

IMAGE DETECTION METHOD AND RELATED APPARATUS, DEVICE, STORAGE MEDIUM AND COMPUTER PROGRAM

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application is based on and claims priority to Chinese patent application No. 202011002322.1, filed on September 22, 2020 and entitled “Image Detection Method and Related Apparatus, Device and Storage medium”, the disclosure of which is hereby incorporated by reference in its entirety.

TECHNICAL FIELD

[0002] The disclosure relates to the technical field of image processing, and in particular to an image detection method, a related apparatus, a device, a storage medium and a computer program.

BACKGROUND

[0003] At present, image processing, especially detection and recognition of an object in an image, has been widely used in various usage scenarios. Taking human face as an example, face detection and recognition in the image have been widely used in finance, border inspection, government, aerospace, electric power, factories, education, medical treatment and other fields. In the related art, a camera is adopted to collect an image or a video stream containing a human face, and the human face is automatically detected in the image, and then face recognition is performed on the detected human face, and corresponding processing is performed based on the recognition result.

SUMMARY

[0004] Embodiments of the disclosure at least provide an image detection method and a related apparatus, a device, a storage medium, and a computer program.

[0005] The embodiments of the disclosure provide an image detection method, which may include that: a first image containing a to-be-detected target is acquired; the first image is detected to obtain a detection result of the first image, the detection result including whether the to-be-detected target in the first image is shaded by a preset object; and a preset operation matched with the detection result is executed.

[0006] Therefore, the first image containing the to-be-detected target is detected so as to obtain the detection result of whether the to-be-detected target is shaded. Then, the preset operation matched with the detection result is executed. It can be judged whether the to-be-detected target is shaded and then the subsequent preset operation matched with the detection result can be performed, which achieves flexible processing based on the shading state of the to-be-detected target in the image.

[0007] In some embodiments, the detection result is obtained by detecting the first image using a neural network.

[0008] Therefore, the neural network trained in advance is used for detection, which makes the detection result more accurate and the detection speed faster.

[0009] In some embodiments, the operation of executing the preset operation matched with the detection result may include that: in the case that the to-be-detected target is not shaded by the preset object, a first reminder is issued. The first reminder is used to give a reminder of shading the to-be-detected target using the preset object.

[0010] Therefore, the first reminder is issued when the to-be-detected target is not shaded, so that the situation that the to-be-detected target is not shaded by the preset object is reminded timely, and then a reminded person can take corresponding measures in time.

[0011] In some embodiments, the detection result may also include whether the shading mode in which the to-be-detected target is shaded by the preset object is a preset shading mode. The operation of executing the preset operation matched with the

detection result may include that: a second reminder is issued when the to-be-detected target is shaded by the preset object but the shading mode does not belong to the preset shading mode. The second reminder is used for giving a reminder of adjusting the shading mode of the preset object.

[0012] Therefore, when the shading mode is incorrect, the second reminder is issued to facilitate timely adjustment of the shading mode of the to-be-detected target.

[0013] In some embodiments, the operation of executing the preset operation matched with the detection result may include that: in the case that the to-be-detected target is shaded by the preset object, at least a first feature of an unshaded part of the to-be-detected target is extracted from the first image as a to-be-recognized feature of the to-be-detected target; and the to-be-detected target is recognized using the to-be-recognized feature, and a recognition result is obtained.

[0014] Therefore, when the to-be-detected target is shaded by the preset object, the feature of the unshaded part is extracted for recognition, which realizes recognition based on a local feature of the to-be-detected target. Moreover, since the local feature is unshaded, it can represent the to-be-detected target, thus ensuring the recognition accuracy to a certain extent.

[0015] In some embodiments, the operation of extracting at least the first feature of the unshaded part of the to-be-detected target from the first image as the to-be-recognized feature of the to-be-detected target may include that: the first feature of the unshaded part of the to-be-detected target is extracted from the first image, and a second feature of a shaded part of the to-be-detected target is acquired; and the first feature and the second feature are taken as to-be-recognized features of the to-be-detected target.

[0016] Therefore, in addition to the adoption of the feature of the unshaded part of the to-be-detected target, the feature of the shaded part is adopted, and thus the feature richness of the to-be-detected target may be improved.

[0017] In some embodiments, the operation of acquiring the second feature of the shaded part of the to-be-detected target may include that: the feature of the shaded part is extracted from the first image as the second feature; or, a preset feature of the shaded part

is acquired as the second feature. The preset feature may include a feature obtained based on at least one reference feature. Each reference feature is extracted from a region, corresponding to the shaded part, in a reference target without the shaded part.

[0018] Therefore, as for a feature determination manner of the shaded part, the feature of the shaded part may be directly extracted. Since the feature of the shaded part can be different with different to-be-detected targets to a certain extent, this manner can improve the recognition accuracy. It may also acquire the preset feature as the feature of the shaded part. This manner does not need to perform feature extraction on the shaded part, thus reducing the loss of processing resources and improving the processing efficiency.

[0019] In some embodiments, the operation of using the to-be-recognized feature to recognize the to-be-detected target and obtaining the recognition result may include at least one of the following operations. In the case that one preset target is included, a first similarity between the to-be-recognized feature and a pre-stored feature of the preset target is acquired, and in the case that the first similarity meets a first preset condition, the recognition result is determined to include that the to-be-detected target passes identity authentication. In the case that multiple preset targets are included, second similarities between the to-be-recognized feature and the pre-stored feature of each preset target are respectively acquired, and the recognition result is determined to include that the identity of the to-be-detected target is determined as the identity of the preset target corresponding to the second similarity meeting a second preset condition.

[0020] Therefore, the first similarity with the pre-stored feature of a specific preset target is calculated, or the similarities with the pre-stored features of multiple preset targets are calculated, so that the to-be-detected target can be compared with a certain specific preset target or the preset targets in a certain database according to actual scenario requirements.

[0021] In some embodiments, the method may include at least one of the following operations. The first preset condition includes that the first similarity is greater than a first similarity threshold. The second preset condition includes that the second similarity is greater than a second similarity threshold.

[0022] Therefore, the first similarity threshold is set in different scenarios, so that the recognition result is more accurate.

[0023] In some embodiments, the method may include at least one of the following items: the first similarity threshold in the case that the to-be-recognized feature includes the second feature of the shaded part of the to-be-detected target is smaller than the first similarity threshold in the case that the to-be-recognized feature does not include the second feature, and the second similarity threshold in the case that the to-be-recognized feature includes the second feature is smaller than the second similarity threshold in the case that the to-be-recognized feature does not include the second feature.

[0024] Therefore, if the to-be-recognized feature includes the second feature, the second feature may be different from a real feature at a key point of the shaded part of the to-be-detected target. Hence, in such a case the recognition accuracy can be improved by appropriately reducing the similarity threshold.

[0025] In some embodiments, before the first similarity between the to-be-recognized feature and the pre-stored feature of the preset target is acquired, the method may further include that: an account is registered for a user in response to an account registration request; a second image meeting a preset quality requirement is determined from at least one of the second images shot by the user, and the feature of a preset part of the user is extracted from the determined second image; and the feature of the preset part is associated with the account, and the feature of the preset part is stored as the pre-stored feature of the preset target.

[0026] Therefore, the second image meeting the quality requirement is determined firstly to extract the feature of the preset part, so as to make the extracted feature more accurate.

[0027] In some embodiments, in the case that the to-be-detected target is shaded by the preset object, before at least the first feature of the unshaded part of the to-be-detected target is extracted from the first image, the method may further include at least one of the following steps: the first image meeting the preset quality requirement is determined as the first image for subsequent feature extraction from multiple first images

containing the to-be-detected target; the first image for subsequent feature extraction is preprocessed; living body detection is performed on the first image for subsequent feature extraction, and it is determined to execute extraction of at least the first feature of the unshaded part of the to-be-detected target from the first image and subsequent steps thereof in the case that a living body detection result indicates that the to-be-detected target is a living body.

[0028] Therefore, before feature extraction, preprocessing is performed first, so that the extracted feature is more accurate. The to-be-detected target is recognized only when the to-be-detected target is a living body, so that the security of recognition is enhanced, and prosthesis attack may be prevented to a certain extent.

