

IMAGE PROCESSING METHOD AND APPARATUS, DEVICE, AND MEDIUM

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims priority to Chinese Patent Application No. 2023110802296, entitled “IMAGE PROCESSING METHOD AND APPARATUS, DEVICE, AND MEDIUM” filed with the China National Intellectual Property Administration on August 25, 2023, which is incorporated herein by reference in its entirety.

FIELD OF THE TECHNOLOGY

[0002] This application relates to the field of artificial intelligence technologies, and in particular, to an image processing method and apparatus, a device, and a medium.

BACKGROUND OF THE DISCLOSURE

[0003] In a scenario of classifying a palm print image, the palm print image is usually first embedded by using a feature extraction network, to extract an embedding feature of the palm print image, and then the palm print image is classified by using the embedding feature.

[0004] The palm print image is photographed by using a camera, and the palm print image may include several regions with different definitions. In an existing application, a palm print image may be directly embedded by using a feature extraction network. In this process, the feature extraction network may pay too much attention to an image of an unclear region in the palm print image, and a feature of the image of the unclear region is usually inaccurate. Consequently, an embedding feature extracted from the palm print image is inaccurate, and further, a classification result of the palm print image is also inaccurate.

SUMMARY

[0005] This application provides an image processing method and apparatus, a device, and a medium, to improve accuracy of extracting an embedding feature of a palm print image, thereby improving accuracy of classifying the palm print image.

[0006] An aspect of this application provides an image processing method. The method includes:

[0007] obtaining a first sample image, the first sample image being divided into a plurality of regions, and the plurality of regions with different image definitions belonging to different definition categories;

[0008] invoking an attention mechanism network to perform attention degree recognition on the plurality of regions, to obtain a first region and a second region, wherein the attention mechanism network includes an attention parameter, and determines, based on the attention parameter, that an attention degree to the first region is higher than an attention degree to the second region;

[0009] adding a first predicted label to the first region and adding a second predicted label to the second region based on the respective attention degrees to the first region and the second region, the first predicted label indicating that the first region belongs to a first predicted definition category, the second predicted label indicating that the second region belongs to a second predicted definition category, and a definition indicated by the first predicted definition category being higher than a definition indicated by the second predicted definition category;

[0010] obtaining a first reference label of the first region and a second reference label of the second region, the first reference label indicating that the first region belongs to a first reference definition category, and the second reference label indicating that the second region belongs to a second reference definition category; and

[0011] adjusting the attention parameter based on a difference between the first predicted label and the first reference label and a difference between the second predicted label and the second reference label, to obtain a trained attention mechanism network, the trained attention mechanism network being configured to, based on the adjusted attention parameter, extract an embedding feature of a palm print image for identifying an owner of a palm print in the palm print image.

[0012] An aspect of this application provides an image processing apparatus. The apparatus includes:

[0013] a first obtaining module, configured to obtain a first sample image, the first sample image being divided into a plurality of regions, and the plurality of regions with different image definitions belonging to different definition categories;

[0014] an invoking module, configured to invoke an attention mechanism network to perform attention degree recognition on the plurality of regions, to obtain a first region and a second region, wherein the attention mechanism network includes an attention parameter, and determines,

based on the attention parameter, that an attention degree to the first region is higher than an attention degree to the second region;

[0015] an addition module, configured to add a first predicted label to the first region and add a second predicted label to the second region based on the respective attention degrees to the first region and the second region, the first predicted label indicating that the first region belongs to a first predicted definition category, the second predicted label indicating that the second region belongs to a second predicted definition category, and a definition indicated by the first predicted definition category being higher than a definition indicated by the second predicted definition category;

[0016] a second obtaining module, configured to obtain a first reference label of the first region and a second reference label of the second region, the first reference label indicating that the first region belongs to a first reference definition category, and the second reference label indicating that the second region belongs to a second reference definition category; and

[0017] an adjustment module, configured to adjust the attention parameter based on a difference between the first predicted label and the first reference label and a difference between the second predicted label and the second reference label, to obtain a trained attention mechanism network, the trained attention mechanism network being configured to, based on the adjusted attention parameter, extract an embedding feature of a palm print image for identifying an owner of a palm print in the palm print image.

[0018] An aspect of this application provides a computer device, including a memory and a processor, the memory having a computer program stored therein, and the computer program, when executed by the processor, causing the processor to perform the method according to an aspect in this application.

[0019] An aspect of this application provides a computer-readable storage medium, the computer-readable storage medium having a computer program stored therein, and the computer program, when executed by a processor, causing the processor to perform the method according to an aspect in this application.

[0020] An aspect of this application provides a computer program product. The computer program product includes a computer program. The computer program is stored in a computer-readable storage medium. A processor of a computer device reads the computer program from the computer-readable storage medium, and the processor executes the computer program, to

cause the computer device to perform the method provided in various optional implementations of the foregoing aspect.

[0021] The first sample image in this application may be divided into a plurality of regions subjected to division. For the plurality of regions, if image definitions of the regions are different, the regions may belong to different definition categories. In other words, in this application, the first sample image may be divided based on different definitions of images of respective parts of the first sample image. In this way, in this embodiment, an attention mechanism network can be invoked to perform attention degree recognition on the plurality of regions, to obtain a first region and a second region of the plurality of regions. The attention mechanism network includes an attention parameter, and determines, based on the attention parameter, that an attention degree to the first region is higher than an attention degree to the second region. Therefore, in this application, a first predicted label can be added to the first region and a second predicted label can be added to the second region based on the respective attention degrees to the first region and the second region. The first predicted label is used for indicating that the first region belongs to a first predicted definition category, the second predicted label is used for indicating that the second region belongs to a second predicted definition category, and a definition indicated by the first predicted definition category is higher than a definition indicated by the second predicted definition category. In other words, in this application, a region to which the attention mechanism network pays more attention may be marked with a predicted label indicating a higher image definition. In addition, in this application, a first reference label of the first region and a second reference label of the second region can be obtained. The first reference label is used for indicating a first reference definition category that the first region actually belongs to, and the second reference label is used for indicating a second reference definition category that the second region actually belongs to. Therefore, in this application, the attention parameter of the attention mechanism network can be adjusted based on a difference between the first predicted label and the first reference label and a difference between the second predicted label and the second reference label, to obtain a trained attention mechanism network,. Subsequently, the trained attention mechanism network can extract an embedding feature of a palm print image based on the adjusted attention parameter. The embedding feature can be used for identifying an owner of a palm print in the palm print image. It can be learned that, the method provided in this embodiment can add, by using an attention degree of the attention mechanism network to the region in the first sample image, the predicted label to the region of the first sample image, and a definition of a definition category indicated by a predicted label added for a region with a higher attention degree may be higher. Further, the attention parameter may be adjusted according to the difference between a real label (for example, the reference label) and the predicted

label of the region. Subsequently, the attention mechanism network can adopt a higher attention degree for a region with a higher definition in the palm print image by using the adjusted attention parameter, and adopt a lower attention degree for a region with a lower definition in the palm print image, so that an image feature of the region with the higher definition in the palm print image can be extracted to a greater extent, and finally a more accurate embedding feature of the palm print image is extracted. More accurate classification on the palm print image (that is, more accurate identification on an owner of the palm print in the palm print image) can be implemented by using the more accurate embedding feature of the palm print image.

BRIEF DESCRIPTION OF THE DRAWINGS

[0022] FIG. 1 is a schematic structural diagram of a network architecture for palm print image processing according to an embodiment of this application.

[0023] FIG. 2 is a schematic diagram of a scenario of adjusting a network parameter according to an embodiment of this application.

[0024] FIG. 3 is a schematic flowchart of an image processing method according to an embodiment of this application.

[0025] FIG. 4 is a schematic flowchart of training an image segmentation network according to an embodiment of this application.

[0026] FIG. 5 is a schematic diagram of a scenario of adjusting an attention parameter according to an embodiment of this application.

[0027] FIG. 6 is a schematic flowchart of a network training method according to an embodiment of this application.

[0028] FIG. 7 is a schematic diagram of a scenario of training a palm print classification network according to an embodiment of this application.

[0029] FIG. 8 is a schematic diagram of a scenario of generating a target embedding feature according to an embodiment of this application.

[0030] FIG. 9 is a schematic diagram of a scenario of palm print prediction according to an embodiment of this application.

[0031] FIG. 10 is a schematic structural diagram of an image processing apparatus according to an embodiment of this application.

[0032] FIG. 11 is a schematic structural diagram of a computer device according to an embodiment of this application.

DESCRIPTION OF EMBODIMENTS

[0033] This application relates to artificial intelligence-related technologies. Artificial intelligence (AI) is a theory, a method, a technology, and an application system that use a digital computer or a machine controlled by the digital computer to simulate, extend, and expand human intelligence, perceive an environment, obtain knowledge, and use knowledge to obtain an optimal result. In other words, the artificial intelligence is a comprehensive technology in computer science. The artificial intelligence attempts to understand an essence of intelligence, and produces a new intelligent machine that can react in a manner similar to the human intelligence. The artificial intelligence is to study the design principles and implementation methods of various intelligent machines, to enable the machines to have the functions of perception, reasoning, and decision-making.

