

HAND IMAGE PROCESSING METHOD AND APPARATUS, DEVICE, STORAGE MEDIUM, AND PROGRAM PRODUCT

RELATED APPLICATION

[0001] This application claims priority to Chinese Patent Application No. 2022114733356, entitled "PALM IMAGE PROCESSING METHOD AND APPARATUS, DEVICE, STORAGE MEDIUM, AND PROGRAM PRODUCT" and filed on November 21, 2022, which is incorporated herein by reference in its entirety.

FIELD OF THE TECHNOLOGY

[0002] Embodiments of this application relate to the field of computer technologies, further to the field of computer vision technologies, and in particular, to a palm image processing method and apparatus, a device, a storage medium, and a program product.

BACKGROUND OF THE DISCLOSURE

[0003] With the development of computer technologies, a palm identification technology is increasingly widely applied, and may be applied to a plurality of scenarios. For example, in a payment scenario or a check in scenario at work, verification may be performed on a user identity through palm identification.

[0004] In the related art, a user provides a physical palm, and a palm identification device collects a palm image on the palm of the user, so that the palm identification device performs encryption on the palm image based on the palm image, and the encrypted palm image is configured for identity verification.

[0005] However, performing encryption and decryption on the palm image requires a large amount of time, which reduces efficiency of palm identification.

SUMMARY

[0006] According to various embodiments of this application, a palm image processing method and apparatus, a device, a storage medium, and a program product are provided.

[0007] According to an aspect, a palm image processing method is provided, performed by a palm identification device, including:

- [0008] obtaining a palm image, and extracting image features of a plurality of sizes from the palm image;
- [0009] fusing the image features of the plurality of sizes into a fused image feature;
- [0010] determining, based on the fused image feature, a palm area defining a palm in the palm image;
- [0011] extracting a palm part image from the palm image based on the palm area; and
- [0012] encrypting the palm part image into encrypted palm part image data, the encrypted palm part image data is for being transmitted to a server which performs identity identification using the encrypted palm part image data.
- [0013] According to another aspect, a method for training a palm part detection model is provided, performed by a computer device, the method including:
- [0014] obtaining a sample palm image and a sample palm part image;
- [0015] extracting sample image features of a plurality of sizes from the sample palm image;
- [0016] fusing the sample image features of the plurality of sizes into a sample fused image feature;
- [0017] determining, based on the sample fused image feature, a sample prediction palm area defining a palm in the sample palm image;
- [0018] extracting a sample prediction palm part image from the sample palm image based on the sample prediction palm area;
- [0019] calculating a loss function value according to the sample palm part image and the sample prediction palm part image; and
- [0020] updating a model parameter of the palm part detection model based on the loss function value.
- [0021] According to another aspect, this application provides a palm image processing apparatus, including:
- [0022] a feature extraction module, configured to obtain a palm image, and extract, to image features of a plurality of sizes from the palm image;

[0023] a feature fusion module, configured to fuse the image features of the plurality of sizes into a fused image feature;

[0024] a prediction module, configured to determine, based on the fused image feature, a palm area defining a palm in the palm image;

[0025] a cropping module, configured to extract a palm part image from the palm image based on the palm area; and

[0026] an encryption module, configured to encrypt the palm part image into encrypted palm part image data, the encrypted palm part image data is for being transmitted to a server which performs identity verification using the encrypted palm part image data.

[0027] According to another aspect, this application provides an apparatus for training a palm part detection model, including a palm part detection model, and further including an obtaining module, a cropping module, a calculation module, and an update module, the palm part detection model including a feature extraction module, a feature fusion module, and a prediction module;

[0028] the obtaining module being configured to obtain a sample palm image and a sample palm part image;

[0029] the feature extraction module being configured to extract sample image features of a plurality of sizes from the sample palm image;

[0030] the feature fusion module being configured to fuse the sample image features of the plurality of sizes into a sample fused image feature;

[0031] the prediction module being configured to determine, based on the sample fused image feature, a sample prediction palm area defining a palm in the sample palm image;

[0032] the cropping module being configured to extract a sample prediction palm part image from the sample palm image based on the sample prediction palm area;

[0033] the calculation module being configured to calculate a loss function value according to the sample palm part image and the sample prediction palm part image; and

[0034] the update module being configured to update a model parameter of the palm part detection model based on the loss function value.

[0035] According to another aspect, this application further provides a computer device, including a memory and a processor, the memory having computer-readable instructions stored therein, the computer-readable instructions, when executed by the processor, causing the processor to implement the operations of the method embodiments in this application.

[0036] According to another aspect, this application further provides a computer-readable storage medium, having computer-readable instructions stored therein, the computer-readable instructions, when executed by a processor, causing the processor to implement the operations of the method embodiments in this application.

[0037] According to another aspect, this application further provides a computer program product, including computer-readable instructions, the computer-readable instructions, when executed by a processor, causing the processor to implement the operations of the method embodiments in this application.

[0038] Details of one or more embodiments of this application are provided in the subsequent accompanying drawings and descriptions. Other features, objectives, and advantages of this application become apparent from the specification, the accompanying drawings, and the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0039] To describe the technical solutions in embodiments of this application more clearly, the following briefly introduces the accompanying drawings required for describing the embodiments. Apparently, the accompanying drawings in the following description show only some embodiments of this application, and a person of ordinary skill in the art may still derive other accompanying drawings from the accompanying drawings without creative efforts.

[0040] FIG. 1 is a schematic diagram of a palm image processing method according to an exemplary embodiment of this application.

[0041] FIG. 2 is a schematic diagram of an architecture of a computer system according to an exemplary embodiment of this application.

[0042] FIG. 3 is a flowchart of a palm image processing method according to an exemplary embodiment of this application.

[0043] FIG. 4 is a flowchart of a palm image processing method according to an exemplary embodiment of this application.

[0044] FIG. 5 is a schematic diagram of image feature fusion according to an exemplary embodiment of this application.

[0045] FIG. 6 is a schematic diagram of finger gap points in a palm according to an exemplary embodiment of this application.

[0046] FIG. 7 is a schematic diagram of a palm image processing method according to an exemplary embodiment of this application.

[0047] FIG. 8 is a schematic flowchart of a palm image processing method applied to a cross-device payment scenario according to an exemplary embodiment of this application.

[0048] FIG. 9 is a schematic flowchart of a palm image processing method applied to a cross-device identity verification scenario according to an exemplary embodiment of this application.

[0049] FIG. 10 is a schematic flowchart of a palm image processing method according to an exemplary embodiment of this application.

[0050] FIG. 11 is a framework diagram of a generation phase of a palm part detection model and a prediction phase of a palm area according to an exemplary embodiment of this application.

[0051] FIG. 12 is a flowchart of a method for training a palm part detection model according to an exemplary embodiment of this application.

