and segmentation algorithm can help users learn which images are recognizable to the algorithm and give a forward or backward prompt to finally achieve actual availability. At the same time, an association between the photos taken by users is calculated through multiple algorithms by using characteristics of the video stream, thus providing more abundant and reliable information to a background algorithm engine and achieving more accurate and robust effects.

20

Brief Description of the Drawings

[0040] The embodiments of this specification can be clearer when described with reference to the accompanying drawings.

[0041] FIG. 1 shows a schematic diagram of an image processing system 100 according to
25 an embodiment of this specification;

[0042] FIG. 2 shows a flowchart of an image processing method according to an embodiment of this specification;

[0043] FIG. 3 shows a schematic diagram of a text prompt on a screen of a mobile phone according to a model classification result;

5 [0044] FIG. 4 schematically shows an effect diagram of the target box and/or target segmentation information displayed on the screen;

[0045] FIG. 5 shows a mapping effect of a mapping matrix; and

[0046] FIG. 6 shows an image processing apparatus 600 according to an embodiment of this specification.

10

Detailed Description

[0047] The embodiments of this specification are described in the following with reference to the accompanying drawings.

[0048] FIG. 1 shows a schematic diagram of an image processing system 100 according to an embodiment of this specification. As shown in FIG. 1, the system 100 includes a mobile device 11 and a server 12. The mobile device 11 is, for example, a mobile phone, a smart device capable of communication, or the like. The server 12 is, for example, a server used by an insurance company to process damage assessment photos. The mobile device 11 includes a camera 111 and a display 113. In addition, the mobile device 11 is provided with a mobile terminal algorithm model 112. The algorithm model 112 includes a classification model 1121 and a target detection and segmentation model 1122. The system is used to guide a user to take damage assessment photos of a vehicle and upload them to the server for processing. For example, in the scene of a vehicle accident, a damage assessor of an insurance company is not required to arrive at the scene, but a vehicle owner in the accident only needs to use the mobile device 11 to take photos according to tips of an APP for claim settlement, and then qualified damage assessment photos can be obtained and uploaded to the server 12.

15

20

25

[0049] In the process of shooting, first of all, the user opens a shooting interface in the APP for auto insurance claim settlement (i.e. issuing an instruction to the camera) through, for example, a button on the display 113 (e.g., a touch screen). On the shooting interface, the APP invokes the camera 111 to acquire a video stream, inputs the video stream into the algorithm model 112 for processing, and displays the video stream on the display 113 at the same time. The algorithm model 112 includes a classification model 1121 and a target detection and segmentation model 1122. The classification model 1121 is used to perform basic classification on input image frames, for example, whether the image frames are blurred and whether the light intensity is sufficient. The target detection and segmentation model 1122 is used to add a target box and/or segmentation information to the input image frames. The segmentation information shows target pixel-level segmentation results. After a current frame in the video stream is input into the algorithm model 112, an image of the current frame is first classified by the classification model 1121. Based on a classification result, a prompt is shown on the display 113 to prompt the user to perform a corresponding operation, such as stabilizing the camera. When the classification result of the classification model 1121 for the image of the current frame is in line with a predetermined criterion, a target box and/or segmentation information are/is added to the image by the target detection and segmentation model 1122. At the same time, the target box and/or the segmentation information are/is displayed on the frame of image shown on the display 113, and the user is prompted accordingly. For example, the user is prompted to move forward or backward. When the result of the target detection and segmentation model 1122 is in line with the predetermined criterion, a prompt is shown in the display 113 to prompt the user to take photos. The user clicks a shoot button on the display 113 to take photos, thereby obtaining qualified damage assessment photos. When the user finishes shooting, all the damage assessment photos taken will be uploaded to the server 12. In the server 12, the damage assessment photos uploaded by the user are processed through a trained damage assessment algorithm model, thereby obtaining a damage assessment result.