[0034] The artificial intelligence technology is a comprehensive discipline, and relates to a wide range of fields, including both hardware-level technologies and software-level technologies. Basic technologies of the artificial intelligence usually include technologies such as a sensor, a dedicated artificial intelligence chip, cloud computing, distributed storage, big data processing technologies, an operating/interaction system, and electromechanical integration. The artificial intelligence software technologies mainly include several major directions such as a computer vision technology, a speech processing technology, a natural language processing technology, and machine learning/deep learning.

[0035] This application mainly relates to machine learning in artificial intelligence. Machine learning (ML) is a multi-domain interdisciplinary subject, relates to multi-domain subjects such as probability theory, statistics, approximation theory, convex analysis, and algorithm complexity theory, and specially studies how a computer simulates or implements a human learning behavior, to obtain new knowledge or skills, and reorganize an existing knowledge structure to continuously improve its performance. The machine learning, as a core of the artificial intelligence, is a fundamental way to make the computer intelligent, and is applied throughout various fields of the artificial intelligence. The machine learning and the deep learning usually include technologies such as an artificial neural network, a belief network, reinforcement learning, transfer learning, inductive learning, and demonstration learning.

[0036] Machine learning involved in this application mainly refers to how to obtain a palm print classification network through training, to accurately classify a palm print category of a palm print image by using the trained palm print classification network. For a specific procedure, refer to the related descriptions in the embodiment corresponding to FIG. 3.

[0037] First, all data (all relevant data such as palm print images) acquired in this application are acquired with the consent and authorization of an object (such as a user, an institution, or an enterprise) to which the data belongs, and the acquisition, use, and processing of relevant data need to comply with the relevant laws, regulations, and standards of relevant countries and regions.

[0038] Referring to FIG. 1, FIG. 1 is a schematic structural diagram of a network architecture for palm print image processing according to an embodiment of this application. As shown in FIG. 1, the network architecture may include a server 200 and a terminal device cluster. The terminal device cluster may include one terminal device or a plurality of terminal devices. A quantity of terminal devices is not limited herein. As shown in FIG. 1, the plurality of terminal devices may specifically include a terminal device 1, a terminal device 2, a terminal device 3, ..., and a terminal device n. As shown in FIG. 1, the terminal device 1, the terminal device 2, the terminal device 3, ..., and the terminal device n may all be in a network connection with the server 200, so that each terminal device can exchange data with the server 200 through the network connection.

[0039] The server 200 shown in FIG. 1 may be an independent physical server, a server cluster or distributed system including a plurality of physical servers, or a cloud server that provides a basic cloud computing service such as a cloud service, a cloud database, cloud computing, a cloud function, cloud storage, a network service, cloud communication, a middleware service, a domain name service, a security service, a content delivery network (CDN), big data, and an artificial intelligence platform. The terminal device may be an intelligent terminal such as a smartphone, a tablet computer, a notebook computer, a desktop computer, an in-vehicle terminal, or a smart television. Communication between the terminal device 1 and the server 200 is used as an example below to describe embodiments of this application in detail.

[0040] Referring to FIG. 2 together, FIG. 2 is a schematic diagram of a scenario of adjusting a network parameter according to an embodiment of this application. As shown in FIG. 2, a server 200 may adjust an attention parameter of an attention mechanism network by using a first sample image. The first sample image may include a plurality of regions (local images belonging to the first

sample image). The attention mechanism network may determine attention degrees of regions in an input image (for example, the first sample image) by using the attention parameter.

[0041] Therefore, the server 200 may add a predicted label for each region in the first sample image by using the attention parameter of the attention mechanism network. The predicted label may be a label of a definition category of each region in the first sample image and determined by using the attention parameter of the attention mechanism network. For example, a region that the attention mechanism network pays more attention may be marked with a label of a definition category of a higher definition.

[0042] The server 200 may further obtain a reference label of each region in the first sample image. The reference label may be used for indicating an actual definition category of each region in the first sample image.

[0043] Therefore, the server 200 may adjust the attention parameter of the attention mechanism network based on a difference between the predicted label and the reference label of each region in the first sample image, so that the attention mechanism network can pay more attention to a feature of an image of a clearer region in the input image by using the adjusted attention parameter.

[0044] Subsequently, the attention mechanism network may be used as a network for extracting a feature of a palm print image in a palm print identification scenario. The attention mechanism network may perform embedding on an input palm print image by using the attention parameter adjusted according to the foregoing manner, to generate an accurate embedding feature of the palm print image, and then accurate classification of a palm print category of the palm print image can be implemented by using the embedding feature. The palm print image may be obtained by the terminal device 1 by photographing a palm print of a user. The terminal device 1 may send the acquired palm print image to the server 200, to request the server 200 to classify a palm print category of the palm print image, so that the server 200 can invoke the attention mechanism network to classify the palm print category of the palm print image by using the adjusted attention parameter. For a specific procedure, refer to the related descriptions in the following embodiments corresponding to FIG. 3 and FIG. 6.

[0045] By using the method in this embodiment of this application, the attention mechanism network can pay more attention to an image feature of a clearer region in an input palm print image, so that a more accurate embedding feature of the palm print image can be extracted, and more accurate classification on the palm print category of the palm print image can also be implemented by using the accurate embedding feature of the palm print image.

[0046] Referring to FIG. 3, FIG. 3 is a schematic flowchart of an image processing method according to an embodiment of this application. An execution body in this embodiment of this application may be an image processing device. The image processing device may be a computer device or a computer device cluster including a plurality of computer devices. The computer device may be a server, a terminal device, or another device. This is not limited. As shown in FIG. 3, the method may include the following operations.

[0047] Operation S101: Obtain a first sample image, where the first sample image is divided into a plurality of regions, and the plurality of regions with different image definitions belong to different definition categories.

[0048] In an embodiment, the image processing device may obtain the first sample image. The first sample image may include a plurality of regions. In other words, the first sample image may be divided into the plurality of regions, and each region may be a local image in the first sample image. In other words, the first sample image may be divided into a plurality of image blocks (namely, the plurality of regions), a size of each image block may be determined according to an actual application scenario (may be a preset division size), and sizes of the image blocks may be the same.

[0049] The plurality of regions may include regions having the same definition, or may include regions having different definitions (namely, image definitions). The definition of the image may refer to sharpness of change of an edge of image details, in other words, clarity of the image details and boundaries thereof. In this application, if image definitions of the regions are different, it may be considered that definition categories of the regions are different. The definition category of the image may be understood as a definition level of the image. The definition level may refer to a definition range obtained by dividing the definition of the image, and one definition level may correspond to one definition range. Each definition category of the image may be used for indicating a definition (an image definition) corresponding to the definition category, for example, used for indicating a definition level corresponding to the definition category. A higher definition level of the image indicates a higher definition indicated by the image (that is, the image is clearer). Otherwise, a lower definition level of the image indicates a lower definition indicated by the image (that is, the image is less clear).

[0050] For example, there may be at least two definition categories of the image. The two may include a blurry definition category and a clear (or high definition) definition category. As the name implies, a definition (blurry definition) indicated by the blurry definition category is lower than a definition (clear definition) indicated by the clear definition category. Alternatively, on this

basis, the definition category may be classified into finer definition categories based on definition degrees. For example, definition categories of the image may include a blurry (for example, low definition) definition category, a standard definition category, a high definition category, an ultra-high definition category, and the like. Definitions indicated by the definition categories may be sequentially ascending. For example, a definition indicated by the standard definition category is higher than a definition indicated by the blurry definition category, a definition indicated by the high definition category is higher than the definition indicated by the standard definition category, and a definition indicated by the ultra-high definition category is higher than the definition indicated by the high definition category.

[0051] The categories of the definition categories of the image and specific definition categories included may be set arbitrarily according to an actual application scenario, and this is not limited. That the image definitions of the regions are different means that the regions have different definition levels.

[0052] In other words, the first sample image may include regions of at least two definition categories, and one region may have one definition category. A specific quantity of first sample images may be determined according to an actual application scenario.

[0053] Operation S102: Invoke an attention mechanism network to perform attention degree recognition on the plurality of regions, to obtain a first region and a second region, where the attention mechanism network includes an attention parameter, and determines, based on the attention parameter, that an attention degree to the first region is higher than an attention degree to the second region.

[0054] In an embodiment of this application, the attention mechanism network may perform attention degree recognition on the input image (for example, the first sample image) in a unit of each region obtained by dividing the input image. Therefore, the image processing device may invoke the attention mechanism network to perform identification (for example, attention degree recognition) on the plurality of regions of the first sample image, to obtain the first region and the second region of the plurality of regions.

[0055] The attention mechanism network is a network (namely, a model) having an attention mechanism. A core idea of the attention mechanism is to simulate an attention process of human beings on an input of the network. To be specific, the network can automatically determine which part of input data needs to be concerned about when data processing is performed, so that the network can effectively capture key information of the input data. For example, the attention mechanism network includes, but is not limited to, a Transformer (a neural network based on a

self-attention mechanism), a GAT (a graph attention mechanism network), and the like. The attention mechanism network may include an attention parameter (which belongs to a weight parameter, and is a network parameter of the attention mechanism network). The attention mechanism network may determine, by using the attention parameter, that an attention degree of the first region is higher than an attention degree (that is, a concern degree) of the second region, or in other words, the attention degree paid to the first region by the attention mechanism network is higher than attention degree paid to the second region. The attention degree paid to each region by the attention mechanism network may be determined by using the attention parameter of the attention mechanism network, as described in the following content.