[0052] FIG. 13 is a flowchart of a method for training a palm part detection model according to an exemplary embodiment of this application.

[0053] FIG. 14 is a block diagram of a palm image processing apparatus according to an exemplary embodiment of this application.

[0054] FIG. 15 is a block diagram of an apparatus for training a palm part detection model according to an exemplary embodiment of this application.

[0055] FIG. 16 is a schematic diagram of a structure of a computer device according to an exemplary embodiment of this application.

DESCRIPTION OF EMBODIMENTS

[0056] The technical solutions in embodiments of this application are clearly and completely described in the following with reference to the accompanying drawings in embodiments of this application. Apparently, the described embodiments are merely some rather than all of embodiments of this application. All other embodiments obtained by a person of ordinary skill in the art based on embodiments of this application without creative efforts shall fall within the protection scope of this application.

[0057] A computer vision (CV) technology is a science that studies how to use a machine to "see", and furthermore, is machine vision that a camera and a computer are used for replacing human eyes to perform identification, measurement, and the like on a target, and further perform graphic processing, so that the computer processes the target into an image more suitable for human eyes to observe, or an image transmitted to an instrument for detection. As a scientific discipline, the computer vision studies related theories and technologies and attempts to establish an artificial palm image identification system that can obtain information from images or multidimensional data. The computer vision technologies generally include technologies such as image processing, image identification, image semantic understanding, image retrieval, video processing, video semantic understanding, video content/behavior identification, three-dimensional object reconstruction, a 3D technology, virtual reality, augmented reality, synchronous positioning, and map construction, and further include common biometric feature identification technologies.

[0058] An embodiment of this application provides a schematic diagram of a palm image processing method. As shown in FIG. 1, the method may be performed by a palm identification device 10. The palm identification device 10 may have an image encryption function.

[0059] For example, the palm identification device 10 includes a camera, and the palm identification device 10 obtains a palm image 11 through a camera. The palm identification device 10 may also be connected to an image collector. The image collector includes a camera. The palm identification device 10 may obtain the palm image 11 collected by the image collector. The palm identification device 10 performs palm prediction on the palm image 11, to obtain a palm part image 12 corresponding to a palm in the palm image 11. The palm identification device 10 performs image encryption on the palm part image 12, to obtain encrypted palm part image data 13. The palm identification device 10 transmits the

encrypted palm part image data 13 to a palm identification server for decryption and palm comparison and identification, to obtain a user identifier 14 corresponding to the palm image 11.

[0060] For example, the palm identification device 10 obtains the palm image 11, and extracts image features of a plurality of sizes from the palm image 11. The palm identification device 10 fuses the image features of the plurality of sizes into a fused image feature. The palm identification device 10 determines a palm area in the palm image 11 based on the fused image feature. The palm identification device 10 extracts a palm part image 12 from the palm image 11 based on the palm area. The palm identification device 10 encrypts the palm part image 12 into encrypted palm part image data 13.

[0061] The encrypted palm part image data 13 is for being transmitted to the palm identification server for decryption and palm comparison and identification, to obtain the user identifier 14 corresponding to the palm image 11.

[0062] The palm part image 12 is an effective area for identification in the palm image 11, or the palm part image 12 is an area in the palm image 11 in which the palm is located, or the palm part image 12 is an area in the palm image 11 that can be configured for palm identification.

[0063] In some embodiments, manners in which the palm identification device 10 transmits the encrypted palm part image data 13 to the palm identification server include at least one of network transmission, data line transmission, and Bluetooth transmission, but is not limited thereto. This is not specifically limited in the embodiments of this application.

[0064] For example, the palm identification device 10 performs feature extraction on the palm image 11, to obtain the image features. The palm identification device 10 performs feature fusion on the image features, to obtain the fused image feature. The palm identification device 10 determines the palm area in the palm image, and extracts the palm part image 12 from the palm image based on the palm area.

[0065] For example, the palm area can be obtained by prediction by the palm part detection model, and the palm part detection model includes a backbone network, a neck network, and a prediction network.

[0066] The palm identification device 10 inputs the palm image 11 into the backbone network; the backbone network performs a slicing operation on the palm image, to obtain slice images of the plurality of sizes; and the palm identification device 10 performs feature

extraction on the slice images of the plurality of sizes, to obtain the image features of the plurality of sizes corresponding to the palm image 11. The size represents a size of a feature, and the plurality of sizes refer to a plurality of sizes.

[0067] For example, a size of the input palm image 11 is 640*640, and sizes of the output image features of the plurality of sizes are: 80*80, 40*40, and 20*20. A size or a size of an image or a feature may be a dimension of a matrix representing the image or the feature.

[0068] For example, the palm identification device 10 inputs the image features of the plurality of sizes into the neck network for feature fusion, to obtain the fused image feature. The palm identification device 10 inputs the fused image feature into the prediction network for prediction, to obtain the palm area in the palm image 11. The palm identification device 10 removes an image area outside the palm area in the palm image 11 and uses a remaining image as the palm part image 12.

[0069] In summary, according to the method provided in this embodiment, the image features of the plurality of sizes corresponding to the palm image are obtained by obtaining the palm image and performing feature extraction; feature fusion is performed on the image features of the plurality of sizes, to obtain the fused image feature; a palm area corresponding to a palm in the palm image is determined based on the fused image feature; the palm part image corresponding to the palm is obtained by cropping the palm image based on the palm area; and image encryption is performed on the palm part image, to obtain the encrypted palm part image data, and the encrypted palm part image data is transmitted to the palm identification server for decryption and palm comparison and identification, to perform identity identification. In this application, the palm part image corresponding to the palm is obtained by cropping the palm image, and image encryption, transmission, and decryption are performed only on the palm part image. This reduces time of identifying the palm image and improves identification efficiency of palm identification.

[0070] FIG. 2 is a schematic diagram of an architecture of a computer system according to an embodiment of this application. The computer system may include: a palm identification device 100 and a palm identification server 200.

[0071] The palm identification device 100 may be an electronic device such as a mobile phone, a tablet computer, a vehicle-mounted terminal (vehicle), a wearable device, a personal computer (PC), a voice interaction device with a palm image identification function,

an appliance with a palm image identification function, a vehicle-mounted terminal, an aircraft, or a vending terminal. A client running an application may be installed in the palm identification device 100. The application may be an application specifically performing palm image identification, or another application providing a palm image identification function. This is not limited in this application. In addition, a form of the application is not limited in this application, including but not limited to an application (App), a sub-application, a web page program, and the like installed in the palm identification device 100. The sub-application is a program running in a running environment provided by a parent application program. The parent application program is an independent native application. The child application runs in dependence on the parent application.

[0072] The palm identification server 200 may be an independent physical server, or may be a server cluster including a plurality of physical servers or a distributed system, or may be a cloud server providing basic cloud computing services, such as a cloud server providing cloud computing services, 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 palm image identification platform. The palm identification server 200 may be a backend server of the foregoing application, and is configured to provide a backend service for a client of the application.