[0050] The structure of the system 100 shown in FIG. 1 is merely schematic, and the system according to the embodiment of this specification is not limited to the structure shown in FIG. 1. For example, the display can be a non-touch screen, and the damage assessment result may not be obtained through an algorithm in the server 12 but manually determined based on the damage assessment photos. In addition, the algorithm model 112 can also include various algorithms for obtaining an association between frames in the video stream to obtain an association between multiple damage assessment photos taken.

[0051] FIG. 2 shows a flowchart of an image processing method according to an embodiment of this specification. The method is performed at a mobile device that includes a camera. The method includes the following steps.

[0052] In step S202, a video stream of an accident vehicle is acquired by the camera according to a user instruction.

[0053] In step S204, an image of a current frame in the video stream is obtained.

[0054] In step S206, it is determined whether the image meets a predetermined criterion by inputting the image into a predetermined classification model, wherein the predetermined classification model includes a convolutional neural network for use in the mobile device.

[0055] In step S208, when the image meets the predetermined criterion, a target box and/or target segmentation information are/is added to the image by inputting the image into a target detection and segmentation model, wherein the target box and the target segmentation information both correspond to at least one of a vehicle part and vehicle damage of the vehicle, and the target detection and segmentation model includes a convolutional neural network for use in the mobile device.

[0056] In step S210, the target box and/or the target segmentation information are/is displayed to the user.

[0057] First of all, in step S202, a video stream of an accident vehicle is acquired by the camera according to a user instruction. As described above, the mobile device is, for example, a

mobile phone. For example, a user can take damage assessment photos of the accident vehicle by using an auto insurance claim settlement APP installed on his/her mobile phone. For example, the user opens a shooting interface through a camera icon in the APP and aims the camera at the accident vehicle at the same time. After the shooting interface is opened, similar to the shooting interface of the camera of the mobile phone, images acquired by the camera as well as a button for taking photos are displayed on the screen of the mobile phone. The images continuously acquired by the camera form a video stream. When the video stream is displayed on the screen of the mobile phone by the APP as described above, the video stream is also real-time input into various algorithm models deployed in the APP on the mobile phone. The camera can be configured to acquire one frame of image at a predetermined interval (for example, 125 ms) to reserve an operation time of the algorithm model.

[0058] In an embodiment, after the user enters the shooting interface, the APP first presents a general shooting flow to the user on the screen to help the user understand steps of the whole flow, and then displays the video stream acquired by the camera on the screen. For example, several processes can be demonstrated in the form of pictures, text, sounds, and the like: operating according to a prompt of the APP, taking photos according to a prompt of the APP, uploading the photos, and so on.

[0059] In step S204, an image of a current frame in the video stream is obtained. The current frame in the video stream is an image currently acquired by the camera, that is, an image currently displayed on the screen of the mobile phone. After acquiring the video stream of the vehicle by the camera, the APP can determine frames in the video stream that can be input into the model according to the model processing time. For example, each frame of image in the video stream can be obtained for model analysis, or one frame of image can be obtained every few frames in the video stream to be input into the model for analysis.

[0060] In step S206, it is determined whether the image meets a predetermined criterion by inputting the image into a predetermined classification model, wherein the predetermined classification model includes a convolutional neural network for use in the mobile device.

[0061] As described above with reference to FIG. 1, a lightweight image classification model can be deployed on the mobile phone for fast local processing of vehicle images in the video stream. For example, the predetermined classification model is a multi-task classification model trained by using a mobile model including a convolutional neural network, e.g., MobileNet v2, ShuffleNet, or SqueezeNet. For example, the classification model can be trained by using a large number of tagged vehicle damage photos, wherein the vehicle damage photo can include multiple tags related to whether the photo is blurred, whether the photo includes vehicle damage, whether the light intensity is sufficient, whether the shooting angle is skewed, whether the shooting distance is appropriate, etc., in order to carry out multi-task learning. In the above mobile model such as MobileNet v2, by optimizing a conventional two-dimensional convolution network, model parameters can be effectively reduced and the operation efficiency can be improved, so that such an algorithm can be deployed in the mobile terminal. The optimization includes, for example, superposing multiple small convolution kernels to achieve the same effect as a large convolution kernel so that parameters used are reduced exponentially; replacing general two-dimensional convolution operations with depth-wise separable convolution to reduce the number of the parameters, and so on.