[0056] The attention mechanism network may perform, by using the foregoing attention parameter, identification (that is, feature learning) on the first sample image in a unit of each region of the first sample image, to identify (that is, learn) an attention weight of the attention mechanism network for each region. The attention mechanism network may have one attention weight for one region, and the attention weight of the attention mechanism network for the region may be used for reflecting an attention degree of the attention mechanism network to the region. A higher attention weight indicates a higher attention degree. Otherwise, a lower attention weight indicates a lower attention degree. A value range of the attention weight may be $[0, 1]$.

[0057] It can be learned from the above that, the attention mechanism network may obtain, by using the attention parameter and in a unit of each image block (that is, each region) of the first sample image, the attention weight of each image block of the first sample image. In other words, which image block of the input image (for example, the first sample image) is paid more attention by the attention mechanism network and which image block of the input image is paid less attention by the attention mechanism network can be learned by using the attention parameter of the attention mechanism network.

[0058] Optionally, the value range $[0, 1]$ of the attention weight may be divided based on a quantity of definition categories of the image, to obtain a plurality of weight ranges of the attention weight through division. One definition category may correspond to one weight range of the attention weight. A higher definition indicated by a definition category indicates a higher weight value in the weight range corresponding to the definition category. A combination of all weight ranges may be the entire value range $[0, 1]$ of the attention weight.

[0059] For example, if there are two definition categories of the image, including a blurry definition category and a clear definition category, the value range of the attention weight may be divided into two weight ranges, for example, divided into a weight range $[0, 0.5)$ and a weight range

[0.5, 1]. The weight range [0, 0.5) may be a weight range corresponding to the blurry definition category, the weight range [0.5, 1] may be a weight range corresponding to the clear definition category, and the weight ranges do not overlap each other.

[0060] Therefore, the attention weight of the first region and the attention weight of the second region may be respectively in the two weight ranges obtained through division, and the attention weight of the first region is greater than the attention weight of the second region, in other words, weight values in a weight range to which the attention weight of the first region belongs are greater than weight values in a weight range to which the attention weight of the second region belongs.

[0061] Any two weight ranges of the plurality of weight ranges obtained by dividing the value range of the attention weight may be referred to as a first weight range and a second weight range, and weight values in the first weight range are greater than weight values in the second weight range. For example, the weight range [0.5, 1] may be the first weight range, and the weight range [0, 0.5) may be the second weight range.

[0062] Therefore, in this application, a region of the plurality of regions of the first sample image and whose attention weight is in a first weight range can be used as the first region, and a region of the plurality of regions whose attention weight is in a second weight range can be used as the second region. There may be one or more first regions and second regions, and specific quantities of the first regions and the second regions may be determined according to an actual application scenario.

[0063] In this embodiment, regions (to be specific, the first region and the second region) of two definition categories are used as an example for description. Actually, regions of all definition categories may be (or need to be) processed in a same processing manner, to adjust the attention parameter of the attention mechanism network. For example, when an image has more than two definition categories, there may also be more than two weight ranges. The attention weight of the attention mechanism network for the first region and the attention weight of the attention mechanism network for the second region may be respectively in different weight ranges. The first region and the second region may be regions whose attention weights are in any two weight ranges of the more than two weight ranges. In addition, weight values in the weight range to which the attention weight of the attention mechanism network for the first region belongs is greater than weight values in the weight range to which the attention weight of the attention mechanism network for the second region belongs. In other words, regions with attention weights falling within two weight ranges of the more than two weight ranges may be respectively used as the corresponding

first region and second region, to perform the related processing in this embodiment of this application.

[0064] For example, if there are three definition categories of the image, including a blurry definition category, a standard definition category, and a clear definition category, the value range of the attention weight may be divided into three weight ranges, for example, may be divided into a weight range $[0, 0.33)$, a weight range $[0.33, 0.66)$, and a weight range $[0.66, 1]$. The weight range $[0, 0.33)$ may correspond to the blurry definition category, the weight range $[0.33, 0.66)$ may correspond to the standard definition category, and the weight range $[0.66, 1]$ may correspond to the clear definition category.

[0065] For a region whose attention weight is within the weight range $[0, 0.33)$ and a region whose attention weight is within the weight range $[0.33, 0.66)$, the region whose attention weight is within the weight range $[0, 0.33)$ may be used as the second region, and the region whose attention weight is within the weight range $[0.33, 0.66)$ may be used as the first region, to perform the related processing described in this embodiment of this application.

[0066] For a region whose attention weight is within the weight range $[0.33, 0.66)$ and a region whose attention weight is within the weight range $[0.66, 1]$, the region whose attention weight is within the weight range $[0.33, 0.66)$ may be used as the second region, and the region whose attention weight is within the weight range $[0.66, 1]$ may be used as the first region, to perform the related processing described in this embodiment of this application.

[0067] Moreover, for a region whose attention weight is within the weight range $[0, 0.33)$ and a region whose attention weight is within the weight range $[0.66, 1]$, the region whose attention weight is within the weight range $[0, 0.33)$ may be used as the second region, and the region whose attention weight is within the weight range $[0.66, 1]$ may be used as the first region, to perform the related processing described in this embodiment of this application.

[0068] When there are more than three definition categories of the image, corresponding processing may also be performed according to the foregoing principle. A specific quantity of definition categories of the image and a quantity of weight ranges that need to be obtained through division can both be determined according to an actual application scenario. This is not limited in this application.

[0069] Operation S103: Add a first predicted label to the first region and add a second predicted label to the second region based on the respective attention degrees to the first region and the second region, where the first predicted label is used for indicating that the first region belongs

to a first predicted definition category, the second predicted label is used for indicating that the second region belongs to a second predicted definition category, and a definition indicated by the first predicted definition category is higher than a definition indicated by the second predicted definition category.

[0070] In an embodiment, the image processing device may add a corresponding predicted label for each region by using a weight range to which an attention weight of the attention mechanism network for each region belongs and a definition category corresponding to each weight range, as described in the following content.

[0071] The image processing device may add the first predicted label to the first region by using the attention weight (namely, the attention degree) of the attention mechanism network to the first region. For example, the image processing device may add the first predicted label for the first region by using the definition category corresponding to the weight range to which the attention weight of the first region belongs. The first predicted label may be used for indicating that the definition category of the first region belongs to the first predicted definition category, and the first predicted definition category is a definition category corresponding to the weight range to which the attention weight of the first region belongs. The first predicted definition category may be understood as a definition category determined (that is, predicted) by the attention mechanism network for the first region based on the attention parameter.

[0072] Similarly, the image processing device may add the second predicted label to the second region by using the attention weight (namely, the attention degree) of the attention mechanism network to the second region. For example, the image processing device may add the second predicted label for the second region by using the definition category corresponding to the weight range to which the attention weight of the second region belongs. The second predicted label may be used for indicating that the definition category of the second region belongs to the second predicted definition category, and the second predicted definition category is a definition category corresponding to the weight range to which the attention weight of the second region belongs. The second predicted definition category may be understood as a definition category determined (that is, predicted) by the attention mechanism network for the second region based on the attention parameter.

[0073] In addition, an image definition indicated by the first predicted definition category is higher than an image definition indicated by the second predicted definition category.

[0074] By using the foregoing process of adding the predicted label (including the first predicted label and the second predicted label) for the first region and the second region, in this

embodiment, a predicted label of a definition category indicating a higher definition may be added for a region that the attention mechanism network pays more attention. Subsequently, the attention parameter of the attention mechanism network is adjusted in this manner, so that the attention mechanism network may pay more attention to a clearer region in the input image and pay less attention to a blurrier region in the input image based on the adjusted attention parameter.

[0075] Operation S104: Obtain a first reference label of the first region and a second reference label of the second region, where the first reference label is used for indicating that the first region belongs to a first reference definition category, and the second reference label is used for indicating that the second region belongs to a second reference definition category.

[0076] In an embodiment, the image processing device may obtain the reference label of the first region and the reference label of the second region. The reference label of the first region may be referred to as the first reference label, and the reference label of the second region may be referred to as the second reference label.

[0077] The first reference label may be a label of an actual definition category of the first region, and the second reference label may be a label of an actual definition category of the second region.

[0078] The first reference label may be used for indicating that the definition category of the first region belongs to the first reference definition category, that is, the actual definition category of the first region may be the first reference definition category. Similarly, the second reference label may be used for indicating that the definition category of the second region belongs to the second reference definition category, that is, the actual definition category of the second region may be the second reference definition category.

[0079] The first reference definition category and the second reference definition category each may be any one of the at least two definition categories set for the image.

[0080] In an embodiment, the first reference label of the first region and the second reference label of the second region may be obtained by using a trained definition classification network, as described in the following content:

[0081] The image processing device may invoke the trained definition classification network to perform classification prediction on a definition of the first region, to obtain the first reference definition category of the first region. The first reference definition category is an actual definition category of the first region predicted by using the trained definition classification network.

[0082] Similarly, the image processing device may invoke the trained definition classification network to perform classification prediction on a definition of the second region, to obtain the second reference definition category of the second region. The second reference definition category is an actual definition category of the second region predicted by using the trained definition classification network.

[0083] Therefore, the image processing device can add the first reference label for the first region by using the first reference definition category. The first reference label is used for indicating that the definition category of the first region is the first reference definition category.