[0073] The cloud technology is a hosting technology that unifies a series of resources such as hardware, software, and networks in a wide area network or a local area network to implement computing, storage, processing, and sharing of data. The cloud technology is a collective name of a network technology, an information technology, an integration technology, a management platform technology, an application technology, and the like based on an application of a cloud computing business mode, and may form a resource pool, which is used as required, and is flexible and convenient. The cloud computing technology becomes an important support. A backend service of a technical network system requires a large amount of computing and storage resources, such as video websites, image websites, and more portal websites. As the internet industry is highly developed and applied, each article may have its own identifier in the future and needs to be transmitted to a backend system for logical processing. Data at different levels is separately processed, and data in various industries requires strong system support, which can only be implemented through cloud computing.

[0074] In some embodiments, the server may be further implemented as a node in a blockchain system. A blockchain is a new application mode of computer technologies such as distributed data storage, point-to-point transmission, a consensus mechanism, and an encryption algorithm. The blockchain is essentially a decentralized database and is a string of data blocks generated through association by using a cryptographic method. Each data block includes information of a batch of network transactions, the information being used for verifying the validity of information of the data block (anti-counterfeiting) and generating a next block. The blockchain may include an underlying blockchain platform, a platform product service layer, and an application service layer.

[0075] The palm identification device 100 and the palm identification server 200 may communicate through a network, such as a wired network or a wireless network.

[0076] FIG. 3 is a flowchart of a palm image processing method according to an exemplary embodiment of this application. The method is performed by a palm identification device. The method may also be performed by a server. The method includes the following operations.

[0077] Operation 302: Obtain a palm image, and perform feature extraction on the palm image, to obtain image features of a plurality of sizes.

[0078] The palm image is an image obtained by capturing a physical palm. The physical palm includes a palm and fingers. The palm is an organism connecting a physical arm and physical fingers. The palm image includes an image of all or a part of the physical palm. For ease of description, an image of the physical palm in the palm image is referred to as a palm subsequently. The palm included in the palm image may be a physical palm of a user whose identity identification is to be performed. The palm in the palm image includes at least a palm, may further include the fingers or a part of the arm, and may further include an image of an environment in which the physical palm is located when the physical palm is captured.

[0079] A size is a size of an image feature, and may be represented by a matrix dimension number, such as 100*100, 80*80, 60*60, or 40*40. The plurality of sizes refer to two or more sizes, and each size is different. A maximum size in the plurality of sizes may be the same as a size of the palm image, and is generally smaller than the size of the palm image.

[0080] The palm image may be obtained by capturing, by the palm identification device, a physical palm of a user whose identity is to be verified, or may be transmitted by another device. The palm identification device may directly collect an image of the physical palm, to obtain the palm image. The palm identification device may also collect the image of the physical palm through an image collector connected to the palm image, to obtain the palm image. The palm identification device can perform feature extraction on the obtained palm image, to obtain image features of a plurality of sizes corresponding to the palm image.

[0081] For example, the palm identification device is a cash register device in a store, and the cash register device in the store captures a physical palm of a user through a camera, to obtain the palm image. Alternatively, the palm identification device is a palm image identification server, and after capturing a palm image of a user, a cash register device in a store transmits the palm image to the palm image identification server.

[0082] For example, the image features of the plurality of sizes may be respectively extracted from low-resolution images of the plurality of sizes by downsampling a high-resolution palm image into the low-resolution images of the plurality of sizes. The high resolution and the low resolution are relative concepts, and the high resolution is a resolution higher than the low resolution.

[0083] For example, a size of the input palm image is 640*640, and sizes of the output image features of the plurality of sizes are: 80*80, 40*40, and 20*20.

[0084] Operation 304: Fuse the image features of the plurality of sizes, to obtain a fused image feature.

[0085] The palm identification device may perform feature fusion on the image features of the plurality of sizes, to obtain the fused image feature.

[0086] The image features of the plurality of sizes refer to features obtained by performing feature extraction on images of the plurality of sizes. Image features at different sizes may express different information. In the image features of the plurality of sizes, image features of larger sizes express more position information, and image features of smaller sizes express more semantic information.

[0087] For example, the palm identification device extracts features of the plurality of sizes, under different receptive fields, and in different categories, and fuses the image features of the plurality of sizes, to obtain the fused image feature corresponding to the palm image.

[0088] Operation 306: Determine, based on the fused image feature, a palm area identifying a palm in the palm image.

[0089] The palm area is an identifier configured for representing a position of the palm in the palm image. The palm area includes the palm in the palm image. A boundary of the palm area encloses the palm in the palm image.

[0090] In some embodiments, a shape of the palm area may be a regular shape, for example, may be at least one of a rectangle, a square, a rhombus, a hexagon, a circle, or a triangle. The shape of the palm area may alternatively be an irregular shape. The regular shape is a shape that can be drawn by using a fixed parameter and based on a fixed rule. The irregular shape cannot be drawn by using the fixed parameter and based on the fixed rule, and can generally only be represented by a set of points that constitute the irregular shape.

[0091] The palm identification device can perform prediction based on the fused image feature by using artificial intelligence, to determine the palm area corresponding to the palm in the palm image.

[0092] In some embodiments, a computer device can determine, based on the fused image feature, the position of the palm area in the palm image that identifies the palm in the palm image. When the palm area is of the regular shape, the position may be represented by a value of the fixed parameter required for drawing the regular shape, for example, may be represented by a position of a fixed point of the palm area and a size parameter of the palm area. The fixed point is a point whose position is fixed relative to the palm area. The size parameter is a parameter representing a size of the palm area. For example, when the palm area is of the rectangle shape, the fixed point may be any corner point or a center point of the rectangle, and the size parameter is a side length of an adjacent side of the palm area.

[0093] Operation 308: Extract a palm part image from the palm image based on the palm area.

[0094] The palm part image is an effective area for identification in the palm image, or the palm part image is an area in the palm image in which the palm is located, or the palm part image is an area in the palm image that can be configured for palm identification.

[0095] For example, the palm identification device removes, based on the position of the palm area, an image area located outside the palm area in the palm image, and uses a remaining image as the palm part image.

[0096] Operation 310: Perform image encryption on the palm part image, to obtain encrypted palm part image data, where the encrypted palm part image data is configured for being transmitted to a server, and the server performing identity identification based on the encrypted palm part image data.

[0097] The encrypted palm part image data is data obtained by encrypting the palm part image. The palm part image can be obtained by decrypting the encrypted palm part image data. The encrypted palm part image data may be a bit stream or a word string, or may be an image with content that is difficult to identify by naked eyes. The palm part image is configured for being transmitted to a server, and the server performs identity identification based on the encrypted palm part image data.