[0062] The trained classification model as described above can, for example, classify the image based on at least one of the following conditions: whether the image is blurred, whether the image includes vehicle damage, whether the light intensity is sufficient, whether the shooting angle is skewed (e.g., whether the shooting angle is a tilted angle such as a high angle or an oblique angle), whether the shooting distance is appropriate, and so on. Thus, it is analyzed based on the classification result of the classification model whether the image meets predetermined basic requirements on vehicle damage assessment photos. It can be appreciated that the image

classification performed by the classification model is not limited to the above listed types, and corresponding classification can be added as required.

[0063] In the case where the image does not meet the predetermined criterion, the APP can prompt the user correspondingly based on the classification result of the model.

5 **[0064]** For example, when the classification model judges that the image is blurred, the user can be provided with the following prompt: the image is blurred; please stabilize the camera. When the classification model judges that the image does not include vehicle damage, the user can be prompted to aim the camera at the damage location. When the classification model judges that the image is not obtained under enough light intensity, the user can be prompted that the light
10 intensity is not enough. When the classification model judges that the shooting angle of the image is excessively skewed, the user can be prompted to take a photo facing the damage location. When the shooting distance is too far, the user can be prompted to approach the vehicle, and so on.

[0065] The above prompt may be in the form of a text prompt displayed on the screen, or in a voice form, or it is possible to give a prompt by displaying a text and playing a corresponding
15 voice at the same time. FIG. 3 shows a schematic diagram of a text prompt on a screen of a mobile phone according to a model classification result. As shown in FIG. 3, after the shooting interface is opened, it is obtained by detection of the classification model that the shooting distance of the image of the current frame is far. Based on the detection result, the prompt "It is too far; please approach the vehicle" is displayed at the bottom of the screen.

20 **[0066]** In step S208, when the image meets the predetermined criterion, a target box and/or target segmentation information are/is added to the image by inputting the image into a target detection and segmentation model, wherein the target box and the target segmentation information both correspond to at least one of a vehicle part and vehicle damage of the vehicle, and the target detection and segmentation model includes a convolutional neural network for use in the mobile
25 device.

[0067] As described above with reference to FIG. 1, a target detection and segmentation model can be deployed on the mobile phone to detect a part and damage of the vehicle in the video stream on the mobile phone, and display the target box and the target segmentation information on the screen of the mobile phone. The target detection and segmentation model is a lightweight
5 convolutional neural network model for use in a mobile terminal, which can be implemented by, for example, MobileNet v2+SSDLite, or MobileNet v2+DeepLab v3, MaskRCNN, etc. In an embodiment, the target detection and segmentation model and the classification model include a shared underlying convolutional neural network. In an embodiment, the target detection and segmentation model can be obtained by training with a large number of labeled (with a target box
10 or segmentation information) vehicle damage images. In the training sample, parts or damage areas of vehicles are labeled to train the target detection and segmentation model for vehicle parts and vehicle damage.

[0068] After the above image meets the basic criterion of the above vehicle damage assessment photos, by inputting the image into, for example, the trained target detection and
15 segmentation model as described above, the target detection and segmentation model can automatically detect vehicle parts and/or vehicle damage in the image, and add the target box and/or the target segmentation information to a target position.

[0069] In step S210, the target box and/or the target segmentation information are/is displayed to the user.