[0084] The image processing device may further add the second reference label for the second region by using the second reference definition category. The second reference label is used for indicating that the definition category of the second region is the second reference definition category.

[0085] For example, a process of obtaining the trained definition classification network through training may include: The image classification network may obtain a third sample image and a definition classification network that needs to be trained, where the third sample image has a definition label, and the definition label is used for indicating a real definition category of the third sample image (that is, an actual definition category of the third sample image). In an embodiment of this application, there may be several third sample images. The several third sample images may include sample images of various definition categories (including all definition categories preconfigured for the image), and are used for training the definition classification network that needs to be trained.

[0086] Further, the image processing device may invoke the definition classification network that needs to be trained, and perform classification prediction on the definition category of the third sample image, to obtain a predicted definition category of the third sample image. The predicted definition category is the definition category of the third sample image obtained through prediction by using the definition classification network that needs to be trained.

[0087] Then, the image processing device may adjust a network parameter of the definition classification network that needs to be trained based on a difference (which may be reflected by a cross entropy loss between the real definition category and the predicted definition category) between the real definition category and the predicted definition category of the third sample image, to obtain the trained definition classification network.

[0088] Moreover, the image processing device may further invoke a trained image segmentation network, to obtain the first reference label of the first region and the second reference label of the second region. The process may include the following operations:

[0089] The trained image segmentation network is a network obtained through training and that can be used for segmenting parts of images with different definition categories in the input image. For example, the trained image segmentation network may include, but is not limited to, U-Net (a convolutional neural network). Therefore, the image processing device may invoke the trained image segmentation network to perform image segmentation on the first sample image, to obtain a plurality of segmented images of the first sample image. Each segmented image may be a local image in the first sample image. One segmented image may correspond to one definition category, that is, one segmented image has one definition category. The definition category corresponding to each segmented image may alternatively be obtained through identification by the trained image segmentation network when performing segmentation to obtain the segmented image.

[0090] Further, the image processing device may use a segmented image having the highest coincidence degree (that is, having the largest image coincidence regions) with the first region in the plurality of segmented images as a first definition matching image of the first region, and further, may add the first reference label for the first region based on a definition category corresponding to the first definition matching image. The first reference definition category indicated by the first reference label may be the definition category corresponding to the first definition matching image.

[0091] Similarly, the image processing device may use a segmented image having the highest coincidence degree (that is, having the largest image coincidence regions) with the second region in the plurality of segmented images as a second definition matching image of the second region, and further, may add the second reference label for the second region based on a definition category corresponding to the second definition matching image. The second reference definition category indicated by the second reference label may be the definition category corresponding to the second definition matching image.

[0092] In an embodiment, the image processing device may use either of the foregoing two manners (the manner based on the trained definition classification network and the manner based on the trained image segmentation network) to pre-add a corresponding reference label for each region in the first sample image, so that in operation S104, the reference label (including the first reference label and the second reference label) pre-added for the first region and the second region can be directly obtained.

[0093] Referring to FIG. 4, FIG. 4 is a schematic flowchart of training an image segmentation network according to an embodiment of this application. As shown in FIG. 4, the procedure may include the following operations.

[0094] 1. It is assumed that definition categories set for an image includes a clear definition category and a blurry definition category, and therefore, the image processing device may prepare a sample image including a clear region label and a blurry region label. The clear region label may be a label for an image of a clear region in the sample image, and the clear region label is a label of the clear definition category. The blurry region label may be a label for an image of a blurry region in the sample image, and the blurry region label is a label of the blurry definition category.

[0095] 2. The image processing device may construct an image segmentation network. For example, the image processing device may construct a U-Net network as the image segmentation network.

[0096] 3. The image processing device may define a loss function of the image segmentation network. For example, the loss function may be a cross-entropy loss between a predicted clear region and a predicted fuzzy region and an actual clear region and an actual blurry region indicated by region labels (including the clear region label and the blurry region label) marked for the sample image.

[0097] 4. The image processing device may perform initialization on a network parameter of the image segmentation network, for example, perform random initialization, to further start training the image segmentation network.

[0098] 5. The image processing device may obtain a sample (belonging to the sample image prepared in operation 1) for training the image segmentation network in a current round.

[0099] 6. The image processing device may input an obtained sample into the image segmentation network for forward propagation, to predict and identify a clear region and a blurry region in the input sample image.

[00100] 7. The image processing device may calculate the loss function of the image segmentation network based on a difference (which may be reflected by a cross-entropy loss) between a clear region obtained through prediction and identification and the actual clear region marked by the clear region label, and a difference (which may be reflected by a cross-entropy loss) between a blurry region obtained through prediction and identification and the actual blurry region marked by the blurry region label.

[00101] 8. The image processing device may perform back propagation in the image segmentation network by using the calculated loss function, to adjust (that is, update) the network parameter of the image segmentation network.

[00102] 9. The image processing device may determine whether a stop condition is currently met. The stop condition may be a condition for determining whether the image segmentation network is trained. For example, the stop condition may be that the network parameter of the image segmentation network is trained to a convergent state, or the stop condition may be that a round number of iterative training on the image segmentation network is greater than or equal to a quantity threshold, and the like. Therefore, if the stop condition is met, the image processing device may perform the following operation 10, or if the stop condition is not met, the image processing device may repeatedly perform the foregoing operation 5, to start a next round of iterative training on the image segmentation network.

[00103] 10. Use the image segmentation network obtained through training at this time as the trained image segmentation network.

[00104] Operation S105: Adjust the attention parameter based on a difference between the first predicted label and the first reference label and a difference between the second predicted label and the second reference label, to obtain a trained attention mechanism network, where the trained attention mechanism network is configured to, based on the adjusted attention parameter, extract an embedding feature of a palm print image for identifying an owner of a palm print in the palm print image.

[00105] Optionally, the image processing device may adjust the attention parameter of the attention mechanism network based on the difference between the first predicted label and the first reference label, and the difference between the second predicted label and the second reference label. The difference between the first predicted label and the first reference label may be reflected by a cross-entropy loss between the first predicted label and the first reference label, and the difference between the second predicted label and the second reference label may be reflected by a cross-entropy loss between the second predicted label and the second reference label.

[00106] In other words, the attention parameter of the attention mechanism network may be modified based on the cross-entropy loss between the first predicted label and the first reference label and the cross-entropy loss between the second predicted label and the second reference label (for example, a sum of the two cross-entropies). An objective of the adjustment may be to make the sum of the cross-entropy tend to be a minimum value (for example, 0).

[00107] Referring to FIG. 5, FIG. 5 is a schematic diagram of adjusting an attention parameter according to an embodiment of this application. As shown in FIG. 5, the first region may have a first predicted label and a first reference label, and the second region may have a second predicted label and a second reference label.

[00108] The image processing device may obtain a difference between the first predicted label and the first reference label of the first region. For example, the difference may be a cross-entropy loss between the first predicted label and the first reference label. In other words, the cross-entropy loss may be used for reflecting the difference between the first predicted label and the first reference label. The image processing device may further obtain a difference between the second predicted label and the second reference label of the second region. For example, the difference may be a cross-entropy loss between the second predicted label and the second reference label. In other words, the cross-entropy loss may be used for reflecting/representing the difference between the second predicted label and the second reference label.

[00109] Further, the image processing device may adjust the attention parameter of the attention mechanism network by using the cross-entropy loss between the first predicted label and the first reference label, and the cross-entropy loss between the second predicted label and the second reference label that are obtained above, to obtain the adjusted attention parameter.

[00110] Based on the foregoing principle, the attention mechanism network, by using the adjusted attention parameter, can pay more attention to a region having a higher definition in an input image (such as a palm print image), and pay less attention to a region having a lower definition in the input image. Therefore, subsequently, the attention mechanism network can more accurately extract an embedding feature of the palm print image by using the adjusted attention parameter. The accurate embedding feature can be used for more accurately classifying a palm print category (such as an owner of a palm print in the palm print image) of the palm print image. The process of classifying the palm print category of the palm print image is a process of identifying the owner of the palm print in the palm print image (for example, identifying a user to which the palm print in the palm print image specifically belongs, to determine an identity of the owner of the palm print in the palm print image). For the specific process, refer to the related descriptions in the following embodiment corresponding to FIG. 6.

[00111] In this application, a first sample image can be obtained. The first sample image is divided into a plurality of regions, and the plurality of regions with different image definitions belong to different definition categories. In addition, an attention mechanism network can be invoked to perform attention degree recognition on the plurality of regions, to obtain a first region

and a second region of the plurality of regions. The attention mechanism network includes an attention parameter, and determines, based on the attention parameter, that an attention degree to the first region is higher than an attention degree to the second region. Further, a first predicted label can be added to the first region and a second predicted label can be added to the second region based on the respective attention degrees of the attention mechanism network to the first region and the second region. The first predicted label is used for indicating that the first region belongs to a first predicted definition category, the second predicted label is used for indicating that the second region belongs to a second predicted definition category, and a definition indicated by the first predicted definition category is higher than a definition indicated by the second predicted definition category. A first reference label of the first region and a second reference label of the second region can be obtained. The first reference label is used for indicating that the first region actually belongs to a first reference definition category, and the second reference label is used for indicating that the second region actually belongs to a second reference definition category. Therefore, the attention parameter can be adjusted based on a difference between the first predicted label and the first reference label and a difference between the second predicted label and the second reference label. The attention mechanism network is configured to extract an embedding feature of a palm print image based on the adjusted attention parameter. The embedding feature is used for identifying an owner of a palm print in the palm print image. It can be learned that, the method provided in this embodiment can add, by using an attention degree of the attention mechanism network to the region in the first sample image, the predicted label to the region of the first sample image, and a definition of a definition category indicated by a predicted label added for a region with a higher attention degree may be higher. Further, the attention parameter may be adjusted according to the difference between a real label (for example, the reference label) and the predicted label of the region. Subsequently, the attention mechanism network can adopt a higher attention degree for a region with a higher definition in the palm print image by using the adjusted attention parameter, and adopt a lower attention degree for a region with a lower definition in the palm print image, so that an image feature of the region with the higher definition in the palm print image can be extracted to a greater extent, and finally a more accurate embedding feature of the palm print image is extracted. More accurate classification on the palm print image (that is, more accurate identification on an owner of the palm print in the palm print image) can be implemented by using the more accurate embedding feature of the palm print image.