[0029] In some embodiments, the operation of determining the first image meeting the preset quality requirement as the first image for subsequent feature extraction from the multiple first images containing the to-be-detected target may include that: for each first image, the quality score of the first image is obtained based on a quality factor of the first image. The quality factor of the first image includes at least one of the following items: pose information of the to-be-detected target relative to a shooting device, parameter information for reflecting the size of the to-be-detected target in the first image, and brightness information of the first image. Based on the quality scores, the first image meeting the preset quality requirement is determined as the first image for subsequent feature extraction. The quality score of the selected first image is higher than those of other first images.

[0030] Therefore, feature extraction is performed by determining the image of which the quality score meets the requirement, so that the extracted feature can better represent the to-be-detected target.

[0031] In some embodiments, the operation of preprocessing the first image for subsequent feature extraction may include that: in the case that the first image includes a plurality of to-be-detected targets, it is determined that the to-be-detected target meeting a preset extraction requirement is in a target region in the first image, and an image part in the first image except the target region is removed; and/or, it is detected that the inclination angle of the to-be-detected target in the first image is greater than a preset

angle, and the first image is rotated until the inclination angle of the to-be-detected target is less than or equal to the preset angle.

[0032] Therefore, when there are multiple to-be-detected targets in the first image, only the to-be-detected target that meets the preset extraction requirement is determined, and the to-be-detected targets that do not meet the requirement are discarded, thus reducing the influence of the to-be-detected targets that do not meet the requirement on the recognition result. Secondly, in the case that the to-be-detected target in the first image has an inclination angle, the to-be-detected target is straightened, thus reducing the influence caused by inclination of the to-be-detected target.

[0033] In some embodiments, the preset extraction requirement may include that the area of the corresponding region of the to-be-detected target is larger than that of the corresponding regions of other to-be-detected targets. Other to-be-detected targets include targets other than the to-be-detected target.

[0034] Therefore, since the larger the area of the to-be-detected target is, the more accurate the extracted feature is, the to-be-detected target with a relatively large area is selected so that a to-be-detected result is more accurate.

[0035] In some embodiments, the to-be-detected target may include a human face. The preset object may include a mask.

[0036] Therefore, it is judged whether the human face is wearing the mask and a corresponding operation is executed. For example, if the human face is not wearing the mask or the way of wearing the mask is inaccurate, a corresponding reminder may be issued, so that the user can make adjustments timely. If the human face wears the mask, the human face is recognized.

[0037] The embodiments of the disclosure provide an image detection apparatus, which may include an image acquisition module, configured to acquire a first image containing a to-be-detected target; a target detection module, configured to detect the first image to obtain a detection result of the first image, the detection result including whether the to-be-detected target in the first image is shaded by a preset object; and an operation execution module, configured to execute a preset operation matched with the

detection result.

[0038] The embodiments of the disclosure provide an electronic device, which may include a memory and a processor. The processor is configured to execute program instructions stored in the memory to implement the image detection method.

[0039] The embodiments of the disclosure provide a computer readable storage medium, in which program instructions is stored. The program instruction, when executed by a processor, implement the image detection method.

[0040] The embodiments of the disclosure provide a computer program, which may include a computer readable code. When running in an electronic device, the computer readable code is executed by a processor in the electronic device to implement the image detection method.

[0041] The embodiments of the disclosure provide an image detection method, related apparatus, a device, a storage medium, and a computer program. A first image containing a to-be-detected target is detected so as to obtain whether the to-be-detected target is shaded. Then, a preset operation matched with a detection result is executed. It can be judged whether the to-be-detected target is shaded and then the subsequent preset operation matched with the detection result can be performed, which achieves flexible processing based on the shading state of the to-be-detected target in the image.

[0042] It is to be understood that the above general descriptions and the following detailed descriptions are only exemplary and explanatory, and do not limit the disclosure.

BRIEF DESCRIPTION OF THE DRAWINGS

[0043] The drawings herein are incorporated into and constitute a part of the specification, which illustrate embodiments in accordance with the disclosure and together with the specification are used to explain the technical solutions of the disclosure.

[0044] FIG. 1 is a schematic flowchart of an embodiment of an image detection

method according to an embodiment of the disclosure.

[0045] FIG. 2 is a schematic diagram of a first image in an embodiment of an image detection method according to an embodiment of the disclosure.

[0046] FIG. 3 is a schematic diagram of a preprocessed first image in an embodiment of an image detection method according to an embodiment of the disclosure.

[0047] FIG. 4 is a schematic structural diagram of an embodiment of an image detection apparatus according to an embodiment of the disclosure.

[0048] FIG. 5 is a schematic structural diagram of an embodiment of an electronic device according to an embodiment of the disclosure.

[0049] FIG. 6 is a schematic structural diagram of an embodiment of a computer readable storage medium according to an embodiment of the disclosure.

DETAILED DESCRIPTION

[0050] Solutions of the embodiments of the present application will now be described in detail in combination with the accompanying drawings.

[0051] In the following descriptions, specific details such as a specific system structure, an interface, and a technology are set forth for purposes of illustration rather than limitation, so as to provide a thorough understanding of the disclosure.

[0052] Herein, term "and/or" is only an association relationship describing associated objects and represents that three relationships may exist. For example, A and/or B may represent three conditions: i.e., independent existence of A, existence of both A and B, and independent existence of B. In addition, the character "/" herein generally indicates that the contextual objects are an "or" relationship. Furthermore, "multiple" herein means two or more than two. In addition, the term "at least one" herein indicates any one of multiple kinds or any combination of at least two of the multiple kinds, for example, including at least one of A, B and C, which may indicate including any one or more

elements selected from a set consisting of A, B and C.

[0053] The image detection method in the embodiment of the disclosure may be applied to a device with image processing capability. In addition, the device may have an image acquisition or video acquisition function. For example, the device may include a component such as a camera for acquiring an image or a video. Or, the device may acquire a video stream or image as required from other devices or access the video stream or image as required from storage resources of other devices by means of data transmission or data interaction with other devices. Herein, other devices have the image acquisition or video acquisition function, and are in communication connection with the device. For example, the device may perform data transmission or data interaction with other devices by means of Bluetooth, wireless network, etc. In the embodiments of the disclosure, the communication mode between them is not limited, and can include but not limited to the above-mentioned situations. In one implementation mode, the device may include a mobile phone, a tablet computer, an interactive screen, etc., which is not limited here.

[0054] Referring to FIG. 1, FIG. 1 is a schematic flowchart of an embodiment of an image detection method according to an embodiment of the disclosure. The image detection method, executed by an electronic device, may include the following steps.

[0055] At S11, a first image containing a to-be-detected target is acquired.

[0056] The first image containing the to-be-detected target may be an initial image containing a target object acquired by directly calling a camera of the device. Of course, it may be an image acquired from other devices, or may be an image subjected to frame selection, brightness adjustment, resolution adjustment, etc. The to-be-detected target may include a human face, a face or limb of another animal, and so on. Therefore, the manner of acquiring the first image containing the to-be-detected target is not limited. Herein, other devices refer to devices that can only be operated using different central processing units, respectively.

[0057] At S12, the first image is detected to obtain a detection result of the first image. The detection result includes whether the to-be-detected target in the first image is

shaded by a preset object.

[0058] The preset object refers to any object that can shade the to-be-detected target, such as a mask, a scarf, glasses, arms, paper or other visible objects.

[0059] The first image is detected. In the detection process, it is necessary to detect whether the first image contains a to-be-detected image. If there is the to-be-detected target, it is judged whether the to-be-detected target is shaded by the preset object. The manner of judging whether the to-be-detected target is shaded by the preset object may be training a shading detection model before detecting the first image. Whether the to-be-detected target in the first image is shaded by the preset object can be known by inputting the first image into the shading detection model (for example, a neural network with a shading detection function). Of course, in some embodiments, the manner of judging whether the to-be-detected target is shaded by the preset object may be judging whether a preset detection position in the to-be-detected target is shaded and whether the object which shades the preset detection position meets a condition of the preset object. The features of a shaded object and a preset shaded object may be extracted, and the similarity is judged so as to obtain the detection result including whether the to-be-detected target is shaded nor not.

[0060] At S13, the preset operation matched with the detection result is executed.

[0061] The detection result may be that the to-be-detected target is shaded but not by the preset object, the to-be-detected target is shaded by the preset object but the shading mode is not a preset mode, the to-be-detected target is shaded by the preset object and the shading mode is the same as the preset mode, or the to-be-detected target is not shaded at all. In the embodiment of the disclosure, it is considered whether the to-be-detected target in the first image is shaded by the preset object. Of course, in some embodiments, the preset object may be set as any object. That is, the shading detection model is adopted, and as long as it is determined that the to-be-detected target is shaded, the corresponding preset operation is executed. The preset operation may be any operation related to target detection, such as recognition.