20 [0070] After the target box and/or the target segmentation information are/is added to the image in the model, the target box and/or the target segmentation information can be displayed, through the screen of the mobile phone, on the currently displayed image. For example, target boxes of different parts (or different damage) can be displayed in different colors, and different parts (or different damage) can be displayed in different colors to show segmentation information
25 of the different parts (or different damage). FIG. 4 schematically shows an effect diagram of the target box and/or target segmentation information displayed on the screen. As shown in FIG. 4,

different gray scales represent different colors, for example, target boxes in different colors are added to the left front wheel and the left front lamp respectively. For another example, the left front fender and the left front door are respectively filled with different colors (the different colors are the target segmentation information). In the embodiment of this specification, target detection and target segmentation can be carried out for each vehicle part (or damage) at the same time; for another example, target detection can be carried out for each vehicle part, target segmentation can be carried out for each damage, etc., in order to distinguish the parts and damage during display.

[0071] The target detection and segmentation model of the mobile phone is used to directly detect, on the mobile phone, vehicle damage and parts in the image of the current frame, and visualize them on the screen of the mobile phone to help the user establish intuitive feelings, and know a rough detection result that the currently taken photo is going to have on the server side, images from which part information can be obtained accurately according to an algorithm, and images from which damage information can be obtained accurately according to the algorithm. According to the target box and the target segmentation information displayed on the screen, the user can appropriately adjust the shooting. For example, the user can move the position of the mobile phone so that the target box corresponding to the damage is displayed on the screen, thereby shooting damage assessment photos.

[0072] In an embodiment, after the target box and/or the target segmentation information are/is added to the image, the user is prompted correspondingly based on the target box and/or the target segmentation information. For example, the user is prompted to move forward or backward based on whether the image includes the target box and/or the target segmentation information of the damage, as well as the detected parts, the amount of damage and other information. When the APP determines, based on a processing result of the target detection and segmentation model for the image, that the image meets predetermined requirements on damage assessment photos, the user can be prompted to take a photo. Thus, by clicking the shoot button on the shooting interface, the user can obtain a vehicle damage assessment photo corresponding to the image of the current

frame. After the user clicks the shoot button, the APP can save the damage assessment photo taken by the user into a phone album of the mobile phone or into the APP. After obtaining the damage assessment photo, the APP can upload the damage assessment photo to the server automatically or upload the photo to the server based on an operation of the user. On the shooting interface, the user can move the mobile phone to aim at multiple damage positions on the vehicle and take photos multiple times based on the tips of the APP as mentioned above to obtain multiple damage assessment photos. The APP can upload the multiple damage assessment photos together after the user shoots all the damage assessment photos or upload a single damage assessment photo each time the damage assessment photo is shot.

10 **[0073]** In an embodiment, the method shown in FIG. 2 further includes obtaining an association between the image of the current frame and a first image based on the video stream after the damage assessment photo corresponding to the image of the current frame is obtained, wherein the first image is an image, shot by the user, of a frame before the current frame in the video stream. In other words, the APP can obtain, based on the video stream, an association
15 between multiple damage assessment photos taken based on the video stream. As described above with reference to FIG. 1, various algorithms for obtaining the association can be deployed on the mobile phone.

[0074] As described above, the user can take photos multiple times on the shooting interface to obtain multiple damage assessment photos. As the video stream includes rich
20 information, it can be used to establish an association among various damage assessment photos. For example, an association between the damage assessment photo corresponding to the image of the current frame and each damage assessment photo taken previously can be obtained.

[0075] In an embodiment, the association includes dynamic information between frames constructed according to multiple FlowNets. For example, a first damage assessment image (photo) corresponding to the image of the current frame is processed by FlowNet1, a second damage
25 assessment image (for example, an image frame before the current frame) is processed by

FlowNet2, and an optical flow between the first damage assessment image and the second damage assessment image can be obtained by merging the output of FlowNet1 and FlowNet2.

[0076] In an embodiment, the association includes a mapping matrix between frames. In an algorithm for obtaining the mapping matrix between frames, an image gradient and an image difference between an image 1 and an image 2 are first calculated, and then a mapping matrix between the image 1 and the image 2 is obtained through the least square optimization and Cholesky decomposition. For example, the image 1 and the image 2 are the first damage assessment image and the second damage assessment image. FIG. 5 shows a mapping effect of a mapping matrix. As shown in FIG. 5, (a) is the image 1, (c) is the image 2, (b) is an image 1' obtained by transforming the image 1 by using a mapping matrix, and the mapping matrix is a mapping matrix from the image 1 to the image 2. Thus, the image 1' transformed through the mapping matrix is basically consistent with the image 2.