[00112] Referring to FIG. 6, FIG. 6 is a schematic flowchart of a network training method according to an embodiment of this application. This application may be applied to a palm print recognition scenario. As shown in FIG. 6, the method may include the following operations:

[00113] Operation S201: Obtain a second sample image, where the second sample image is obtained by photographing a sample palm print, and has a palm print label indicating a real category of the sample palm print.

[00114] In an embodiment, the image processing device may further obtain the second sample image. The second sample image may be obtained by photographing the sample palm print (which may be a palm print of any user). The second sample image may have the palm print label. The palm print label may be used for indicating the real category of the sample palm print. The real category may be used for indicating an actual owner of the sample palm print (for example, a user to which the sample palm print actually belongs, that is, a user of the palm print).

[00115] The attention mechanism network may be included in a palm print classification network. The attention parameter (namely, the adjusted attention parameter) obtained by adjusting the attention parameter of the attention mechanism network by using the process described in the foregoing embodiment corresponding to FIG. 3 may remain unchanged in the palm print classification network, so that in a subsequent process of training the palm print classification network, the adjusted attention parameter does not need to be adjusted again. The palm print classification network may be a network configured to identify (that is, classify) the palm print in the palm print image, to identify the owner of the palm print in the palm print image (for example, the user to which the palm print belongs). For example, the palm print classification network includes, but is not limited to, a decision tree model, a support vector machine model, and the like.

[00116] In other words, the process in the foregoing embodiment corresponding to FIG. 3 may be understood as a previous operation of adjusting the attention parameter of the attention mechanism network in the palm print classification network. After the attention parameter of the attention mechanism network in the palm print classification network is adjusted, overall training may be performed on the palm print classification network. In the training process, the adjusted attention parameter may not be updated, but a network parameter other than the adjusted attention parameter in the palm print classification network may be updated, as described in the following content.

[00117] Operation S202: Invoke the trained attention mechanism network to perform embedding on the second sample image based on the adjusted attention parameter, to generate a sample embedding feature of the second sample image.

[00118] In an embodiment, the attention mechanism network may belong to a sub-network in the palm print classification network that is configured to extract a feature from the input image. In addition to the adjusted attention parameter, network parameters of the attention mechanism

network may further include another network parameter used for performing embedding (that is, feature extraction) on the input image.

[00119] Therefore, the image processing device may invoke the foregoing attention mechanism network in the palm print classification network, perform embedding (that is, feature extraction) on the second sample image (to be specific, the input image of the attention mechanism network in this case) based on the foregoing adjusted attention parameter and the another network parameter used for performing embedding on the input image, and may generate the sample embedding feature of the second sample image. In an embodiment, the sample embedding feature may be a feature vector. The sample embedding feature is a feature obtained by the attention mechanism network by learning the second sample image.

[00120] In an embodiment, there may be a plurality of second sample images, the plurality of second sample images may be obtained by photographing the sample palm print from a plurality of photographing angles, and one photographing angle is used for photographing to obtain one second sample image of the sample palm print. Therefore, the process of generating the sample embedding feature in this application may include the following operations:

[00121] The image processing device may invoke the attention mechanism network to separately perform embedding on each second sample image based on the adjusted attention parameter and the another network parameter used for performing embedding on the input image, to generate an image embedding feature of each second sample image. One second sample image may have an image embedding feature, and the image embedding feature of the second sample image is a feature learned by the attention mechanism network from the second sample image.

[00122] Further, the image processing device may fuse image embedding features of a plurality of second sample images, to generate the sample embedding feature. The sample embedding feature may be a feature obtained by fusing features of images (for example, the plurality of second sample images) of the sample palm print photographed from a plurality of angles.

[00123] Because the second sample images photographed from the photographing angles may include images of different clear regions of the sample palm print, the sample embedding feature is obtained by fusing the second sample images obtained from the photographing angles, so that the sample embedding feature includes features of the images of the clear regions of the sample palm print from the different photographing angles. In addition, the features of the images of the clear regions are more accurate than features of images of blurry regions. Therefore, the sample embedding feature generated in this manner is more accurate, and more accurate classification on

the palm print category of the second sample image can also be implemented by using the sample embedding feature subsequently.

[00124] In an embodiment, the manner of fusing the image embedding features of the plurality of second sample images includes, but is not limited to, the following manner: stitching (for example, horizontal stitching) the image embedding features of the plurality of second sample images, to generate the sample embedding feature; or summing the image embedding features of the plurality of second sample images, to generate the sample embedding feature. Optionally, the generated sample embedding feature may also be a feature vector.

[00125] The process of summing the image embedding features of the plurality of second sample images may include: Dimensions of the image embedding features of the second sample images may be the same, and therefore, feature values on a same position in the image embedding features of the second sample images may be added together, to generate the sample embedding feature.

[00126] Operation S203: Invoke a classification sub-network to perform classification prediction on the sample palm print based on the sample embedding feature, to obtain a classification predicted result of the sample palm print.

[00127] In an embodiment, the palm print classification network may further include the classification sub-network. The classification sub-network may be configured to classify a palm print in the palm print image, to output a probability (which may constitute a probability distribution) that the palm print in the palm print image belongs to each user (which may be referred to as an object) in a user set (which may be a full set of users that need to be identified in an application scenario of palm print identification). In an embodiment, the classification sub-network may be a fully connected network.

[00128] Therefore, the palm print classification network may invoke the classification sub-network to perform classification prediction on the sample palm print by using the foregoing sample embedding feature, to obtain the classification predicted result of the sample palm print. The classification predicted result may be a probability distribution (which may be represented as a vector) formed by the predicted probability that the sample palm print belongs to each of the users in the user set.

[00129] Operation S204: Adjust, based on a difference between the real category and the classification predicted result, a network parameter in the palm print classification network other than the adjusted attention parameter, to obtain a trained palm print classification network.

[00130] In an embodiment, the image processing device may adjust network parameters (which may include a network parameter used for embedding the image and a network parameter of the classification sub-network that are of the attention mechanism network and other than the adjusted attention parameter) other than the adjusted attention parameter in the palm print classification network by using a difference (which may be represented by a cross-entropy loss) between the real category (which may also be represented as a vector, and in the vector, a probability at a position of a user indicated by the real category may be 1, and a probability at a position of another user may be 0) and the classification predicted result of the sample palm print, to obtain the trained palm print classification network.

[00131] For example, iterative training may be continuously performed on the palm print classification network by using the foregoing process, until the network parameters (the network parameters other than the adjusted attention parameter) of the palm print classification network reaches a convergent state, or a quantity of iterative training times on the palm print classification network is greater than or equal to a quantity threshold. In this case, the palm print classification network obtained through training may be used as the trained palm print classification network.

[00132] An attention mechanism network in the trained palm print classification network is the trained attention mechanism network, and a classification sub-network in the trained palm print classification network is the trained classification sub-network.

[00133] Referring to FIG. 7. FIG. 7 is a schematic diagram of a scenario of training a palm print classification network according to an embodiment of this application. As shown in FIG. 7, first, the image processing device may adjust the attention parameter of the attention mechanism network in the palm print classification network by using the first sample image.

[00134] Further, network parameters (including feature embedding features of the attention mechanism network other than the adjusted attention parameter, and the network parameter of the classification sub-network used for performing palm print classification) other than the adjusted attention parameter in the palm print classification network may be adjusted by using the second sample image, and finally the trained palm print classification network may be obtained.

[00135] Through the foregoing process, the palm print classification network can be obtained through training, and the trained palm print classification network may include a trained attention mechanism network and a trained classification sub-network. Subsequently, accurate classification on the palm print image can be implemented by using the trained palm print classification network, as described in the following content.

[00136] The image processing device may obtain a to-be-classified palm print image, and the palm print image may be obtained by photographing a target palm print (which may be a palm print of any user). In an embodiment, there may alternatively be a plurality of to-be classified palm print images, the plurality of palm print images may be obtained by photographing the target palm print from a plurality of photographing angles, and one photographing angle may be used for photographing to obtain one palm print image of the target palm print.

[00137] The image processing device may invoke the trained attention mechanism network to perform embedding on the to-be-classified palm print image, to generate a target embedding feature of the to-be-classified palm print image. A generation principle of the target embedding feature is the same as that of the sample embedding feature, and may include:

[00138] The image processing device may invoke the trained attention mechanism network to separately perform embedding on each to-be-classified palm print image, to generate an image embedding feature of the palm print image, and may further fusing (for example, stitching or summing) image embedding features of a plurality of to-be-classified palm print images, to generate the target embedding feature.