[0062] According to the above solution, the first image containing the to-be-detected

target is detected to find out whether the to-be-detected target is shaded or not. Then, the preset operation matched with the detection result is executed. It can be judged whether the to-be-detected target is shaded and then the subsequent preset operation matched with the detection result can be performed, which achieves flexible processing based on the shading state of the to-be-detected target in the image.

[0063] In some embodiments, the detection result is obtained by detecting the first image using the neural network. Before the first image is detected, a preset object shading model is trained, so that the trained preset object shading model can detect whether the to-be-detected target in the first image is shaded by the preset object. The preset object may include one or more, for example, two or three different objects. When there are multiple preset objects, it indicates that one preset object shading model can not only judge whether the to-be-detected target is shaded by the preset objects, but also detect by which preset object the to-be-detected target is shaded. The to-be-detected target may be a human face. The preset object may be a mask. Correspondingly, the preset object shading model is a mask detection model. The mask detection model may detect whether the to-be-detected target is wearing the mask. Of course, in some embodiments, it may also detect whether the mask wearing mode is correct or not while detecting that the to-be-detected target is wearing the mask. The neural network trained in advance is used for detection, which makes the detection result more accurate and the detection speed faster.

[0064] In the case that the to-be-detected target is not shaded by the preset object, a first reminder is issued. The first reminder is used to give a reminder of shading the to-be-detected target using the preset object. The first reminder may have multiple reminding modes, including a face frame selecting mode. If it is detected that the to-be-detected target is not shaded by the preset object, a face region will be selected in the form of face frame. At this time, the face frame may have a warning color, such as red or yellow. The first reminder may also be a combination of the face frame and suggestive words, such as “you are not wearing a mask”, “please wear a mask”. Of course, it may also be a voice reminder, a flashing reminder light or other forms. Of course, these forms may be used together or separately, which is not specified here. For example, when the mask detection model detects that the human face is not wearing the mask, the first

reminder is issued to remind the human face to wear the mask to shade the nose and mouth of the human face. The first reminder is issued when the to-be-detected target is not shaded, so that the situation that the to-be-detected target is not shaded by the preset object is timely reminded, and then a reminded person can take corresponding measures in time.

[0065] In some embodiments, the detection result further includes whether the shading mode in which the to-be-detected target is shaded by the preset object is the preset shading mode. When the neural network of the preset object shading model is trained, the preset shading mode may be marked in training samples. The preset shading mode may be the correct shading mode, and the preset object shading model is trained hereby so that the trained preset object shading model can judge whether the shading mode of the preset object is the preset shading mode when detecting that the to-be-detected target is shaded by the preset object. A second reminder is issued in the case that the to-be-detected target is shaded by the preset object and the shading mode does not belong to the preset shading mode. The second reminder is used to provide a reminder of adjusting the shading mode of the preset object. For example, when the to-be-detected target is the human face, the preset object is the mask, and the preset shading mode is the correct mask wearing mode. When it is detected that the human face is wearing the mask, it is judged whether the mask wearing mode is the correct mask wearing mode. If not, the second reminder is issued to remind the human face to adjust the shading mode of the preset object. Of course, in some embodiments, the preset shading mode may include multiple shading modes, such as a correct shading mode, a first wrong shading mode, a second wrong shading mode, etc. When it is detected that the shading mode in which the to-be-detected target is shaded by the preset object is the first wrong shading mode, a reminder corresponding to the first wrong shading mode is issued. When it is detected that the shading mode in which the to-be-detected target is shaded by the preset object is the second wrong shading mode, a reminder corresponding to the second wrong shading mode is issued to remind the to-be-detected target to adjust the shading mode to be the correct shading mode. For example, likewise when the to-be-detected target is the human face, the preset object is the mask, and the correct shading method is that the mask covers both the nose and the mouth, the first wrong shading method is that the mask covers the

nose but does not cover the mouth, and the reminder corresponding to the first wrong shading method is to remind the human face to cover the mouth at the same time. The second wrong shading method is that the mask covers the mouth but does not cover the nose, and the reminder corresponding to the second wrong shading mode is to remind the human face to cover the nose at the same time. When the shading mode is incorrect, the second reminder is issued, so as to adjust the shading mode of the to-be-detected target in time. The mode of the second reminder is similar to the mode of the first reminder, and it may also be in the form of matching the face frame with a text reminder as well as the face frame with a voice reminder, a separate text reminder or a separate voice reminder or a flashing reminder light, etc. Of course, if there are multiple different preset shading modes, multiple text reminders or voices are correspondingly set. For example, if the preset shading mode is the first wrong shading mode, the text reminder corresponds to the first wrong shading mode.

[0066] In some embodiments, in some service scenarios, the to-be-detected target is recognized in the case that the to-be-detected target is shaded by the preset object. In other service scenarios, if the to-be-detected target is not shaded by the preset object, the to-be-detected target is not recognized. For example, if it is detected that the human face is not wearing the mask in public places (such as high-speed rail or airplane) in which the mask needs to be worn in a certain period or special period, the human face is not recognized, and the face without the mask will not be able to enter a station through face recognition. Of course, according to the needs of the service scenarios, even if it is detected that the to-be-detected target is not shaded by the preset object, the to-be-detected target may still be recognized. For recognizing the to-be-detected target, it is necessary to perform feature extraction on the to-be-detected target. Before the feature extraction, the first image meeting the preset quality requirement may be determined as the first image for subsequent feature extraction from multiple first images containing the to-be-detected target. Specifically, the quality score of each first image may be obtained based on a quality factor of the respective first image. The quality factor of the first image includes at least one of the following items: pose information of the to-be-detected target relative to a shooting device, parameter information for reflecting the size of the to-be-detected target in the first image, and brightness information of the first image. The

pose information of the to-be-detected target relative to the shooting device may be the angle information of the to-be-detected target relative to the shooting device. Here, the angle information of the to-be-detected target relative to the shooting device may be the angle information of the to-be-detected target relative to a lens during shooting. For example, taking the lens as the origin, a three-dimensional coordinate system is established. Herein, a connecting line between the lens and the center of the earth is the X axis, an extension line in front of the lens perpendicular to the X axis is the Y axis, and what is perpendicular to the X axis and Y axis is the Z axis. The three-dimensional coordinate system is only for representing the angle between the to-be-detected target and the shooting device. In some embodiments, the selection of the origin or the selection of three directions of the three-dimensional coordinate system may be different from the embodiments of the disclosure. The angle may be divided into angles in XYZ directions relative to the lens. For example, if the to-be-detected target faces the lens, the angles in the XYZ directions are all 0 degrees. If the to-be-detected target faces a first image acquisition component broadside, the angle of the to-be-detected target relative to the first image acquisition component in the X direction is 90 degrees, the angle in the Y direction is 0 degree and the angle in the Z direction is 0 degree. Here, since the to-be-detected target has rotated 90 degrees around the X axis, the angle between the to-be-detected target and the X axis, that is, relative to the X direction is 90 degrees. Of course, the smaller the angle in each direction is, the better it is. The parameter information for reflecting the size of the to-be-detected target in the first image includes the area occupied by the to-be-detected target in the first image. The area may be expressed by the size of the region occupied by the to-be-detected target in the first image. Of course, the premise is that the to-be-detected target is completely included in the first image. If only a part of the to-be-detected target is included in the first image, the score of the quality factor about the size of the to-be-detected target in this first image is relatively low. Higher brightness information of the first image is not always the better. But the closer to the brightness of natural light at the current moment it is, the better it is, and relatively higher the score of this quality factor is. The weights of the three quality factors are set according to the relation of the influence degree of the three quality factors on the image quality. For example, the weight of the angle is set to 0.4, and the other two are set to 0.3 respectively. Of course, this is only an example, and the weights of various

quality factors may be set according to requirements. In some embodiments, in addition to these three quality factors, factors such as the fuzziness of the first image may also be included, and any factor that can affect the image quality can be used to calculate the quality score of the image. The feature extraction is performed by selecting the image of which the quality score meets the requirement, so that the extracted feature can better represent the to-be-detected target. Of course, the setting of weights may take into account the actual requirements of image detection accuracy, as well as the processing capacity and resource occupation of an image detection device. For example, in some embodiments, if the processing capacity of the image detection device is high and few resources are occupied, multiple quality factors may be considered to calculate the quality score. However, if the processing capacity of the image detection device is too low, several quality factors may be appropriately adopted to calculate the quality score. For example, the appropriate quality factors may be selected according to the time or memory space occupation required to calculate each quality factor. Therefore, it may be chosen flexibly how many quality factors to use or which ones to use. Of course, in some embodiments, it is also possible to determine a lower quality score threshold. If the quality score of the first image is lower than the quality score threshold, it will be excluded, and the first image of which the quality score is higher than the quality score threshold will be retained.