[0077] In an embodiment, the association includes a shooting position and angle transformation relation between frames calculated by using a SLAM technology.

[0078] After the association between multiple damage assessment images (photos) taken is obtained as described above, the association can be uploaded to the server as relevant information. The server assesses damage based on the damage assessment images and the association between the damage assessment images, so that a more accurate and reliable damage assessment result can be obtained.

[0079] FIG. 6 shows an image processing apparatus 600 according to an embodiment of this specification. The apparatus is performed at a mobile device that includes a camera. The apparatus includes:

[0080] an acquisition unit 601 configured to acquire a video stream of an accident vehicle by the camera according to a user instruction;

[0081] a first obtaining unit 602 configured to obtain an image of a current frame in the video stream;

[0082] a determination unit 603 configured to determine whether the image meets a predetermined criterion by inputting the image into a predetermined classification model, wherein the predetermined classification model includes a convolutional neural network for use in the mobile device;

5 [0083] an adding unit 604 configured to add a target box and/or target segmentation information to the image by inputting the image into a target detection and segmentation model when the image meets the predetermined criterion, wherein the target box and the target segmentation information both correspond to at least one of vehicle parts and vehicle damage of the vehicle, and the target detection and segmentation model includes a convolutional neural
10 network for use in the mobile device; and

[0084] a display unit 605 configured to display the target box and/or the target segmentation information to the user.

[0085] In an embodiment, the image processing apparatus 600 further includes a first prompting unit 606 configured to prompt the user correspondingly based on a classification result
15 of the model when the image does not meet the predetermined criterion.

[0086] In an embodiment, the image processing apparatus 600 further includes a presentation unit 607 configured to present a shooting flow to the user before the video stream of the accident vehicle is acquired by the camera.

[0087] In an embodiment, the image processing apparatus 600 further includes a second
20 prompting unit 608 configured to prompt the user correspondingly based on the target box and/or the target segmentation information after the target box and/or the target segmentation information are/is added.

[0088] In an embodiment, in the image processing apparatus 600, the second prompting unit 608 is further configured to prompt the user to move forward or backward based on the target
25 box and/or the target segmentation information.

[0089] In an embodiment, in the image processing apparatus, the second prompting unit 608 is further configured to prompt the user to shoot based on the target box and/or the target segmentation information to obtain a damage assessment photo corresponding to the image of the current frame.

5 [0090] In an embodiment, the image processing apparatus 600 further includes a first uploading unit 609 configured to upload the damage assessment photo to a server after the damage assessment photo corresponding to the image of the current frame is obtained.

[0091] In an embodiment, the image processing apparatus 600 further includes a second obtaining unit 610 configured to obtain an association between the image of the current frame and
10 a first image based on the video stream after the damage assessment photo corresponding to the image of the current frame is obtained, wherein the first image is an image, shot by the user, of a frame before the current frame in the video stream.

[0092] In an embodiment, the image processing apparatus further includes a second uploading unit 611 configured to upload the association to a server after the association between
15 the image of the current frame and the first image is obtained.

[0093] In another aspect, a computing device is provided in this specification, including a memory and a processor, wherein executable codes are stored in the memory, and when the processor executes the executable codes, the image processing method according to any of the above items is implemented.

20 [0094] In the image processing solution according to the embodiments of this specification, blurred images, images obtained under poor light intensity, non-vehicle damage pictures and other images in an actual video stream can be effectively filtered out through an image classification algorithm with low computational load, thus helping users confirm which images are available. An image detection and segmentation algorithm can help users learn which images are recognizable
25 to the algorithm and give a forward or backward prompt to finally achieve actual availability. At the same time, an association between the photos taken by users is calculated through multiple

algorithms by using characteristics of the video stream, thus providing more abundant and reliable information to a background algorithm engine and achieving more accurate and robust effects.