[00139] Because the trained attention mechanism network includes the adjusted attention parameter, the trained attention mechanism network can pay more attention to an image of a clear region in the palm print image by using the adjusted attention parameter. Therefore, a more accurate target embedding feature of the to-be-classified palm print image can be generated.

[00140] Referring to FIG. 8, FIG. 8 is a schematic diagram of a scenario of generating a target embedding feature according to an embodiment of this application. As shown in FIG. 8, the to-be-classified palm print images may include a palm print image 1 to a palm print image m obtained by photographing from a plurality of photographing angles, and the image processing device may invoke the trained attention mechanism network to perform embedding on the to-be-classified palm print images, to generate an image embedding feature 1 of the palm print image 1, an image embedding feature 2 of a palm print image 2, an image embedding feature 3 of a palm print image 3, ..., and an image embedding feature m of the palm print image m.

[00141] Further, feature fusion may be performed on the image embedding feature 1, the image embedding feature 2, the image embedding feature 3, ..., and the image embedding feature m that are generated, to generate the target embedding feature.

[00142] Further, the image processing device may invoke the trained classification sub-network to perform classification prediction on the target palm print by using the generated

target embedding feature, to obtain a target category of the target palm print. The target category may represent a user corresponding to a maximum probability predicted by the trained classification sub-network in probabilities that the target palm print belongs to users in the user set. The user is an owner of the predicted target palm print, that is, the target category may be used for indicating the owner of the target palm print obtained through classification.

[00143] Referring to FIG. 9, FIG. 9 is a schematic diagram of a scenario of palm print prediction according to an embodiment of this application. As shown in FIG. 9, the image processing device may input the to-be-classified palm print image to the trained palm print classification network, to invoke the trained attention mechanism network in the trained palm print classification image to extract the target embedding feature of the to-be-classified palm print image.

[00144] Further, the trained classification sub-network in the trained palm print classification image may be invoked to obtain the target category of the target palm print in the to-be-classified palm print image through classification prediction by using the generated target embedding feature.

[00145] In a feasible implementation, the palm print image may be a palm print image photographed in real time by using a palm print capturing device (for example, a device that initiates an execution request for a target service) after a service initiator (which may be a service account, and the service account may belong to a service user) initiates (for example, initiates in a service platform) the execution request for the target service.

[00146] Therefore, if the target category is used for indicating that the owner of the target palm print is the service initiator (that is, a service user to which the service initiator belongs), the target service may be executed (dealt) for the service initiator, that is, the target service is provided for the service initiator.

[00147] Through the foregoing process, after accurate classification and identification of the palm print image is implemented, corresponding security processing can be performed on a related service.

[00148] Referring to FIG. 10, FIG. 10 is a schematic structural diagram of an image processing apparatus according to an embodiment of this application. As shown in FIG. 10, the image processing apparatus 1 may include: a first obtaining module 11, an invoking module 12, an addition module 13, a second obtaining module 14, and an adjustment module 15.

[00149] The first obtaining module 11 is configured to obtain a first sample image, the first sample image being divided into a plurality of regions, and the plurality of regions with different image definitions belonging to different definition categories.

[00150] The invoking module 12 is configured to invoke an attention mechanism network to perform attention degree recognition on the plurality of regions, to obtain a first region and a second region, the attention mechanism network including an attention parameter, and determining, based on the attention parameter, that an attention degree to the first region is higher than an attention degree to the second region.

[00151] The addition module 13 is configured to add a first predicted label to the first region and add a second predicted label to the second region based on the respective attention degrees of the attention mechanism network to the first region and the second region, the first predicted label indicating that the first region belongs to a first predicted definition category, the second predicted label indicating that the second region belongs to a second predicted definition category, and a definition indicated by the first predicted definition category being higher than a definition indicated by the second predicted definition category.

[00152] The second obtaining module 14 is configured to obtain a first reference label of the first region and a second reference label of the second region, the first reference label indicating that the first region belongs to a first reference definition category, and the second reference label indicating that the second region belongs to a second reference definition category.

[00153] The adjustment module 15 is configured to adjust the attention parameter based on a difference between the first predicted label and the first reference label and a difference between the second predicted label and the second reference label, to obtain a trained attention mechanism network, the trained attention mechanism network being configured to, based on the adjusted attention parameter, extract an embedding feature of a palm print image for identifying an owner of a palm print in the palm print image.

[00154] In an embodiment, the manner in which the invoking module 12 invokes the attention mechanism network to perform attention degree recognition on the plurality of regions, to obtain the first region and the second region includes:

[00155] invoking the attention mechanism network to determine an attention weight of each region of the plurality of regions based on the attention parameter, an attention weight of any region being used for reflecting an attention degree of the attention mechanism network to the region; and

[00156] using a region of the plurality of regions whose attention weight is in a first weight range as the first region, and using a region of the plurality of regions whose attention weight is in a second weight range as the second region;

[00157] a weight value in the first weight range being greater than a weight value in the second weight range.

[00158] In an embodiment, the attention mechanism network is included in a palm print classification network, the adjusted attention parameter remain unchanged in the palm print classification network, the palm print classification network further includes a classification sub-network, and the apparatus 1 is further configured to:

[00159] obtain a second sample image, the second sample image being obtained by photographing a sample palm print, and having a palm print label indicating a real category of the sample palm print;

[00160] invoke the trained attention mechanism network to perform embedding on the second sample image based on the adjusted attention parameter, to generate a sample embedding feature of the second sample image;

[00161] invoke the classification sub-network to perform classification prediction on the sample palm print based on the sample embedding feature, to obtain a classification predicted result of the sample palm print; and

[00162] adjust, based on a difference between the real category and the classification predicted result, a network parameter in the palm print classification network other than the adjusted attention parameter, to obtain a trained palm print classification network.

[00163] In an embodiment, there are a plurality of second sample images, the plurality of second sample images are obtained by photographing the sample palm print from a plurality of photographing angles, and one photographing angle is used for photographing to obtain one second sample image of the sample palm print.

[00164] A manner in which the apparatus 1 invokes the trained attention mechanism network to perform embedding on the second sample image based on the adjusted attention parameter, to generate the sample embedding feature of the second sample image includes:

[00165] invoking the trained attention mechanism network to separately perform embedding on each second sample image based on the adjusted attention parameter, to generate an image embedding feature of each second sample image; and

[00166] fusing image embedding features of the plurality of second sample images, to generate the sample embedding feature.

[00167] In an embodiment, a manner in which the apparatus 1 fuses image embedding features of the plurality of second sample images, to generate the sample embedding feature includes:

[00168] stitching the image embedding features of the plurality of second sample images, to generate the sample embedding feature; or

[00169] summing the image embedding features of the plurality of second sample images, to generate the sample embedding feature.

[00170] In an embodiment, the trained palm print classification network includes the trained attention mechanism network and a trained classification sub-network, and the apparatus 1 is further configured to:

[00171] obtaining a to-be-classified (target) palm print image that is obtained by photographing a target palm print;

[00172] invoking the trained attention mechanism network, to perform embedding on the palm print image, to generate a target embedding feature of the palm print image; and

[00173] invoking the trained classification sub-network to perform classification prediction on the target palm print based on the target embedding feature, to obtain a target category of the target palm print.

[00174] In an embodiment, there are a plurality of palm print images obtained by photographing the target palm print from a plurality of photographing angles, and one photographing angle is used for photographing to obtain one palm print image of the target palm print.

[00175] A manner in which the apparatus 1 invokes the trained attention mechanism network, to perform embedding on the palm print image, to generate the target embedding feature of the palm print image includes:

[00176] invoking the trained attention mechanism network, to separately perform embedding on each palm print image, to generate an image embedding feature of each palm print image; and

[00177] fusing image embedding features of the plurality of palm print images, to generate the target embedding feature.

[00178] In an embodiment, the target category is used for indicating an owner of the target palm print, and the palm print image is obtained by photographing after a service initiator initiates an execution request for a target service.

[00179] The apparatus 1 is further configured to:

[00180] execute the target service on the service initiator if the target category indicates that the owner of the target palm print is the service initiator.

[00181] In an embodiment, a manner in which the second obtaining module 14 obtains the first reference label of the first region and the second reference label of the second region includes:

[00182] invoking a trained definition classification network to perform classification prediction on a definition of the first region, to obtain a first reference definition category of the first region;

[00183] invoking the trained definition classification network to perform classification prediction on a definition of the second region, to obtain a second reference definition category of the second region; and

[00184] adding the first reference label for the first region based on the first reference definition category, and adding the second reference label for the second region based on the second reference definition category.

[00185] In an embodiment, the apparatus 1 is further configured to:

[00186] obtain a third sample image and a definition classification network that is to be trained, the third sample image having a definition label indicating a real definition category of the third sample image;

[00187] invoke the definition classification network that is to be trained to perform classification prediction on a definition category of the third sample image, to obtain a predicted definition category of the third sample image; and

[00188] adjust, based on a difference between the real definition category and the predicted definition category, a network parameter of the definition classification network that is to be trained, to obtain the trained definition classification network.