[0098] The palm identification device may access a network and transmit the encrypted palm part image data to the server through the network. The palm identification device may also transmit the encrypted palm part image data to a terminal through a point-to-point connection between the palm identification device and the terminal, and then the terminal accessing the network transmits the encrypted palm part image data to the server. When the palm identification device is the server, the encrypted palm part image data may be transmitted to another server through a direct cable or the network.

[0099] In some embodiments, the server may decrypt the encrypted palm part image data, to obtain the palm part image, and then perform identity identification based on the palm part image obtained by decryption.

[00100] In some embodiments, the server may decrypt the encrypted palm part image data, to obtain the palm part image, and search a database for a palm part image matching the palm part image obtained by decryption. When the matching palm part image is found, a user identifier associated with the matching palm part image is determined, to perform identity identification. When no matching palm part image is found, it may be determined that identity identification fails.

[00101] In some embodiments, the server may decrypt the encrypted palm part image data, to obtain the palm part image, extract a palm feature from the palm part image, and then search the database for a palm feature matching the extracted palm feature. When the matching palm feature is found, the user identifier associated with the matching palm feature is determined, to perform identity identification.

[00102] In some embodiments, after determining the user identifier by performing identity identification based on the encrypted palm part image data, the server can further verify a permission corresponding to the user identifier. When verifying that the user identifier has a permission for a preset function, the server may further execute or trigger the preset function. The preset function is, for example, access control or another function that needs to verify a user permission.

[00103] In some embodiments, manners in which the palm identification device transmits the encrypted palm part image data to the palm identification server include at least one of network transmission, data line transmission, and Bluetooth transmission, but is not limited thereto. This is not specifically limited in embodiments of this application.

[00104] The palm identification device may perform decryption and palm comparison and identification based on the encrypted palm part image data, to perform identity identification. The palm comparison and identification may be specifically comparing and identifying the palm feature in the palm part image with a preset palm feature in the database.

[00105] The preset palm feature is a palm feature of a palm of a stored user identifier, and each preset palm feature has a corresponding user identifier, representing that the preset palm feature belongs to the user identifier and is the palm feature of the palm of the user. The user identifier may be any user identifier. For example, the user identifier is a user identifier registered in a payment application, or the user identifier is a user identifier registered in an enterprise.

[00106] In this embodiment of this application, the palm identification server includes the database. The database includes a plurality of preset palm features and the user identifier corresponding to each preset palm feature. In the database, the preset palm features and the user identifiers may be in a one-to-one correspondence, or one user identifier may correspond to at least two preset palm features.

[00107] For example, a plurality of users register in the payment application, a preset palm feature of each user is bound with a corresponding user identifier, and the palm features of the plurality of users and the corresponding user identifiers are correspondingly stored in the database. Subsequently, when the user uses the payment application, a target user identifier is determined by performing palm comparison and identification on a first palm area of an infrared image and a second palm area of a color image and the preset palm features in the database, to implement identity verification on the user.

[00108] For example, the palm identification device transmits the encrypted palm part image data to a palm identification server for decryption and palm comparison and identification, to obtain a user identifier corresponding to the palm image.

[00109] In summary, according to the method provided in this embodiment, the image features of the plurality of sizes corresponding to the palm image are obtained by obtaining the palm image and performing feature extraction; feature fusion is performed on the image features of the plurality of sizes, to obtain the fused image feature; prediction is performed based on the fused image feature, to obtain the palm area corresponding to the palm in the palm image; the palm part image corresponding to the palm is obtained by cropping the palm image based on the palm area; and image encryption is performed on the palm part image, to obtain the encrypted palm part image data, and the encrypted palm part image data is transmitted to the palm identification server for decryption and palm comparison and identification, to obtain a user identifier corresponding to the palm image. In this application, the palm part image corresponding to the palm is obtained by cropping the palm image, and image encryption, transmission, and decryption are performed only on the palm part image. This reduces time of identifying the palm image and improves identification efficiency of palm identification.

[00110] FIG. 4 is a flowchart of a palm image processing method according to an exemplary embodiment of this application. The method is performed by a palm identification device. The method includes the following operations.

[00111] Operation 402: Obtain a palm image, and perform feature extraction on the palm image, through a backbone network in a palm part detection model, to obtain image features of a plurality of sizes.

[00112] The palm image is a palm image of a to-be-determined user identifier.

[00113] For example, the palm identification device captures a physical palm of the user, to obtain the palm image. The palm image includes the palm. The palm may be a left palm of the user or a right palm of the user.

[00114] For example, the palm identification device is an internet of things device. The internet of things device captures the left palm of the user through a camera, to obtain the palm image. The internet of things device may be a cash register device in a store. For another example, when the user shops for a transaction in the store, the user extends a palm

toward a camera of a payment terminal in the store, and the payment terminal in the store captures the physical palm of the user through the camera, to obtain the palm image.

[00115] In some embodiments, the palm identification device establishes a communication connection with another device, and receives, through the communication connection, a palm image transmitted by the another device. For example, the palm identification device is a payment application server, and the another device may be a payment terminal. After the payment terminal captures the physical palm of the user, to obtain the palm image, the palm image is transmitted to the payment application server through the communication connection between the payment terminal and the payment application server, so that the payment application server can determine a user identifier of the palm image.

[00116] In some embodiments, the palm area is obtained by prediction by a palm part detection model, and the palm part detection model includes a backbone network; and the performing feature extraction on the palm image, to obtain image features of a plurality of sizes includes: performing a slicing operation on the palm image through the backbone network, to obtain slice images of the plurality of sizes; and respectively performing feature extraction on the slice images of the plurality of sizes through the backbone network, to obtain the image features of the plurality of sizes.

[00117] The palm part detection model is an artificial intelligence model, may be a deep learning model, or may be a neural network model. The palm part detection model has at least a function of determining, based on a fused image feature, a palm area identifying a palm in the palm image. The backbone network is a structure configured for extracting image features of a plurality of sizes from the palm image in the palm part detection model. The slicing operation is an operation of segmenting the palm image to generate images of smaller sizes. The smaller size means that a size of the generated image is smaller than the size of the palm image. Pixels in the slice image are derived from the palm image.

[00118] The palm identification device may perform a slicing operation through the backbone network, to obtain the slice images of the plurality of sizes. The palm identification device may perform feature extraction on the slice images of the plurality of sizes, to obtain the image features of the plurality of sizes corresponding to the palm image.

[00119] In some embodiments, the performing a slicing operation on the palm image through the backbone network, to obtain slice images of the plurality of sizes includes:

determining, through the backbone network and based on the palm image, a slice image of a maximum size in the plurality of sizes; and using, through the backbone network, the slice image of the maximum size as a first layer, and starting downsampling layer by layer, to obtain the slice images of the plurality of sizes including the slice image at the first layer.