[0067] In some embodiments, before the first image is subjected to feature extraction and recognized, the first image for subsequent feature extraction may be preprocessed. The preprocessing mode may be, in the case that the first image includes multiple to-be-detected targets, determining that the to-be-detected target meeting the preset extraction requirement is in a target region of the first image, and removing an image part except the target region in the first image. Here, the target region may be a region containing one to-be-detected target. That is, when the first image contains multiple to-be-detected targets, instead of recognizing the complete first image, only the target region of the to-be-detected target meeting the preset extraction requirement is recognized. Therefore, the noise generated by other to-be-detected targets in the recognition process is reduced to a certain extent, thus reducing the influence of the to-be-detected targets that do not meet the requirement on the recognition result. The preset extraction requirement may be that

the area of the region corresponding to the to-be-detected target is larger than that of the regions corresponding to other to-be-detected targets. Other to-be-detected targets include targets other than the to-be-detected target. If there are multiple to-be-detected targets in the first image, the areas occupied by the multiple to-be-detected targets may be different from each other, and the recognition rate of the to-be-detected targets with a relatively large area is relatively high in the recognition process. Therefore, the to-be-detected targets with a relatively large area(s) is/are selected for recognition. If there are multiple to-be-detected targets with the same area, the to-be-detected target of which the center is closer to the center of the first image may be recognized. Or, in other embodiments, the target regions corresponding to all the to-be-detected targets may be acquired respectively for target detection. Of course, all the to-be-detected targets mentioned in the latter refer to the to-be-detected targets of which areas are tied for the first or are larger than a preset area extraction threshold.

[0068] In some embodiments, before the first image is subjected to feature extraction and recognized, the operation of preprocessing the first image for subsequent feature extraction may also include that upon detection that the inclination angle of the to-be-detected target in the first image is greater than a preset angle, and the first image is rotated until the inclination angle of the to-be-detected target is less than or equal to the preset angle. In some embodiments, in addition to rotating the whole first image, the rotation mode may also be only rotating the to-be-detected target, or rotating the target region containing the to-be-detected target. Therefore, the mode of straightening the to-be-detected target is not limited here. The preset angle may be within 0 to 180 degrees clockwise or counterclockwise. In the embodiment of the disclosure, the preset angle is optionally set to be 0 degree. In some embodiments, the preset angle may also be 30 degrees, 35 degrees, etc. The mode of judging whether the to-be-detected target inclines by the preset angle may be to acquire an included angle between a connecting line of a preset first key point and a preset second key point in the to-be-detected target and a vertical line, and judge whether the included angle is larger than the preset angle. If so, the first image is rotated so that the included angle is smaller than or equal to the preset angle, and the rotated preset first key point is located on the upper side of the preset second key point. Here, the upper side is determined relative to a bottom edge of the first

image. Of course, the inclination angle may also be the inclination angle of the to-be-detected target relative to a certain position of the first image, for example, the inclination angle of the to-be-detected target relative to the center of the first image. Of course, the preset angle here may be set according to the requirements of different scenarios. For example, it may be determined according to the area of the first image occupied by the region where the to-be-detected target is located. For example, when the area of the region where the to-be-detected target is located is larger than a first area preset value, the preset angle may be set to be larger than 30 degrees. When the area of the region where the to-be-detected target is located is smaller than a second area preset value, the preset angle may be set to be smaller than 30 degrees. The larger the area of the region where the to-be-detected target is located is, the larger the area of the to-be-detected target is. That is, the less the to-be-detected target is affected by the angle, the more tolerant the inclination angle of the to-be-detected target is. Otherwise, the more the to-be-detected target is affected by the inclination angle, the stricter the inclination angle of the to-be-detected target is. Of course, this is only an example. In some embodiments, the other correspondence between areas and preset angles may also be set, which is not specified here.

[0069] For example, referring to FIG. 2 and FIG. 3, FIG. 2 is a schematic diagram of a first image in an embodiment of an image detection method according to an embodiment of the disclosure. FIG. 3 is a schematic diagram of a preprocessed first image in an embodiment of an image detection method according to an embodiment of the disclosure. As shown in FIG. 2, a lower part of a to-be-detected target 21 in a first image 20 is shaded by a preset object 22, and the to-be-detected target 21 is apparently inclined to the left. That is, an included angle between a connecting line of an upper left corner point (a first preset key point) and a lower left corner point (a second preset key point) of the to-be-detected target 21 and a vertical line is 30 degrees. That is, if the inclination angle of the to-be-detected target 21 is 30 degrees, which is greater than the preset angle of 0 degrees, the first image 20 is rotated rightwards, namely clockwise by 30 degrees, and the rotated first image is shown in FIG. 3. In FIG. 3, the included angle between the connecting line of the upper left corner point (the first preset key point) and the lower left corner point (the second preset key point) of the to-be-detected target 21 and the vertical

line is 0 degree, which is equal to the preset angle of 0 degree.

[0070] In the case that the inclination angle of the to-be-detected target in the first image is larger than the preset angle, the to-be-detected target is straightened, so that the influence caused by inclination of the to-be-detected target is reduced in the subsequent process of living body detection or target recognition on the to-be-detected target.

[0071] In some embodiments, before the first image is subjected to feature extraction and recognized, the living body detection may also be performed on the first image for subsequent feature extraction. If the result of the living body detection is that the to-be-detected target is a living body, it is determined to execute at least extraction of a first feature of an unshaded part of the to-be-detected target from the first image and its subsequent steps. If there are multiple to-be-detected targets in the first image, the to-be-detected target with the largest area is selected for the living body detection. The target region corresponding to the to-be-detected target may be input into a living body detection model for the living body detection. The living body detection model is trained by using several images containing the to-be-detected target shaded by the preset object. The to-be-detected target is recognized only when the to-be-detected target is the living body. Thus, the security of recognition is enhanced and prosthesis attack may be prevented to a certain extent.

[0072] In the process of recognizing the to-be-detected target, the first feature of the unshaded part of the to-be-detected target is at least extracted from the first image as a to-be-recognized feature of the to-be-detected target. The first feature refers to the feature of an unshaded key point in the to-be-detected target. The first feature of the unshaded part of the to-be-detected target may be extracted from the first image. A second feature of a shaded part of the to-be-detected target is acquired. The first feature and the second feature are taken as to-be-recognized features of the to-be-detected target. The second feature here refers to the feature of the key point of the shaded part in the to-be-detected target. There are two modes to acquire the second feature of the shaded part. One is to extract the feature of the shaded part from the first image as the second feature. That is, although this part is shaded by the preset object, the second feature of the shaded part is extracted in the manner of extracting the first feature without shading. That is, the same

processing mechanism is adopted regardless of whether the to-be-detected target is shaded by the preset object, that is, whether it is shaded by the preset object does not affect the feature extraction process. Of course, in some embodiments, if there is no preset object that shades the to-be-detected target, this mode may still be used to extract the feature of the key point in the to-be-detected target. For example, when the to-be-detected target is the human face, and the preset object is the mask, the second feature extraction mode is to extract the features of each key point on the human face as if the human face is not covered by the mask. That is to say, the same processing mechanism is adopted for the human face wearing the mask and the human face not wearing the mask, that is, wearing the mask may not affect the feature extraction process. Another mode is to acquire a preset feature of the shaded part as the second feature. The preset feature may be the feature obtained based on at least one reference feature. Each reference feature is extracted from a region, corresponding to the shaded part, in a reference target without the shaded part. That is to say, before the first image is recognized, the reference features are preset for the key point of the shaded part, that is, the feature of the shaded part is filled up. For example, a plurality of features corresponding to preset parts in the to-be-detected target of which the detection result is no shaded part are extracted in advance, and then the average value of the plurality of extracted features is filled up as the reference feature of the shaded part by the object. A plurality of features corresponding to the preset parts in the to-be-detected target of which the detection result is not being shaded by the preset object may be extracted in advance, and then the average value of the plurality of extracted features is filled up as the reference feature of the shaded part by the preset object. For example, features corresponding to preset parts in a plurality of faces without masks are extracted in advance, that is, features of parts wearing masks, such as the nose, the mouth, etc., and the average value of the plurality of extracted features is filled up as preset reference features of parts covered by the masks. As for the feature determination manner of the shaded part, the feature of the shaded part may be directly extracted. Since the feature of the shaded part can be different with different to-be-detected targets to a certain extent, this manner can improve the recognition accuracy. It may also acquire the preset feature as the feature of the shaded part. This manner does not need to perform feature extraction on the shaded part, thus reducing the loss of processing resources and improving the processing efficiency.