[0095] Various embodiments in this specification are described in a progressive manner. The same or similar parts between the embodiments may be referenced to one another. Each
5 embodiment focuses on a part that is different from other embodiments. Particularly, the system embodiment is described in a relatively simple manner because it is similar to the method embodiment, and for related parts, reference can be made to the partial description in the method embodiment.

[0096] Specific embodiments of this specification are described in the foregoing. Other
10 embodiments fall within the scope of the appended claims. Under some circumstances, the actions or steps described in the claims may be performed in a sequence different from that in the embodiments and still can achieve a desired result. In addition, the processes depicted in the accompanying drawings are not necessarily required to follow the specific sequence shown or a consecutive sequence to achieve the desired result. Multitask processing and parallel processing
15 are also possible or may be advantageous in some implementations.

[0097] Those of ordinary skill in the art should be further aware that units and algorithm steps of the examples described in combination with the embodiments disclosed herein can be implemented by electronic hardware, computer software or a combination thereof. In order to clearly describe interchangeability between the hardware and the software, compositions and steps
20 of the examples have been generally described according to functions in the above description. Whether the functions are executed by hardware or software depends on specific applications and design constraints of the technical solution. Those of ordinary skill in the art can implement the described functions by using different methods for each specific application, but such an implementation should not be considered as going beyond the scope of this application.

25 [0098] Steps of the method or algorithm described in combination with the embodiments disclosed herein can be implemented by hardware, a processor-executed software module, or a

combination thereof. The software module may be built in a RAM, a memory, a read-only memory (ROM), an electrically programmable ROM, an electrically erasable programmable ROM, a register, a hard disk, a removable disk, a CD-ROM, or any other forms of storage mediums well-known in the technical field.

5 **[0099]** The specific implementations above further describe the objectives, the technical solutions and beneficial effects of the present invention in detail. It should be understood that the above descriptions are merely specific implementations of the present invention and are not intended to limit the protection scope of the present invention. Any modification, equivalent replacement and improvement made without departing from the spirit and principle of the present
10 invention shall all fall within the protection scope of the present invention.

CLAIMS

1. An image processing method, which is performed at a mobile device that comprises a camera, the method comprising:

acquiring a video stream of an accident vehicle by the camera according to a user instruction;

5 obtaining an image of a current frame in the video stream;

determining whether the image meets a predetermined criterion by inputting the image into a predetermined classification model, wherein the predetermined classification model comprises a convolutional neural network for use in the mobile device;

10 adding a target box and/or target segmentation information to the image by inputting the image into a target detection and segmentation model when the image meets the predetermined criterion, wherein the target box and the target segmentation information both correspond to at least one of vehicle parts and vehicle damage of the vehicle, and the target detection and segmentation model comprises a convolutional neural network for use in the mobile device; and

displaying the target box and/or the target segmentation information to the user.

15 2. The image processing method of claim 1, further comprising: prompting the user correspondingly based on a classification result of the model when the image does not meet the predetermined criterion.

3. The image processing method of claim 1, wherein the predetermined classification model classifies the image based on at least one of the following conditions: whether the image is blurred, whether the image comprises vehicle damage, whether the light intensity is sufficient, whether the shooting angle is skewed, and whether the shooting distance is appropriate.

4. The image processing method of claim 1, further comprising presenting a shooting flow to the user before the video stream of the accident vehicle is acquired by the camera.

5. The image processing method of claim 1, further comprising prompting the user correspondingly based on the target box and/or the target segmentation information after the target box and/or the target segmentation information are/is added.

6. The image processing method of claim 5, wherein the prompting the user correspondingly based on the target box and/or the target segmentation information comprises prompting the user to move forward or backward based on the target box and/or the target segmentation information.

7. The image processing method of claim 5, wherein the prompting the user correspondingly based on the target box and/or the target segmentation information comprises prompting the user to shoot based on the target box and/or the target segmentation information to obtain a damage assessment photo corresponding to the image of the current frame.