[00120] A computer device may first determine, based on the palm image, the slice image of the maximum size in the plurality of sizes. The size of the image feature of the maximum size may be smaller than or equal to the size of the palm image. The computer device may downsample the palm image once, to generate the slice image of the maximum size in the plurality of sizes.

[00121] Downsampling the palm image is a process of sampling at least one pixel in the palm image. Pixels obtained through downsampling are spliced in sequence in the palm image, so that a slice image at a smaller size can be obtained.

[00122] If a quantity of the plurality of sizes is N , the plurality of sizes are N sizes. The N sizes are all different sizes, and therefore, there is a maximum size in the N sizes. The slice image of the maximum size is used as a first layer, downsampling layer by layer is started from the first layer, an n^{th} layer is downsampled to obtain a slice image at an $(n+1)^{\text{th}}$ layer, and an $(N-1)^{\text{th}}$ layer is downsampled to obtain a slice image at the N^{th} layer. n is a positive integer from 1 to $N-1$. Finally, a total of N slice images at the first layer to the N^{th} layer may be obtained.

[00123] In this embodiment, the slice images of the plurality of sizes are obtained by downsampling layer by layer, which provides a new way to obtain the slice images of the plurality of sizes by segmenting the palm image.

[00124] In some embodiments, the performing a slicing operation on the palm image through the backbone network, to obtain slice images of the plurality of sizes includes: downsampling and splicing, through the backbone network, pixels in the palm image of the plurality of sizes, to obtain the slice images of the plurality of sizes. There are two adjacent pixels in slice images of different sizes, and a quantity of pixels between the two pixels at which sampling occurs in the palm image is different.

[00125] In this embodiment, the palm identification device may sample the pixels from the palm image of each size by using a sampling interval of the size, and splice the sampled pixels according to a relative position relationship in the palm image, to obtain a slice image of each size. The sampling interval is a quantity of pixels that are spaced, including intervals

of pixels in two directions: a row and a column. The size is negatively correlated with the sampling interval. A larger size indicates a smaller sampling interval. A smaller size indicates a larger sampling interval. Pixels that are not sampled in the palm image may be concentrated to a channel of an image feature.

[00126] For example, of a maximum size in the plurality of sizes, in the palm image, sampling is performed on every other column. That is, a pixel is selected every other pixel in the row direction, and the selected pixels are spliced, to obtain the slice image. Information in the spliced slice image is not lost, but size information in the palm image is concentrated on the channel, and a convolution operation is performed on an obtained new image, to obtain a downsampling feature map without information loss.

[00127] For another example, for a second largest size in the plurality of sizes, in the palm image, sampling is performed in every two columns. That is, pixels are selected every two pixels, and the selected pixels are spliced, to obtain the slice image. Information in the spliced slice image is not lost, but size information in the palm image is concentrated on the channel, and a convolution operation is performed on an obtained new image, to obtain a downsampling feature map without information loss.

[00128] For example, a size of an original image is: $640 \times 640 \times 3$. Through the slicing operation, a feature map of $320 \times 320 \times 12$ is obtained. 640×640 in the original image is configured for representing width*height, and 3 in the original image is configured for representing a length (also referred to as a quantity of channels) of a feature vector corresponding to each pixel.

[00129] Operation 404: Perform, through a neck network of the palm part detection model, feature fusion on the image features of the plurality of sizes, to obtain the fused image feature.

[00130] The palm area is obtained by prediction by the palm part detection model, and the palm part detection model includes a neck network; and the fusing the image features of the plurality of sizes, to obtain a fused image feature includes: performing, through the neck network, feature fusion on the image features of the plurality of sizes, to obtain the fused image feature.

[00131] The neck network belongs to the palm part detection model and is a structure configured for performing feature fusion on the image features of the plurality of sizes to obtain the fused image feature. The neck network is connected to the backbone network, and

is configured for receiving the image features of the plurality of sizes output by the backbone network.

[00132] The image features of the plurality of sizes refer to features obtained by performing feature extraction on images of the plurality of sizes. Image features at different sizes may express different information. In the image features of the plurality of sizes, image features of larger sizes express more position information, and image features of smaller sizes express more semantic information.

[00133] For example, the palm identification device extracts features of the plurality of sizes, under different receptive fields, and in different categories, and fuses the image features of the plurality of sizes, to obtain the fused image feature corresponding to the palm image.

[00134] In some embodiments, the performing, through the neck network, feature fusion on the image features of the plurality of sizes, to obtain the fused image feature includes: performing, through the neck network, fusion for the same size on the image features of the plurality of sizes, to obtain the fused image feature. The fusion of the image features of the same size may be summation or averaging.

[00135] The palm identification device fuses the image features of the plurality of sizes, which increases detection accuracy of a small target, so that a palm area can be detected more accurately. In this way, the palm part image can be extracted more accurately. This further reduces a data amount of the encrypted palm part image data, improves identity identification efficiency, and ensures security of identity identification.

[00136] In some embodiments, the performing, through the neck network, feature fusion on the image features of the plurality of sizes, to obtain the fused image feature includes: performing, through the neck network, feature fusion based on the image features of the plurality of sizes, to obtain first intermediate features of the plurality of sizes; performing, through the neck network, feature fusion of the plurality of sizes based on the first intermediate features of the plurality of sizes, to obtain second intermediate features of the plurality of sizes; and performing, through the neck network, feature fusion on the second intermediate features of the plurality of sizes, to obtain the fused image feature.

[00137] The feature fusion of the plurality of sizes refers to performing feature fusion for different sizes respectively. The first intermediate features of the plurality of sizes may be in a one-to-one correspondence with the image features of the plurality of sizes. The first intermediate features of the plurality of sizes may be in a one-to-one correspondence with the

second intermediate features of the plurality of sizes. The image features of the plurality of sizes may be consistent with the first intermediate features of the plurality of sizes. The first intermediate features of the plurality of sizes may be consistent with the second intermediate features of the plurality of sizes. The performing feature fusion on the second intermediate features of the plurality of sizes may be fusing, for the same size, the second intermediate features of the plurality of sizes. The fusion may be summation or averaging.

[00138] In this embodiment, features of different sizes are fused and are finally fused into the fused image feature, which can fully express detailed features in the palm image, so that a palm area is detected more accurately. In this way, the palm part image can be extracted more accurately. This further reduces a data amount of the encrypted palm part image data, improves identity identification efficiency, and ensures security of identity identification.

[00139] In some embodiments, the plurality of sizes are N sizes, N is a positive integer greater than 1, and the extracting, based on the image features of the plurality of sizes, first intermediate features of the plurality of sizes includes: determining, based on an image feature ranking first in a positive order of size, a first intermediate feature ranking first in the positive order of size; upsampling layer by layer starting from the first intermediate feature ranking first in the positive order of size, and fusing a result of upsampling a first intermediate feature ranking n^{th} in the positive order of size with an image feature ranking $(n+1)^{\text{th}}$ in the positive order of size, to obtain a first intermediate feature ranking $(n+1)^{\text{th}}$ in the positive order of size; and determining, when obtaining a first intermediate feature at an N^{th} layer, the first intermediate features of the plurality of sizes from a first layer to the N^{th} layer. n is a positive integer from 1 to $N-1$; and the positive order of size is in ascending order of size.