[0073] In some embodiments, after the to-be-recognized feature of the to-be-detected target is obtained, the to-be-detected target is recognized using the to-be-recognized feature. A recognized scenario may be divided into a 1: 1 scenario and a 1: N scenario. 1:1 refers to the comparison between two features, while 1: N refers to the comparison between one feature and multiple features. In the 1: 1 scenario, that is, in the case that one preset target is included, a first similarity between the to-be-recognized feature and a pre-stored feature of the preset target is acquired, and in the case that the first similarity meets a first preset condition, the recognition result is determined to include that the to-be-detected target passes the identity authentication. The first preset condition may be that the first similarity is greater than a first similarity threshold. The first similarity threshold in the case that the to-be-recognized feature includes the second feature of the shaded part of the to-be-detected target is smaller than the first similarity threshold in the case that the to-be-recognized feature does not include the second feature. If the to-be-recognized feature includes the second feature, the second feature may be different from a real feature at a key point of the shaded part of the to-be-detected target. Hence, in such a case the recognition accuracy can be improved by appropriately reducing the similarity threshold. In the case that the to-be-recognized feature does not include the second feature, the selection of the first similarity threshold may be determined according to the ratio of the number of shaded key points to the number of overall key points of the to-be-detected target. For example, if the key points of the shaded part are one-third of the number of the overall key points of the to-be-detected target, the first similarity threshold may be determined to be one-third of the similarity threshold for the recognition of the unshaded to-be-detected target. At this time, it may be set that when the to-be-recognized feature includes the second feature, the first similarity threshold may be smaller than the first similarity threshold in the case that the to-be-recognized feature does not include the second feature by 0.1 or other values, which is not specified here. For example, the similarity threshold for the recognition of the unshaded to-be-detected target may be between 0.6 and 1. Of course, this is only an example. In some embodiments, the first similarity threshold in the case that the to-be-recognized feature includes the second feature of the shaded part of the to-be-detected target may also be equal to the first similarity threshold in the case that the to-be-recognized feature does not include the second feature. If the to-be-recognized feature includes the second feature, it may also be

determined according to the above method or according to an actual situation.

[0074] In some embodiments, before the first similarity between the to-be-recognized feature and the pre-stored feature of the preset target is acquired, the association between a user account and the pre-stored feature of the preset target is first established. The implementation mode is as follows: an account is registered for a user in response to an account registration request. The account here may be some electronic payment accounts, and any Application that can perform target recognition may respond to the account registration request and register the account for the user. The user may register by a mobile phone number in the corresponding Application. After successful registration, the user acquires information such as a user name and a password. The second image meeting the preset quality requirement is determined from at least one of the second images shot by the user, and the feature of the preset part of the user is extracted from the determined second image. The preset part here is the same as the preset part of the to-be-detected target. The step of selecting the second image meeting the preset quality requirement is the same as the above step of selecting the first image meeting the preset quality requirement. Finally, the feature of the preset part is associated with the account, and the feature of the preset part is stored as the pre-stored feature of the preset target. That is, the preset part of the user is the preset target.

[0075] In the 1: N scene, that is, in the case that multiple preset targets are included, the second similarity between the to-be-recognized feature and the pre-stored feature of each preset target is acquired, and the recognition result is determined to include that the identity of the to-be-detected target is determined as the identity of the preset target of which the second similarity meets the second preset condition. The second preset condition may be that the second similarity is greater than a second similarity threshold. In general, meeting the second preset condition means that, the second similarity is not only greater than the second similarity threshold, but also a parameter which takes the maximum value among all the second similarities. That is, the identity of the preset target corresponding to the maximum second similarity is selected as the identity of the to-be-detected target. The second similarity threshold in the case that the to-be-recognized feature includes the second feature is smaller than the second similarity threshold in the case that the to-be-recognized feature does not include the second feature. If the to-be-

recognized feature includes the second feature, the second feature may be different from the real feature of the to-be-detected target at the key point of the shaded part. Hence, in such a case, the recognition accuracy can be improved by appropriately reducing the similarity threshold. The method for determining the second similarity threshold is the same as the method for determining the first similarity threshold.

[0076] 1: N may be a scenario involving multiple human faces. For example, a face recognition gate is installed at the entrance and exit of an office building or a company. In such a scenario, it is necessary to register everyone who needs to enter and exit the building or company, and form a face database. When the registered person appears at the gate, the camera on the gate detects and captures the human face, and compares the captured human face with the face database. When the comparison is successful, the gate is opened. When an unregistered person appears at the gate, the comparison is unsuccessful, and the gate does not respond.

[0077] According to the above solution, the first image containing the to-be-detected target is detected to find out whether the to-be-detected target is shaded or not. Then, the preset operation matched with the detection result is executed. It can be judged whether the to-be-detected target is shaded and then the subsequent preset operation matched with the detection result can be performed, which implements flexible processing based on the shading state of the to-be-detected target in the image.

[0078] Based on the foregoing embodiments, the embodiment of the disclosure further provides an image detection method. The method utilizes the model detection capability based on a deep learning algorithm, and starting from a human face verification scenario, realizes the image detection flow in the 1: 1 and 1: N scenarios. Meanwhile, a method for checking whether a mask is worn or not is provided, and an implementation mode of face comparison and face retrieval in the mask wearing process is provided.

[0079] (1) The face verification scenario mainly includes a 1: 1 scenario and a 1: N scenario.

[0080] For example, in a payment scenario, 1: 1 indicates that a face photo captured in real time is subjected to 1: 1 check with a database picture bound by a member. If it is

confirmed that they are the same user, the authentication is passed. For another example, the 1: N scenario is more related to face retrieval. For example, a face recognition gate is installed at the entrance and exit of the office building or the company. In such a scenario, it is necessary to register every user who needs to enter and exit the building or company, and form a face database. When the user appears at the gate, the camera on the gate detects and captures the human face, and compares the captured face image with the pictures in the face database. When the comparison is successful, the gate is opened. When the comparison is unsuccessful, the gate has no response, that is, the gate remains closed.

[0081] In some embodiments, regardless of in the 1: 1 scenario or the 1: N scenario, it is necessary to take precautions against a non-living body, for example, the attack from a dummy such as a photo (including a photo taken of a person, an electronic composite photo, etc.) or a mask.

[0082] (2) The image detection method in the 1: 1 scenario may be implemented in the following mode.

[0083] First, the user has an account, such as an electronic payment account. The user registers with the mobile phone number in an Application (APP) corresponding to the electronic payment account. After successful registration, the user acquires information such as the user name and the password, thus obtaining an electronic payment account. Secondly, the Application guides the user to perform the operation of binding the human face through some activities. When the user passes the living body authentication, face recognition is performed. When the face quality in a video meets a face collection demand, the face image of the user is collected. The frame with the highest quality in the shooting process may be selected, and the quality evaluation criteria include one or more of the dimensions such as the angle of the human face, the intensity of illumination and the size of the human face.

[0084] The collected face image of the user may be associated with the account, specifically, it may be associated with an identifier of the account and taken as a comparison picture of the face database. When the user spends money online, a payment link is entered after an order is confirmed. At this time, if the user has bound the face

image and selected face payment, a face capture link is entered. During face detection, the wearing of the mask (that is, face attribute detection) is checked. If the mask is worn, the follow-up process is continued. If the mask is not worn, the user may be reminded through voice play or text display to wear the mask.

[0085] In the case that the user wears the mask, face detection is performed. During the shooting process, a frame with the highest quality is selected, and the selected frame is subjected to face alignment (for example, if the human face has a certain inclination angle, face straightening is performed), and a face feature value is extracted based on the aligned face picture.