8. The image processing method of claim 7, further comprising uploading the damage assessment photo to a server after the damage assessment photo corresponding to the image of the current frame is obtained.

9. The image processing method of claim 7, further comprising obtaining an association between the image of the current frame and a first image based on the video stream after the damage assessment photo corresponding to the image of the current frame is obtained, wherein the first image is an image, shot by the user, of a frame before the current frame in the video stream.

10. The image processing method of claim 9, wherein the association comprises at least one of the following relations between the image of the current frame and the first image: an optical flow, a mapping matrix, and a position and angle transformation relation.

11. The image processing method of claim 9, further comprising uploading the association to a server after the association between the image of the current frame and the first image is obtained.

12. The image processing method of claim 1, wherein the predetermined classification model and the target detection and segmentation model comprise a shared convolutional neural network for use in the mobile device.

13. An image processing apparatus, which is implemented at a mobile device that comprises a camera, the apparatus comprising:

an acquisition unit configured to acquire a video stream of an accident vehicle by the camera according to a user instruction;

a first obtaining unit configured to obtain an image of a current frame in the video stream;

a determination unit configured to determine whether the image meets a predetermined criterion by inputting the image into a predetermined classification model, wherein the predetermined classification model comprises a convolutional neural network for use in the mobile
5 device;

an adding unit configured to add a target box and/or target segmentation information to the image by inputting the image into a target detection and segmentation model when the image meets the predetermined criterion, wherein the target box and the target segmentation information both correspond to at least one of vehicle parts and vehicle damage of the vehicle, and the target
10 detection and segmentation model comprises a convolutional neural network for use in the mobile device; and

a display unit configured to display the target box and/or the target segmentation information to the user.

14. The image processing apparatus of claim 13, further comprising a first prompting unit
15 configured to prompt the user correspondingly based on a classification result of the model when the image does not meet the predetermined criterion.

15. The image processing apparatus of claim 13, wherein the predetermined classification model classifies the image based on at least one of the following conditions: whether the image is blurred, whether the image comprises vehicle damage, whether the light intensity is sufficient,
20 whether the shooting angle is skewed, and whether the shooting distance is appropriate.

16. The image processing apparatus of claim 13, further comprising a presentation unit configured to present a shooting flow to the user before the video stream of the accident vehicle is acquired by the camera.

17. The image processing apparatus of claim 13, further comprising a second prompting unit
25 configured to prompt the user correspondingly based on the target box and/or the target

segmentation information after the target box and/or the target segmentation information are/is added.

18. The image processing apparatus of claim 17, wherein the second prompting unit is further configured to prompt the user to move forward or backward based on the target box and/or the target segmentation information.

19. The image processing apparatus of claim 17, wherein the second prompting unit is further configured to prompt the user to shoot based on the target box and/or the target segmentation information to obtain a damage assessment photo corresponding to the image of the current frame.

20. The image processing apparatus of claim 19, further comprising a first uploading unit configured to upload the damage assessment photo to a server after the damage assessment photo corresponding to the image of the current frame is obtained.

21. The image processing apparatus of claim 19, further comprising a second obtaining unit configured to obtain an association between the image of the current frame and a first image based on the video stream after the damage assessment photo corresponding to the image of the current frame is obtained, wherein the first image is an image, shot by the user, of a frame before the current frame in the video stream.

22. The image processing apparatus of claim 21, wherein the association comprises at least one of the following relations between the image of the current frame and the first image: an optical flow, a mapping matrix, and a position and angle transformation relation.

23. The image processing apparatus of claim 21, further comprising a second uploading unit configured to upload the association to a server after the association between the image of the current frame and the first image is obtained.

24. The image processing apparatus of claim 13, wherein the predetermined classification model and the target detection and segmentation model comprise a shared convolutional neural network for use in the mobile device.

25. A computing device, comprising a memory and a processor, wherein executable codes are stored in the memory, and when the processor executes the executable codes, the method of any of claims 1 to 12 is implemented.