[00140] The ascending order is referred to as the positive order, and the positive order of size is in ascending order of size. Upsampling is a process of increasing the size. The upsampling may use a size at a next layer as a target, and adjacent feature elements in a feature at each layer are interpolated, to obtain a feature at the next layer.

[00141] In some embodiments, the determining, based on an image feature ranking first in a positive order of size, a first intermediate feature ranking first in the positive order of size includes: performing at least one time of convolution on the image feature ranking first in the positive order of size, to obtain the first intermediate feature ranking first in the

positive order of size. In addition to performing convolution, channel adjustment may be further performed. A quantity of channels of the first intermediate features that are obtained after the channel adjustment is performed may be consistent.

[00142] In some embodiments, a size of a result of upsampling the first intermediate feature ranking n^{th} in the positive order of size may be the same as a size of an image feature ranking $(n+1)^{\text{th}}$ in the positive order of size. The result of upsampling the first intermediate feature ranking n^{th} in the positive order of size and the image feature ranking $(n+1)^{\text{th}}$ in the positive order of size are fused, and the two may be spliced, or summation or averaging may be performed on the two at the same size.

[00143] For example, FIG. 5 is a schematic diagram of image feature fusion. A palm identification device performs feature extraction on a palm image 501, to obtain image features 502 at a plurality of sizes corresponding to the palm image 501. The image features 502 include a feature layer T1, a feature layer T2, a feature layer T3, and a feature layer T4. After the image feature 502 is obtained, one time of convolution is performed on the feature layer T4 and channel adjustment is performed to obtain a feature layer P4. The feature layer P4 is upsampled and is combined with the feature layer T3, and then convolution is performed and channel adjustment is performed to obtain a feature layer P3. The feature layer P3 is upsampled and is combined with the feature layer T2, and then convolution is performed and channel adjustment is performed to obtain a feature layer P2. The feature layer P2 is upsampled and is combined with the feature layer T1, and then $1*1$ convolution is performed and channel adjustment is performed to obtain a feature layer P1.

[00144] In some embodiments, the plurality of sizes are N sizes, N is a positive integer greater than 1, and the performing feature fusion of the plurality of sizes based on the first intermediate features of the plurality of sizes, to obtain second intermediate features of the plurality of sizes includes: determining, based on a first intermediate feature ranking first in a reverse order of size, a second intermediate feature ranking first in the reverse order of size; downsampling layer by layer starting from the second intermediate feature ranking first in the reverse order of size, and fusing a result of downsampling a second intermediate feature ranking m^{th} in the reverse order of size with a first intermediate feature ranking $(m+1)^{\text{th}}$ in the reverse order of size, to obtain a second intermediate feature ranking $(n+1)^{\text{th}}$ in the reverse order of size, where m is an integer from 1 to N-1; and the reverse order of size is in descending order of size; and determining, when obtaining the first intermediate feature at the

N^{th} layer, the first intermediate features of the plurality of sizes from the first layer to the N^{th} layer.

[00145] The descending order is referred to as the reverse order, and the reverse order of size is in descending order of size. Downsampling is a process of reducing a size. The downsampling may use a size at a next layer as a target, and feature elements in a feature at each layer are used at intervals, to obtain a feature at the next layer. The first intermediate feature may be referred to as a semantic image feature. The second intermediate feature may be referred to as a position image feature. The semantic image features represent more semantic information, and the semantic image features represent more position information.

[00146] In some embodiments, the determining, based on a first intermediate feature ranking first in a reverse order of size, a second intermediate feature ranking first in the reverse order of size includes: performing at least one time of convolution on the first intermediate feature ranking first in the reverse order of size, to obtain the second intermediate feature ranking first in the reverse order of size. In addition to performing convolution, channel adjustment may be further performed. A quantity of channels of the first intermediate features after the channel adjustment is performed may be consistent.

[00147] A size of a result of downsampling the second intermediate feature ranking m^{th} in the reverse order of size may be consistent with a size of a first intermediate feature ranking $(m+1)^{\text{th}}$ in the reverse order of size. The result of downsampling the second intermediate feature ranking m^{th} in the reverse order of size and the first intermediate feature ranking $(m+1)^{\text{th}}$ in the reverse order of size are fused, and the two may be spliced, or summation or averaging may be performed on the two at the same size.

[00148] For example, the palm identification device may perform one time of convolution on the feature layer P1 and perform channel adjustment to obtain a feature layer N1. The feature layer N1 is downsampled and is combined with the feature layer P2, and then convolution is performed and channel adjustment is performed to obtain a feature layer N2. The feature layer N2 is downsampled and is combined with the feature layer P3, and then convolution is performed and channel adjustment is performed to obtain a feature layer N3. The feature layer N3 is downsampled and is combined with the feature layer P4, and then convolution is performed and channel adjustment is performed to obtain a feature layer N4. The palm identification device fuses features at a feature layer N1–4, to obtain a fused image feature 503.

[00149] In some embodiments, an activation function in the convolution is a logistic activation function.

[00150] In some embodiments, the palm identification device upsamples the image features of the plurality of sizes in ascending order of size, to obtain the first intermediate feature; the palm identification device downsamples the image features of the plurality of sizes in descending order of size, to obtain the second intermediate feature; and the palm identification device performs feature fusion on the first intermediate feature and the second intermediate feature, to obtain the fused image feature.

[00151] Operation 406: Determine, through a prediction network of the palm part detection model based on the fused image feature, a palm area identifying a palm in the palm image.

[00152] The palm area is an identifier configured for representing a position of the palm in the palm image.

[00153] In some embodiments, the palm area is determined by the palm part detection model, and the palm part detection model includes a prediction network; and the determining, based on the fused image feature, a palm area identifying a palm in the palm image includes: performing grid division on the fused image feature, to obtain a plurality of grid features of the fused image feature; and inputting the plurality of grid features into the prediction network, and performing prediction on the plurality of grid features, to obtain the palm area.

[00154] The grid feature is a gridded fused image feature obtained by performing grid division on the fused image feature. For example, the palm identification device performs grid division on the fused image feature, to obtain the plurality of grid features. The palm identification device separately performs prediction based on each grid feature through the prediction network, to obtain a prediction result corresponding to each grid feature. Further, the palm identification device may determine the position of the palm area in the palm image based on the prediction results respectively corresponding to the plurality of grid features.