[0086] There are two modes to extract and compare the face feature value. The first mode is to extract the face feature value with the mask part. For example, the face feature value without the mask in the pictures of the face database is A, the human face with the mask is subjected to feature extraction in the same way, the extracted feature value is A1, and then the feature vector of the feature value A1 with the mask and the feature vector of the feature value A without the mask are compared. The second mode is to extract the face feature value of a visible part above the mask. Assuming that the whole face has 128 key points and 64 key points above the mask, the feature value extracted from the 64 key points of the visible part above the mask is compared with the feature value extracted from the corresponding 64 key points in the pictures of the face database. Furthermore, the generated feature value is compared with the feature value of the picture in the face database of the user, and when it exceeds a comparison threshold (for example, the comparison threshold is set to be 0.8, and it is considered as the same user if the similarity exceeds 0.8), the comparison is passed and the face verification process ends.

[0087] (3) The image detection method in the 1: N scenario may be implemented in the following modes.

[0088] First, in an office building personnel management service system, the photo(s) of each user who has permission to enter the building (forming the comparison pictures of the face database in a passing scenario of the building) are entered. At this time, the pictures in the face database are in one-to-one correspondence to the building users. Secondly, the pictures in the face database are bound and associated with a face

recognition device of the gate, so that the face recognition device of the gate may read the pictures in the face database. When a person needs to pass through the gate, the face recognition device on the gate enters a face capture state after detecting the face information (at this time, a camera module on the face recognition device may be turned on all the time to track the human face, and the face frame always moves along with the movement of the human face).

[0089] Then, during face detection, the wearing of the mask is checked (that is, the face attribute detection), and different operations are conducted according to different check results. For specific operation contents, the content descriptions in the above 1: 1 scenario may be referred to.

[0090] There are also two modes to extract and retrieve the face feature value in the 1: N scenario. The first mode is to extract the face feature value including the mask part. Assuming that there are five face feature values A, B, C, D and E without the mask in the comparison pictures of the face database, the human face with the mask is subjected to the feature extraction in the same way, and the five feature values A1, B1, C1, D1 and E1 are extracted. When the feature value A1 with the mask is retrieved, the distance between the feature vector of the feature value A1 and the feature vector of the feature value A is the shortest, so the corresponding feature value A is retrieved. The second mode is to extract the face feature value of the visible part above the mask. Herein, the generation part of the feature value is the same as the method in the above-mentioned 1: 1 scenario, which may be seen in the above descriptions. After the feature value is generated, the generated feature value is subjected to face feature retrieval in the pictures of the face database of all users in the whole building (1: N search). When there is a feature value exceeding the comparison threshold, it is considered that the retrieval is successful. After successful retrieval, the face recognition device transmits an opening signal, the gate opens, and the face verification process ends.

[0091] In the embodiment of the disclosure, through the above-mentioned method, before face recognition, it is possible to check whether the user wears the mask and give a reminder. In the case that the user wears the mask, accurate face recognition is conducted and the face verification process is completed, which greatly improves the

pass efficiency and reduces the risks brought by taking off the mask.

[0092] The image detection method may be performed by an image detection device, for example, the image detection method may be executed by terminal equipment or a server or other processing equipment. Herein, the terminal equipment may be User Equipment (UE), mobile equipment, a terminal, a cellular phone, a cordless phone, Personal Digital Assistant (PDA), a handheld device, a computing device, a vehicle-mounted device, a wearable device, etc. In some possible implementations, the image detection method may be implemented in a manner that a processor calls a computer readable instruction stored in a memory.

[0093] Referring to FIG. 4, FIG. 4 is a schematic structural diagram of an embodiment of an image detection apparatus according to an embodiment of the disclosure. The image detection apparatus 40 includes an image acquisition module 41, a target detection module 42 and an operation execution module 43. The image acquisition module 41 is configured to acquire a first image containing a to-be-detected target. The target detection module 42 is configured to detect the first image to obtain a detection result of the first image, the detection result including whether the to-be-detected target in the first image is shaded by a preset object. The operation execution module 43 is configured to execute a preset operation matched with the detection result.

[0094] According to the above solution, the first image containing the to-be-detected target is detected to find out whether the to-be-detected target is shaded or not. Then, the preset operation matched with the detection result is executed. It can be judged whether the to-be-detected target is shaded and then the subsequent preset operation matched with the detection result can be performed, which achieves flexible processing based on the shading state of the to-be-detected target in the image.

[0095] In some embodiments, the detection result is obtained by detecting the first image by the target detection module 42 using the neural network.

[0096] According to the above solution, the neural network trained in advance is used for detection, which makes the detection result more accurate and the detection speed faster.

[0097] In some embodiments, the operation that the operation execution module 43 executes the preset operation matched with the detection result includes that: in the case that the to-be-detected target is not shaded by the preset object, a first reminder is issued. The first reminder is used to give a reminder of shading the to-be-detected target using the preset object.

[0098] According to the above solution, the first reminder is issued when the to-be-detected target is not shaded, so that a reminder can be given timely in the situation that the to-be-detected target is not shaded by the preset object, and then a reminded person can take corresponding measures in time.

[0099] In some embodiments, the detection result also includes whether the shading mode in which the to-be-detected target is shaded by the preset object is a preset shading mode. The operation that the operation execution module 43 executes the preset operation matched with the detection result includes that: a second reminder is issued when the to-be-detected target is shaded by the preset object but the shading mode does not belong to the preset shading mode. The second reminder is used for giving a reminder of adjusting the shading mode of the preset object.

[00100] According to the above solution, when the shading mode is incorrect, the second reminder is issued, so that the shading mode of the to-be-detected target can be adjusted in time.

[00101] In some embodiments, the operation that the operation execution module 43 executes the preset operation matched with the detection result includes that: in the case that the to-be-detected target is shaded by the preset object, at least a first feature of an unshaded part of the to-be-detected target is extracted from the first image as a to-be-recognized feature of the to-be-detected target; and the to-be-detected target is recognized using the to-be-recognized feature, and a recognition result is obtained.

[00102] According to the above solution, when the to-be-detected target is shaded by the preset object, the feature of the unshaded part is extracted for recognition, which realizes recognition based on a local feature of the to-be-detected target. Moreover, since the local feature is unshaded, it can represent the to-be-detected target, thus ensuring the

recognition accuracy to a certain extent.

[00103] In some embodiments, the operation that the operation execution module 43 extracts at least the first feature of the unshaded part of the to-be-detected target from the first image as the to-be-recognized feature of the to-be-detected target include that: the first feature of the unshaded part of the to-be-detected target is extracted from the first image, and a second feature of a shaded part of the to-be-detected target is acquired; and the first feature and the second feature are taken as to-be-recognized features of the to-be-detected target.

[00104] According to the above solution, in addition to the adoption of the feature of the unshaded part of the to-be-detected target, the feature of the shaded part is also combined, and then the feature richness of the to-be-detected target may be improved.

[00105] In some embodiments, the operation that the operation execution module 43 acquires the second feature of the shaded part of the to-be-detected target includes that: the feature of the shaded part is extracted from the first image as the second feature; or, a preset feature of the shaded part is acquired as the second feature. The preset feature includes a feature obtained based on at least one reference feature. Each reference feature is extracted from a region, corresponding to the shaded part, in a reference target without the shaded part.

[00106] According to the above solution, as for a feature determination manner of the shaded part, the feature of the shaded part may be directly extracted. Since the feature of the shaded part can be different with different to-be-detected targets to a certain extent, this manner can improve the recognition accuracy. It may also acquire the preset feature as the feature of the shaded part. This manner does not need to perform feature extraction on the shaded part, thus reducing the loss of processing resources and improving the processing efficiency.

[00107] In some embodiments, the operation that the operation execution module 43 uses the to-be-recognized feature to recognize the to-be-detected target and obtains the recognition result includes at least one of the following operations. In the case that one preset target is included, a first similarity between the to-be-recognized feature and a pre-

stored feature of the preset target is acquired, and in the case that the first similarity meets a first preset condition, the recognition result is determined to include that the to-be-detected target passes identity authentication. In the case that multiple preset targets are included, second similarities, each between the to-be-recognized feature and the pre-stored feature of a respective one of the preset targets are acquired, and the recognition result is determined to include that the identity of the to-be-detected target is determined as the identity of the preset target corresponding to the second similarity meeting a second preset condition.