[00189] In an embodiment, a manner in which the second obtaining module 14 obtains the first reference label of the first region and the second reference label of the second region includes:

[00190] invoking a trained image segmentation network to perform image segmentation on the first sample image, to obtain a plurality of segmented images of the first sample image, one segmented image corresponding to one definition category;

[00191] using a segmented image having a highest coincidence degree with the first region in the plurality of segmented images as a first definition matching image of the first region, and adding

the first reference label for the first region based on a definition category corresponding to the first definition matching image; and

[00192] using a segmented image having a highest coincidence degree with the second region in the plurality of segmented images as a second definition matching image of the second region, and adding the second reference label for the second region based on a definition category corresponding to the second definition matching image;

[00193] the first reference definition category being a definition category corresponding to the first definition matching image, and the second reference definition category being a definition category corresponding to the second definition matching image.

[00194] According to an embodiment of this application, operations involved in the image processing method shown in FIG. 3 may be executed by the modules in the image processing apparatus 1 shown in FIG. 10. For example, operation S101 shown in FIG. 3 may be performed by the first obtaining module 11 shown in FIG. 10; operation S102 shown in FIG. 3 may be performed by the invoking module 12 shown in FIG. 10; operation S103 shown in FIG. 3 may be performed by the addition module 13 shown in FIG. 10; operation S104 shown in FIG. 3 may be performed by the second obtaining module 14 shown in FIG. 10; and operation S105 shown in FIG. 3 may be performed by the adjustment module 15 shown in FIG. 10.

[00195] In this application, a first sample image can be obtained. The first sample image is divided into a plurality of regions, and the plurality of regions with different image definitions belong to different definition categories. In addition, an attention mechanism network can be invoked to perform attention degree recognition on the plurality of regions, to obtain a first region and a second region. The attention mechanism network includes an attention parameter, and determines, based on the attention parameter, that an attention degree to the first region is higher than an attention degree to the second region. Further, a first predicted label can be added to the first region and a second predicted label can be added to the second region based on the respective attention degrees to the first region and the second region. The first predicted label is used for indicating that the first region belongs to a first predicted definition category, the second predicted label is used for indicating that the second region belongs to a second predicted definition category, and a definition indicated by the first predicted definition category is higher than a definition indicated by the second predicted definition category. A first reference label of the first region and a second reference label of the second region can be obtained. The first reference label is used for indicating that the first region actually belongs to a first reference definition category, and the second reference label is used for indicating that the second region actually belongs to a second

reference definition category. Therefore, the attention parameter can be adjusted based on a difference between the first predicted label and the first reference label and a difference between the second predicted label and the second reference label, to obtain a trained attention mechanism network,. The trained attention mechanism network is configured to extract an embedding feature of a palm print image based on the adjusted attention parameter. The embedding feature is used for identifying an owner of a palm print in the palm print image. It can be learned that, the apparatus provided in this embodiment can add, by using an attention degree of the attention mechanism network to the region in the first sample image, the predicted label to the region of the first sample image, and a definition of a definition category indicated by a predicted label added for a region with a higher attention degree may be higher. Further, the attention parameter may be adjusted according to the difference between a real label (for example, the reference label) and the predicted label of the region. Subsequently, the attention mechanism network can adopt a higher attention degree for a region with a higher definition in the palm print image by using the adjusted attention parameter, and adopt a lower attention degree for a region with a lower definition in the palm print image, so that an image feature of the region with the higher definition in the palm print image can be extracted to a greater extent, and finally a more accurate embedding feature of the palm print image is extracted. More accurate classification on the palm print image (that is, more accurate identification on an owner of the palm print in the palm print image) can be implemented by using the more accurate embedding feature of the palm print image.

[00196] According to an embodiment of this application, the modules in the image processing apparatus 1 shown in FIG. 10 may be separately or wholly combined into one unit or several units, or one (or some) of the units herein may further be divided into a plurality of functionally smaller subunits, which can implement same operations without affecting implementation of the technical effects of this embodiment of this application. The foregoing modules are divided based on logical functions. In an actual application, a function of one module may also be implemented by a plurality of units, or functions of a plurality of modules are implemented by one unit. In other embodiments of this application, the image processing apparatus 1 may also include other units. In an actual application, the functions may alternatively be cooperatively implemented by other units and may be implemented with collaboration by a plurality of units.

[00197] According to an embodiment of this application, a computer program that can perform operations in the corresponding method shown in embodiments of this application may be run on a general computer device (the computer device may include processing elements and storage elements such as a central processing unit (CPU), a random access storage medium (RAM),

and a read-only storage medium (ROM)), to construct the image processing apparatus 1 shown in FIG. 10. The computer program may be recorded in, for example, a computer-readable recording medium, and may be loaded into the foregoing computer device by using the computer-readable recording medium, and run in the computer device.

[00198] Referring to FIG. 11, FIG. 11 is a schematic structural diagram of a computer device according to an embodiment of this application. As shown in FIG. 11, a computer device 1000 may include a processor 1001, a network interface 1004, and a memory 1005. In addition, in some embodiments, the computer device 1000 may further include: a user interface 1003 and at least one communication bus 1002. The communication bus 1002 is configured to implement connection and communication between these components. The user interface 1003 may include a display and a keyboard. In an embodiment, the user interface 1003 may further include a standard wired interface or wireless interface. In an embodiment, the network interface 1004 may include a standard wired interface or a standard wireless interface (such as a Wi-Fi interface). The memory 1005 may be a high-speed RAM memory, or may be a non-volatile memory, such as at least one magnetic disk memory. In an embodiment, the memory 1005 may alternatively be at least one storage apparatus that is located far away from the foregoing processor 1001. As shown in FIG. 11, the memory 1005, which is a type of computer storage medium, may include an operating system, a network communication module, a user interface module, and a device control application program.

[00199] In the computer device 1000 shown in FIG. 11, the network interface 1004 may provide a network communication function; the user interface 1003 is mainly an interface configured to provide input for a user; and the processor 1001 may be configured to invoke the device control application program stored in the memory 1005, to implement the following operations:

[00200] obtaining a first sample image, the first sample image being divided into a plurality of regions, and the plurality of regions with different image definitions belonging to different definition categories;

[00201] invoking an attention mechanism network to perform attention degree recognition on the plurality of regions, to obtain a first region and a second region, wherein the attention mechanism network includes an attention parameter, and determines, based on the attention parameter, that an attention degree to the first region is higher than an attention degree to the second region;

[00202] adding a first predicted label to the first region and adding a second predicted label to the second region based on the respective attention degrees to the first region and the second

region, the first predicted label indicating that the first region belongs to a first predicted definition category, the second predicted label indicating that the second region belongs to a second predicted definition category, and a definition indicated by the first predicted definition category being higher than a definition indicated by the second predicted definition category;

[00203] obtaining a first reference label of the first region and a second reference label of the second region, the first reference label indicating that the first region belongs to a first reference definition category, and the second reference label indicating that the second region belongs to a second reference definition category; and

[00204] adjusting the attention parameter based on a difference between the first predicted label and the first reference label and a difference between the second predicted label and the second reference label, to obtain a trained attention mechanism network, the trained attention mechanism network being configured to, based on the adjusted attention parameter, extract an embedding feature of a palm print image for identifying an owner of a palm print in the palm print image.

[00205] The computer device 1000 described in this embodiment of this application may perform the descriptions of the foregoing image processing method in the embodiments of this application, or may perform the descriptions of the foregoing image processing apparatus 1 in the foregoing embodiment corresponding to FIG. 10. Details are not described herein again. In addition, descriptions of beneficial effects of using the same method are not described herein again.

[00206] In addition, this application further provides a computer-readable storage medium, having a computer program stored therein. When executing the computer program, a processor can perform the descriptions of the image processing method in embodiments of this application. Therefore, details are not described herein again. In addition, descriptions of beneficial effects of using the same method are not described herein again. For technical details not disclosed in the embodiment of the computer storage medium in this application, refer to the descriptions of method embodiments of this application.

[00207] In an example, the foregoing computer program may be deployed to be executed on one computer device, on a plurality of computer devices located at one place, or on a plurality of computer devices distributed at a plurality of places and interconnected through a communication network. The plurality of computer devices distributed at the plurality of places and interconnected through the communication network may form a blockchain system.

[00208] The foregoing computer-readable storage medium may be an internal storage unit of the foregoing computer device, for example, a hard disk or memory of the computer device. The computer-readable storage medium may be an external storage device of the computer device, for example, a removable hard disk drive, a smart media card (SMC), a secure digital (SD) card, or a flash card equipped on the computer device. Further, the computer-readable storage medium may alternatively include both an internal storage unit and an external storage device of the computer device. The computer-readable storage medium is configured to store the computer program and other programs and data that are required by the computer device. The computer-readable storage medium may be further configured to temporarily store data that has been outputted or is to be outputted.

[00209] This application provides a computer program product, including a computer program, the computer program being stored in a computer-readable storage medium. A processor of a computer device reads the computer program from the computer-readable storage medium, and the processor executes the computer program, to cause the computer device to perform the descriptions of the foregoing image processing method in embodiments of this application. Therefore, details are not described herein again. In addition, descriptions of beneficial effects of using the same method are not described herein again. For technical details not disclosed in the embodiment of the computer-readable storage medium in this application, refer to the descriptions of the method embodiments of this application.