[00155] In some embodiments, the palm identification device is equipped with an infrared camera, and the palm image processing method further includes: obtaining an infrared image collected of a palm the same as the palm image, where the infrared image is an image obtained by the infrared camera imaging, based on infrared light, the palm; identifying a palm area from the infrared image; and the extracting a palm part image from the palm image based on the palm area includes: determining an overlapping area between

the palm area in the palm image and the palm area in the infrared image; and extracting the palm part image from the palm image based on the overlapping area.

[00156] In this embodiment, the palm identification device is further equipped with the infrared camera. The palm identification device obtains an infrared image corresponding to the same physical palm. The palm identification device performs area identification on the infrared image, to determine the palm area in the infrared image. The infrared image is an image obtained by the infrared camera imaging, based on infrared light, the palm.

[00157] In some embodiments, the identifying a palm area from the infrared image includes: detecting a finger gap point in the infrared image; and determining the palm area in the infrared image based on the finger gap point.

[00158] For example, the palm identification device performs detection on the finger gap point in the infrared image, and determines the palm area in the infrared image based on the finger gap point.

[00159] FIG. 6 is a schematic diagram of a finger gap point in a palm. The finger gap point is a first finger gap point 601 between an index finger and a middle finger, or the finger gap point is a second finger gap point 602 between the middle finger and a ring finger, or the finger gap point is a third finger gap point 603 between the ring finger and a little finger.

[00160] Because a palm area in a palm image may exist in any area in the palm image, to determine a position of the palm area in the palm image, finger gap point detection is performed on the palm image, to obtain at least one finger gap point in the palm image, so that the palm area can be determined subsequently according to the at least one finger gap point.

[00161] Operation 408: Extract a palm part image from the palm image based on the palm area.

[00162] The palm part image is an effective area for identification in the palm image, or the palm part image is an area in the palm image in which the palm is located, or the palm part image is an area in the palm image that can be configured for palm identification.

[00163] In some embodiments, the palm identification device obtains at least three finger gap points in the infrared image; the finger gap points are sequentially connected, to obtain a finger gap point connecting line; and the palm identification device extracts the infrared image based on the finger gap point connecting line to obtain the palm area.

[00164] In some embodiments, the determining an overlapping area between the palm area in the palm image and the palm area in the infrared image includes: obtaining a coordinate parameter of the palm area in the palm image, and obtaining a coordinate parameter of the palm area in the infrared image; and determining the overlapping area between the palm area and the palm area based on the coordinate parameter of the palm area and the coordinate parameter of the palm area.

[00165] The palm identification device obtains the coordinate parameter of the palm area in the palm image, and obtains the coordinate parameter of the palm area in the infrared image; and the palm identification device obtains the palm part image corresponding to the palm by cropping the palm image based on the overlapping area between the coordinate parameter of the palm area and the coordinate parameter of the palm area.

[00166] Operation 410: Perform image encryption on the palm part image, to obtain encrypted palm part image data, where the encrypted palm part image data is configured for being transmitted to a server, and the server performs identity identification based on the encrypted palm part image data.

[00167] The encrypted palm part image data is configured for being transmitted to the palm identification server for decryption and palm comparison and identification, to obtain the user identifier corresponding to the palm image. The palm comparison and identification refers to comparing and identifying the palm feature in the palm part image with a preset palm feature in the database.

[00168] As one of biological features, the palm has biological uniqueness and distinction. Compared with facial identification, which is currently widely applied to fields such as identity verification, payment, access control, and bus riding, the palm is not affected by makeup, a mask, ink glasses, and the like, which can improve accuracy of user identity verification. In some scenarios, such as a pandemic prevention and control scenario, a mask needs to be worn to cover the mouth and the nose. In this case, performing identity verification by using a palm image may be a better option.

[00169] Cross-device registration identification is a capability that is very important for user experience. For two types of associated devices, a user may register in a device of one type, a user identifier of the user is bound with a palm feature of the user, and then the user may perform identity verification on a device of the other type. Because a mobile phone and an internet of things device have a large difference in image style and image quality, the

user can directly perform, through cross-device registration and identification, identity verification on the internet of things device end after performing registration on the mobile phone end, and the user does not need to perform registration on the two types of devices. For example, after perform registration on the mobile phone end, the user can directly perform identity verification on a device in a store, and does not need to perform registration on the device in the store, to avoid information leaking of the user.

[00170] In summary, according to the method provided in this embodiment, the image features of the plurality of sizes corresponding to the palm image are obtained by obtaining the palm image and performing feature extraction; feature fusion is performed on the image features of the plurality of sizes, to obtain the fused image feature; prediction is performed based on the fused image feature, to obtain the palm area corresponding to the palm in the palm image; the palm part image corresponding to the palm is obtained by cropping the palm image based on the palm area; and image encryption is performed on the palm part image, to obtain the encrypted palm part image data, and the encrypted palm part image data is transmitted to the palm identification server for decryption and palm comparison and identification, to obtain a user identifier corresponding to the palm image. In this application, the palm part image corresponding to the palm is obtained by cropping the palm image, and image encryption, transmission, and decryption are performed only on the palm part image. This reduces time of identifying the palm image and improves identification efficiency of palm identification.

[00171] According to the method provided in this embodiment, the slicing operation is performed on the palm image through the backbone network and feature extraction is performed, to obtain the image features of the plurality of sizes. Based on the foregoing method, features of the plurality of sizes, under different receptive fields, and in different categories are obtained. This improves accuracy of prediction of the palm area.

[00172] According to the method provided in this embodiment, the image features of the plurality of sizes are upsampled through the neck network in ascending order of size, to obtain the first intermediate feature; the image features of the plurality of sizes are downsampled in descending order of size, to obtain the second intermediate feature; and feature fusion is performed on the first intermediate feature and the second intermediate feature, to obtain the fused image feature. Based on the foregoing method, the image features of the plurality of sizes can be fused. This increases accuracy of detecting a small target in the palm image.

[00173] According to the method provided in this embodiment, a gridded fused image feature is obtained by performing grid division on the fused image feature. Through the prediction network, prediction is performed based on the gridded fused image feature, to obtain the palm area. Based on the foregoing method, accuracy of prediction of the palm area is improved, and security of identity identification is further ensured based on the palm.

[00174] According to the method provided in this embodiment, an infrared image corresponding to a same palm is obtained, area identification is performed on the infrared image, to determine a palm area in the infrared image, and based on the overlapping area between the palm area and the palm area in the infrared image, a palm part image corresponding to the palm is obtained by cropping the palm image. Based on the foregoing method, through the palm area in the infrared image and the palm area, a palm part image corresponding to the palm is jointly obtained, so that accuracy of extracting the palm part image is improved, and the obtained palm part image is more accurate.

[00175] FIG. 7 is a schematic diagram of a palm image processing method according to an exemplary embodiment of this application. The method is performed by a palm identification device. The method includes:

[00176] obtaining, by the palm identification device, a palm image 701 corresponding to a palm, and inputting, by the palm identification device, the obtained palm image 701 into a backbone network 702 for feature extraction, to obtain image features at a plurality of sizes corresponding to the palm image 701.