[00108] According to the above solution, the first similarity with the pre-stored feature of a specific preset target is calculated, or the similarities with the pre-stored features of multiple preset targets are calculated, so that the to-be-detected target can be compared with a certain specific preset target or the preset targets in a certain database according to actual scenario requirements.

[00109] In some embodiments, the first preset condition includes that the first similarity is greater than a first similarity threshold. The second preset condition includes that the second similarity is greater than a second similarity threshold.

[00110] According to the above solution, the first similarity threshold is set in different scenarios, so that the recognition result is more accurate.

[00111] In some embodiments, the first similarity threshold in the case that the to-be-recognized feature includes the second feature of the shaded part of the to-be-detected target is smaller than the first similarity threshold in the case that the to-be-recognized feature does not include the second feature. The second similarity threshold in the case that the to-be-recognized feature includes the second feature is smaller than the second similarity threshold in the case that the to-be-recognized feature does not include the second feature.

[00112] According to the above solution, if the to-be-recognized feature includes the second feature, the second feature may be different from the real feature at the key point of the shaded part of the to-be-detected target. Hence, in such a case, the recognition accuracy can be improved by appropriately reducing the similarity threshold.

[00113] In some embodiments, the image detection apparatus 40 further includes a pre-storage module (not shown in the figure). Before the operation execution module 43 acquires the first similarity between the to-be-recognized feature and the pre-stored feature of the preset target, the pre-storage module is configured to register an account for a user in response to an account registration request; determine a second image meeting a preset quality requirement from at least one of the second images shot by the user, and extract the feature of a preset part of the user from the determined second image; and associate the feature of the preset part with the account, and store the feature of the preset part as the pre-stored feature of the preset target.

[00114] According to the above solution, the second image meeting the quality requirement is determined firstly, to extract the feature of the preset part, so as to make the extracted feature more accurate.

[00115] In some embodiments, in the case that the to-be-detected target is shaded by the preset object, before at least the first feature of the unshaded part of the to-be-detected target is extracted from the first image, the operation execution module 43 is further configured to execute at least one of the following steps: the first image meeting the preset quality requirement is determined as the first image for subsequent feature extraction from multiple first images containing the to-be-detected target; the first image for subsequent feature extraction is preprocessed; living body detection is performed on the first image for subsequent feature extraction, and it is determined to execute at least extraction of the first feature of the unshaded part of the to-be-detected target from the first image and subsequent steps thereof in the case that a living body detection result indicates that the to-be-detected target is a living body.

[00116] According to the above solution, before feature extraction, preprocessing is performed firstly, so that the extracted feature is more accurate. The to-be-detected target is recognized only when the to-be-detected target is a living body, so that the security of recognition is enhanced, and prosthesis attack may be prevented to a certain extent.

[00117] In some embodiments, the operation that the operation execution module 43 determines the first image meeting the preset quality requirement as the first image for subsequent feature extraction from the multiple first images containing the to-be-detected

target includes that: the quality score of each first image is correspondingly obtained based on a quality factor of each first image. The quality factor of the first image includes at least one of the following items: pose information of the to-be-detected target relative to a shooting device, parameter information for reflecting the size of the to-be-detected target in the first image, and brightness information of the first image. Based on the quality scores, the first image meeting the preset quality requirement is determined as the first image for subsequent feature extraction. The quality score of the selected first image is higher than that of other first images.

[00118] According to the above solution, feature extraction is performed by determining the image of which the quality score meets the requirement, so that the extracted feature can better represent the to-be-detected target.

[00119] In some embodiments, the operation that the operation execution module 43 preprocesses the first image for subsequent feature extraction includes that: in the case that the first image includes a plurality of to-be-detected targets, it is determined that the to-be-detected target meeting a preset extraction requirement is in a target region in the first image, and an image part in the first image except the target region is removed; and/or, it is detected that the inclination angle of the to-be-detected target in the first image is greater than a preset angle, and the first image is rotated until the inclination angle of the to-be-detected target is less than the preset angle.

[00120] According to the above solution, when there are multiple to-be-detected targets in the first image, only the to-be-detected target that meets the preset extraction requirement is determined, and the to-be-detected targets that do not meet the requirement are discarded, thus reducing the influence of the to-be-detected targets that do not meet the requirement on the recognition result. Secondly, in the case that the to-be-detected target in the first image has an inclination angle, the to-be-detected target is straightened, thus reducing the influence caused by inclination of the to-be-detected target.

[00121] In some embodiments, the preset extraction requirement includes that the area of the corresponding region of the to-be-detected target is larger than that of the corresponding regions of other to-be-detected targets. Other to-be-detected targets

include targets other than the to-be-detected target.

[00122] According to the above solution, since the larger the area of the to-be-detected target is, the more accurate the extracted feature is, the to-be-detected target with larger area is selected so that the to-be-detected result is more accurate.

[00123] In some embodiments, the to-be-detected target includes a human face. The preset object includes a mask.

[00124] According to the above solution, it is judged whether the human face is wearing the mask and a corresponding operation is executed. For example, if the human face is not wearing the mask or the way of wearing the mask is inaccurate, a corresponding reminder may be issued, so that the user can make adjustments timely. If the human face wears the mask, the human face is recognized.

[00125] According to the above solution, the first image containing the to-be-detected target is detected to find out whether the to-be-detected target is shaded or not. Then, the preset operation matched with the detection result is executed. It can be judged whether the to-be-detected target is shaded and then the subsequent preset operation matched with the detection result can be performed, which achieves flexible processing based on the shading state of the to-be-detected target in the image.

[00126] Referring to FIG. 5, FIG. 5 is a schematic structural diagram of an embodiment of an electronic device according to an embodiment of the disclosure. The electronic device 50 includes a memory 51 and a processor 52. The processor 52 is configured to execute a program instruction stored in the memory 51 to implement the steps in any of the image detection method embodiments. In an implementation scenario, the electronic device 50 may include, but is not limited to, a microcomputer and a server. In addition, the electronic device 50 may also include mobile devices such as a notebook computer and a tablet computer, which are not limited here.

[00127] The processor 52 is configured to control the processor itself and the memory 51 to implement the steps in any of the image detection method embodiments. The processor 52 may also be called a Central Processing Unit (CPU). The processor 52 may be an integrated circuit chip with signal processing capability. The processor 52 may also

be a general-purpose processor, a Digital Signal Processor (DSP), an Application Specific Integrated Circuit (ASIC), a Field-Programmable Gate Array (FPGA) or other programmable logic devices, a discrete gate or transistor logic device, or a discrete hardware component. The general-purpose processor may be a microprocessor or the processor may be any conventional processor, etc. Furthermore, the processor 52 may be jointly implemented by the integrated circuit chip.

[00128] According to the above solution, the first image containing the to-be-detected target is detected to find out whether the to-be-detected target is shaded or not. Then, the preset operation matched with the detection result is executed. It can be judged whether the to-be-detected target is shaded and then the subsequent preset operation matched with the detection result can be performed, which achieves flexible processing based on the shading state of the to-be-detected target in the image.

[00129] Referring to FIG. 6, FIG. 6 is a schematic structural diagram of an embodiment of a computer readable storage medium according to an embodiment of the disclosure. The computer readable storage medium 60 stores a program instruction 61 capable of being executed by a processor. The program instruction 61 is configured to implement the steps in any of the image detection method embodiments.

[00130] According to the above solution, the first image containing the to-be-detected target is detected to find out whether the to-be-detected target is shaded or not. Then, the preset operation matched with the detection result is executed. It can be judged whether the to-be-detected target is shaded and then the subsequent preset operation matched with the detection result can be performed, which achieves flexible processing based on the shading state of the to-be-detected target in the image.

[00131] In some embodiments, the embodiments of the disclosure provide a computer program, which includes a computer readable code. When running in an electronic device, the computer readable code is executed by a processor in the electronic device to implement the above method.

[00132] In some embodiments, the functions or included modules of the apparatus provided by the embodiments of the disclosure may be used to execute the method

described in the above method embodiments, and its implementation may refer to the descriptions of the above method embodiments for brevity.

[00133] The above descriptions of various embodiments tend to emphasize the differences between the various embodiments, and their similarities may be referred to each other for brevity.

[00134] In several embodiments provided by the disclosure, it should be understood that the disclosed methods and apparatuses may be implemented in other ways. For example, the apparatus embodiment described above is only schematic, and for example, division of the modules or units is only logic function division, and other division manners may be adopted during practical implementation. For example, units or components may be combined or integrated into another system, or some characteristics may be neglected or not executed. In addition, coupling or direct coupling or communication connection between each displayed or discussed component may be indirect coupling or communication connection, implemented through some interfaces, of the device or the units, and may be electrical and mechanical or adopt other forms.