[00210] In the specification, claims, and accompanying drawings of embodiments of this application, the terms "first" and "second" are intended to distinguish between different objects but do not indicate a particular order. In addition, the term "include" or any variation thereof is intended to cover a non-exclusive inclusion. For example, a process, method, apparatus, product, or device that includes a series of steps or modules is not limited to the listed steps or units; and instead, further includes an operation or module that is not listed, or further includes another operation or unit that is intrinsic to the process, method, apparatus, product, or device.

[00211] A person of ordinary skill in the art may be aware that, units and algorithm operations of the examples described in the foregoing embodiments provided in this specification may be implemented by electronic hardware, computer software, or a combination thereof. To clearly describe the interchangeability between the hardware and the software, the foregoing has generally described compositions and operations of each example based on functions. Whether the functions are executed in a mode of hardware or software depends on particular applications and design constraint conditions of the technical solutions. A person skilled in the art may use different

methods to implement the described functions for each particular application, but it should not be considered that the implementation goes beyond the scope of the present application.

[00212] The foregoing disclosures are merely embodiments of this application, and it is clear that, the scope of the claims of this application is not limited thereto. Therefore, any equivalent modification made according to the claims of this application shall fall within the scope of this application.

CLAIMS

1. An image processing method, the method comprising:

obtaining a first sample image, the first sample image being divided into a plurality of regions, and the plurality of regions with different image definitions belonging to different definition categories;

invoking an attention mechanism network to perform attention degree recognition on the plurality of regions, to obtain a first region and a second region, wherein the attention mechanism network includes an attention parameter, and determines, based on the attention parameter, that an attention degree to the first region is higher than an attention degree to the second region;

adding a first predicted label to the first region and adding a second predicted label to the second region based on the respective attention degrees to the first region and the second region, the first predicted label indicating that the first region belongs to a first predicted definition category, the second predicted label indicating that the second region belongs to a second predicted definition category, and a definition indicated by the first predicted definition category being higher than a definition indicated by the second predicted definition category;

obtaining a first reference label of the first region and a second reference label of the second region, the first reference label indicating that the first region belongs to a first reference definition category, and the second reference label indicating that the second region belongs to a second reference definition category; and

adjusting the attention parameter based on a difference between the first predicted label and the first reference label and a difference between the second predicted label and the second reference label, to obtain a trained attention mechanism network, the trained attention mechanism network being configured to, based on the adjusted attention parameter, extract an embedding feature of a palm print image for identifying an owner of a palm print in the palm print image.

2. The method according to claim 1, wherein the invoking the attention mechanism network to perform the attention degree recognition on the plurality of regions, to obtain the first region and the second region comprises:

invoking the attention mechanism network to determine an attention weight of each region of the plurality of regions based on the attention parameter, an attention weight of any region being used for reflecting an attention degree of the attention mechanism network to the region; and

using a region of the plurality of regions whose attention weight is in a first weight range as the first region, and using a region of the plurality of regions whose attention weight is in a second

weight range as the second region;

a weight value in the first weight range being greater than a weight value in the second weight range.

3. The method according to claim 1 or 2, wherein the trained attention mechanism network is included in a palm print classification network, the adjusted attention parameter remain unchanged in the palm print classification network, the palm print classification network further includes a classification sub-network, and the method further comprises:

obtaining a second sample image, the second sample image being obtained by photographing a sample palm print, and having a palm print label indicating a real category of the sample palm print;

invoking the trained attention mechanism network to perform embedding on the second sample image based on the adjusted attention parameter, to generate a sample embedding feature of the second sample image;

invoking the classification sub-network to perform classification prediction on the sample palm print based on the sample embedding feature, to obtain a classification predicted result of the sample palm print; and

adjusting, based on a difference between the real category and the classification predicted result, a network parameter in the palm print classification network other than the adjusted attention parameter, to obtain a trained palm print classification network.

4. The method according to any one of claims 1 to 3, wherein there are a plurality of second sample images obtained by photographing the sample palm print from a plurality of photographing angles; and

the invoking the trained attention mechanism network to perform embedding on the second sample image based on the adjusted attention parameter, to generate the sample embedding feature of the second sample image comprises:

invoking the trained attention mechanism network to separately perform embedding on each second sample image based on the adjusted attention parameter, to generate an image embedding feature of each second sample image; and

fusing image embedding features of the plurality of second sample images, to generate the sample embedding feature.

5. The method according to any one of claims 1 to 4, wherein the fusing the image embedding features of the plurality of second sample images, to generate the sample embedding feature

comprises:

stitching the image embedding features of the plurality of second sample images, to generate the sample embedding feature; or

summing the image embedding features of the plurality of second sample images, to generate the sample embedding feature.

6. The method according to any one of claims 1 to 5, wherein the trained palm print classification network includes the trained attention mechanism network and a trained classification sub-network, and the method further comprises:

obtaining a target palm print image that is obtained by photographing a target palm print;

invoking the trained attention mechanism network, to perform embedding on the palm print image, to generate a target embedding feature of the palm print image; and

invoking the trained classification sub-network to perform classification prediction on the target palm print based on the target embedding feature, to obtain a target category of the target palm print.

7. The method according to any one of claims 1 to 6, wherein there are a plurality of palm print images obtained by photographing the target palm print from a plurality of photographing angles; and

the invoking the trained attention mechanism network, to perform embedding on the palm print image, to generate the target embedding feature of the palm print image comprises:

invoking the trained attention mechanism network, to separately perform embedding on each palm print image, to generate an image embedding feature of each palm print image; and

fusing image embedding features of the plurality of palm print images, to generate the target embedding feature.

8. The method according to any one of claims 1 to 7, wherein the target category is used for indicating an owner of the target palm print, and the palm print image is obtained by photographing after a service initiator initiates an execution request for a target service; and

the method further comprises:

executing the target service on the service initiator if the target category indicates that the owner of the target palm print is the service initiator.

9. The method according to any one of claims 1 to 8, wherein the obtaining the first reference

label of the first region and the second reference label of the second region comprises:

invoking a trained definition classification network to perform classification prediction on a definition of the first region, to obtain a first reference definition category of the first region;

invoking the trained definition classification network to perform classification prediction on a definition of the second region, to obtain a second reference definition category of the second region; and

adding the first reference label for the first region based on the first reference definition category, and adding the second reference label for the second region based on the second reference definition category.

10. The method according to any one of claims 1 to 9, further comprising:

obtaining a third sample image and a definition classification network that is to be trained, the third sample image having a definition label indicating a real definition category of the third sample image;

invoking the definition classification network that is to be trained to perform classification prediction on a definition category of the third sample image, to obtain a predicted definition category of the third sample image; and

adjusting, based on a difference between the real definition category and the predicted definition category, a network parameter of the definition classification network that is to be trained, to obtain the trained definition classification network.

11. The method according to any one of claims 1 to 10, wherein the obtaining the first reference label of the first region and the second reference label of the second region comprises:

invoking a trained image segmentation network to perform image segmentation on the first sample image, to obtain a plurality of segmented images of the first sample image, one segmented image corresponding to one definition category;

using a segmented image having a highest coincidence degree with the first region in the plurality of segmented images as a first definition matching image of the first region, and adding the first reference label for the first region based on a definition category corresponding to the first definition matching image; and

using a segmented image having a highest coincidence degree with the second region in the plurality of segmented images as a second definition matching image of the second region, and adding the second reference label for the second region based on a definition category

corresponding to the second definition matching image;

the first reference definition category being a definition category corresponding to the first definition matching image, and the second reference definition category being a definition category corresponding to the second definition matching image.

12. An image processing apparatus, the apparatus comprising:

a first obtaining module, configured to obtain a first sample image, the first sample image being divided into a plurality of regions, and the plurality of regions with different image definitions belonging to different definition categories;

an invoking module, configured to invoke an attention mechanism network to perform attention degree recognition on the plurality of regions, to obtain a first region and a second region, wherein the attention mechanism network includes an attention parameter, and determines, based on the attention parameter, that an attention degree to the first region is higher than an attention degree to the second region;

an addition module, configured to add a first predicted label to the first region and add a second predicted label to the second region based on the respective attention degrees to the first region and the second region, the first predicted label indicating that the first region belongs to a first predicted definition category, the second predicted label indicating that the second region belongs to a second predicted definition category, and a definition indicated by the first predicted definition category being higher than a definition indicated by the second predicted definition category;

a second obtaining module, configured to obtain a first reference label of the first region and a second reference label of the second region, the first reference label indicating that the first region belongs to a first reference definition category, and the second reference label indicating that the second region belongs to a second reference definition category; and

an adjustment module, configured to adjust the attention parameter based on a difference between the first predicted label and the first reference label and a difference between the second predicted label and the second reference label, to obtain a trained attention mechanism network, the trained attention mechanism network being configured to, based on the adjusted attention parameter, extract an embedding feature of a palm print image for identifying an owner of a palm print in the palm print image.

13. A computer program product, comprising a computer program, the computer program being executed by a processor to perform the method according to any one of claims 1 to 11.

14. A computer device, comprising a memory and a processor, the memory having a computer program stored therein, and the computer program, when executed by the processor, causing the processor to perform the method according to any one of claims 1 to 11.

15. A computer-readable storage medium, having a computer program stored therein, the computer program being adapted to be loaded by a processor to perform the method according to any one of claims 1 to 11.