[00135] In addition, each function unit in each embodiment of the disclosure may be integrated into a processing unit, or each unit may exist independently, or two or more than two units may also be integrated into a unit. The integrated unit may be implemented in a hardware form and may also be implemented in form of software functional unit. When being implemented in form of software functional unit and sold or used as an independent product, the integrated unit may be stored in a computer-readable storage medium. Based on such an understanding, the technical solutions of the embodiments of the disclosure essentially, or the part contributing to the related art, or all or part of the technical solutions may be implemented in a form of a software product. The computer software product is stored in a storage medium and includes several instructions for instructing a computer device (which may be a personal computer, a server, or a network device) or a Processor to perform all or some of the steps of various implementation modes and methods described in the embodiments of the disclosure. The above-mentioned storage medium includes: various media capable of storing program codes such as a U disk, a mobile hard disk, a Read-Only Memory (ROM), a Random

Access Memory (RAM), a magnetic disk, or an optical disk.

INDUSTRIAL PRACTICABILITY

[00136] The embodiments of the disclosure provide an image detection method and related apparatus, a device, a storage medium, and a computer program. The method includes that: a first image containing a to-be-detected target is acquired; the first image is detected to obtain a detection result of the first image, the detection result including whether the to-be-detected target in the first image is shaded by a preset object; and a preset operation matched with the detection result is executed. According to the image detection method provided by the embodiments of the disclosure, it can be judged whether the to-be-detected target is shaded and then the subsequent preset operation matched with the detection result is performed, which achieves flexible processing based on the shading state of the to-be-detected target in the image.

CLAIMS

1. An image detection method, comprising:

acquiring a first image containing a to-be-detected target;

detecting the first image to obtain a detection result of the first image, the detection result comprising whether the to-be-detected target in the first image is shaded by a preset object; and

executing a preset operation matched with the detection result.

2. The method of claim 1, wherein the detection result is obtained by detecting the first image using a neural network.

3. The method of claim 1 or 2, wherein executing the preset operation matched with the detection result comprises:

in the case that the to-be-detected target is not shaded by the preset object, issuing a first reminder, the first reminder being configured to give a reminder of shading the to-be-detected target using the preset object.

4. The method of any one of claims 1-3, wherein the detection result comprises whether a shading mode in which the to-be-detected target is shaded by the preset object is a preset shading mode; wherein executing the preset operation matched with the detection result comprises:

in the case that the to-be-detected target is shaded by the preset object and the shading mode does not belong to the preset shading mode, issuing a second reminder, the second

reminder being configured to give a reminder of adjusting the shading mode of the preset object.

5. The method of any one of claims 1-4, wherein executing the preset operation matched with the detection result comprises:

in the case that the to-be-detected target is shaded by the preset object, extracting at least a first feature of an unshaded part of the to-be-detected target from the first image as a to-be-recognized feature of the to-be-detected target; and

recognizing the to-be-detected target using the to-be-recognized feature and obtaining a recognition result.

6. The method of claim 5, wherein extracting at least the first feature of the unshaded part of the to-be-detected target from the first image as the to-be-recognized feature of the to-be-detected target comprises:

extracting the first feature of the unshaded part of the to-be-detected target from the first image, and acquiring a second feature of a shaded part of the to-be-detected target; and

taking the first feature and the second feature as the to-be-recognized features of the to-be-detected target.

7. The method of claim 6, wherein acquiring the second feature of the shaded part of the to-be-detected target comprises:

extracting a feature of the shaded part from the first image as the second feature; or

acquiring a preset feature of the shaded part as the second feature, the preset feature comprising a feature obtained based on at least one reference feature, and each of the at least

one reference feature being extracted from a region, corresponding to the shaded part, in a reference target without the shaded part.

8. The method of any one of claims 5-7, wherein recognizing the to-be-detected target using the to-be-recognized feature and obtaining the recognition result comprises at least one of the following items:

in the case that there is one preset target, acquiring a first similarity between the to-be-recognized feature and a pre-stored feature of the preset target, and in the case that the first similarity meets a first preset condition, determining that the recognition result comprises that the to-be-detected target passes identity authentication; or

in the case that there are multiple preset targets, acquiring a second similarity between the to-be-recognized feature and the pre-stored feature of each of the preset targets, respectively, and determining that the recognition result comprises that the identity of the to-be-detected target is determined as the identity of the preset target corresponding to the second similarity meeting a second preset condition.

9. The method of claim 8, comprising at least one of the following items:

the first preset condition comprising that the first similarity is greater than a first similarity threshold; and

the second preset condition comprising that the second similarity is greater than a second similarity threshold.

10. The method of claim 9, comprising at least one of the following items:

the first similarity threshold in the case that the to-be-recognized feature comprises the second feature of the shaded part of the to-be-detected target being smaller than the first

similarity threshold in the case that the to-be-recognized feature does not comprise the second feature; or

the second similarity threshold in the case that the to-be-recognized feature comprises the second feature being smaller than the second similarity threshold in the case that the to-be-recognized feature does not comprise the second feature.

11. The method of any one of claims 8-10, wherein before acquiring the first similarity between the to-be-recognized feature and the pre-stored feature of the preset target, the method further comprises:

registering an account for a user in response to an account registration request;

determining a second image meeting a preset quality requirement from at least one second image shot by the user, and extracting a feature of a preset part of the user from the determined second image; and

associating the feature of the preset part with the account, and saving the feature of the preset part as the pre-stored feature of the preset target.

12. The method of any one of claims 5-11, wherein in the case that the to-be-detected target is shaded by the preset object, before extracting at least the first feature of the unshaded part of the to-be-detected target from the first image, the method further comprises at least one of the following steps:

determining the first image meeting a preset quality requirement as the first image for subsequent feature extraction from multiple first images containing the to-be-detected target;

preprocessing the first image for subsequent feature extraction; and

performing living body detection on the first image for subsequent feature extraction, and in the case that a result of the living body detection is that the to-be-detected target is a

living body, determining to extract at least the first feature of the unshaded part of the to-be-detected target from the first image and perform subsequent steps.

13. The method of claim 12, wherein determining the first image meeting the preset quality requirement as the first image for subsequent feature extraction from the multiple first images containing the to-be-detected target comprises:

for each of the first images, obtaining a quality score of the first image based on a quality factor of the first image, the quality factor of the first image comprising at least one of the following items: pose information of the to-be-detected target relative to a shooting device, parameter information for reflecting a size of the to-be-detected target in the first image, and brightness information of the first image; and

determining the first image meeting the preset quality requirement as the first image for subsequent feature extraction based on the quality score, the quality score of a selected first image being higher than quality scores of other ones of the first images.

14. The method of claim 12 or 13, wherein preprocessing the first image for subsequent feature extraction comprises at least one of the following:

in the case that the first image comprises multiple to-be-detected targets, determining that the to-be-detected target meeting the preset extraction requirement is in a target region of the first image, and removing an image part except the target region in the first image; or

detecting that an inclination angle of the to-be-detected target in the first image is greater than a preset angle, and rotating the first image until the inclination angle of the to-be-detected target is less than or equal to the preset angle.

15. The method of claim 14, wherein the preset extraction requirement comprises that an area of a region corresponding to the to-be-detected target is larger than an area of regions

corresponding to other to-be-detected targets, the other to-be-detected targets comprising targets other than the to-be-detected target.

16. The method of claim 10, wherein the to-be-detected target comprises a human face, and the preset object comprises a mask.

17. An image detection apparatus, comprising:

an image acquisition module, configured to acquire a first image containing a to-be-detected target;

a target detection module, configured to detect the first image to obtain a detection result of the first image, the detection result comprising whether the to-be-detected target in the first image is shaded by a preset object; and

an operation execution module, configured to execute a preset operation matched with the detection result.

18. An electronic device, comprising a memory and a processor, wherein the processor is configured to execute program instructions stored in the memory to implement the method of any one of claims 1-16.

19. A computer readable storage medium, storing program instructions thereon, wherein the program instructions, when executed by a processor, implement the method of any one of claims 1-16.

20. A computer program, comprising a computer readable code, wherein the computer readable code, when running in an electronic device, causes a processor of the electronic device to implement the method of any one of claims 1-16.