[00177] The palm identification device performs, through a neck network 703, feature fusion on the image features of the plurality of sizes, to obtain a fused image feature.

[00178] The palm identification device performs prediction through a prediction network 704 based on the fused image feature, to obtain a palm area corresponding to a palm in the palm image 701. The palm identification device obtains a palm part image 705 corresponding to the palm by cropping the palm image 701 based on the palm area.

[00179] The palm identification device inputs the palm part image 705 into an encryption network 706 for image encryption, to obtain encrypted palm part image data 707.

[00180] The palm identification device transmits the encrypted palm part image data 707 through a network 708 to a decryption network 709 in a palm identification server for image decryption and performs palm comparison and identification through a verification network 710, to obtain a user identifier 711 corresponding to the palm image 701.

[00181] In summary, according to the method provided in this embodiment, the palm part image corresponding to the palm is obtained by cropping the palm image, and image encryption, transmission, and decryption are performed only on the palm part image. This reduces time of identifying the palm image and improves identification efficiency of palm identification.

[00182] FIG. 8 is a schematic diagram of cross-device payment based on a palm image processing method according to an exemplary embodiment of this application. The method relates to a user terminal 801, a merchant terminal 803, and a payment application server 802.

[00183] A payment application is installed in the user terminal 801. The user terminal 801 logs in to the payment application based on a user identifier, and establishes a communication connection with the payment application server 802. Through the communication connection, the user terminal 801 may interact with the payment application server 802. The payment application is installed in the merchant terminal 803. The merchant terminal 803 logs in to the payment application based on a merchant identifier, and establishes a communication connection with the payment application server 802. Through the communication connection, the merchant terminal 803 may interact with the payment application server 802.

[00184] A cross-device payment procedure includes the following.

[00185] 1. A user holds the user terminal 801 at home, captures a physical palm of the user through the user terminal 801, to obtain a palm image of the user, logs in to the payment application based on the user identifier, and transmits a palm image registration request to the payment application server 802, where the palm image registration request carries the user identifier and the palm image.

[00186] 2. The payment application server 802 receives the palm image registration request transmitted by the user terminal 801, performs feature extraction on the palm image, to obtain a palm feature of the palm image, stores the palm feature and the user identifier in correspondence, and transmits a palm image binding success notification to the user terminal 801.

[00187] After storing the palm feature and the user identifier in correspondence, the payment application server 802 uses the palm feature as a preset palm feature, and may subsequently determine a corresponding user identifier by using the stored preset palm feature.

[00188] 3. The user terminal 801 receives the palm image binding success notification, displays the palm image binding success notification, and prompts the user that the palm image is bound to the user identifier.

[00189] The user completes registration of the palm image through interaction between the user terminal 801 of the user and the payment application server 802, and may subsequently implement automatic payment through the palm image.

[00190] 4. When the user purchases goods in a store to perform a transaction, the merchant terminal 803 captures the physical palm of the user, to obtain the palm image, and based on the payment application that the merchant identifier logs in to, transmits a payment request to the payment application server 802, where the payment request carries the merchant identifier, a consumption amount, and the palm image.

[00191] 5. After receiving the payment request, the payment application server 802 performs palm comparison and identification on the palm image, determines the user identifier of the palm image, determines an account of the user identifier in the payment application, completes transfer through the account, and transmits a payment completion notification to the merchant terminal 803 after the transfer is completed.

[00192] After the user performs palm image registration by using the user terminal 801, the user can directly perform payment through the palm on the merchant terminal 803, and the user does not need to perform palm image registration on the merchant terminal 803. This implements an effect of cross-device palm image identification, and improves convenience.

[00193] 6. The merchant terminal 803 receives the payment completion notification, displays the payment completion notification, and prompts the user that payment is completed, so that the user completes a transaction with the merchant for an item, and the user can bring the item away.

[00194] In addition, in the foregoing embodiment, the cross-device payment process is implemented through the user terminal 801 and the merchant terminal 803. The merchant terminal 803 may be further replaced with a payment device on a bus, and a cross-device bus riding payment solution is implemented according to the foregoing operations.

[00195] FIG. 9 is a schematic diagram of cross-device identity verification based on a palm image processing method according to an exemplary embodiment of this application.

The method relates to a user terminal 901, an access control device 903, and an access control server 902.

[00196] The user terminal 901 establishes a communication connection with the access control server 902. Through the communication connection, the user terminal 901 can interact with the access control server 902. The access control device 903 establishes a communication connection with the access control server 902. Through the communication connection, the access control device 903 can interact with the access control server 902.

[00197] A cross-device identity verification procedure includes the following.

[00198] 1. A user holds the user terminal 901 at home, captures a physical palm of the user through the user terminal 901, to obtain a palm image of the user, and transmits a palm registration request to the access control server 902, where the palm registration request carries a user identifier and the palm image.

[00199] 2. The access control server 902 receives the palm registration request transmitted by the user terminal 901, performs feature extraction on the palm image, to obtain a palm feature of the palm image, stores the palm feature and the user identifier in correspondence, and transmits a palm binding success notification to the user terminal 901.

[00200] After storing the palm feature and the user identifier in correspondence, the access control server 902 may use the palm feature as a preset palm feature, and may subsequently determine a corresponding user identifier by using the stored preset palm feature.

[00201] 3. The user terminal 901 receives the palm binding success notification, displays the palm binding success notification, and prompts the user that the palm image is bound to the user identifier.

[00202] The user completes registration of the palm image through interaction between the user terminal 901 of the user and the access control server, and may subsequently implement automatic door opening through the palm image.

[00203] 4. When the user comes home from the outside, the access control device 903 captures the physical palm of the user, to obtain the palm image of the user, and transmits an identity verification request to the access control server 902, where the identity verification request carries the verification palm image.

[00204] 5. The access control server 902 receives the identity verification request transmitted by the access control device 903, identifies the verification palm image, to obtain the user identifier of the palm image, determines that the user is a registered user, and transmits a verification success notification to the access control device 903.

[00205] 6. The access control device 903 receives the verification success notification transmitted by the access control server 902, and controls, based on the verification success notification, a home door to be opened, so that the user can enter the room.

[00206] The foregoing embodiment implements the cross-device identity verification process through the user terminal 901 and the access control device 903.

[00207] It may be learned from the foregoing cross-device identity verification scenario that whether in a palm registration stage of interaction between the user terminal 901 and the access control server 902, or in a palm image processing stage of interaction between the another terminal device and the server, after obtaining the palm image, the user terminal 901 or the another terminal device extracts and encrypts an area in which the palm is located in the palm image, transmits the palm image of the area in which the palm is located to the server, so that the server performs palm comparison and identification. In addition, in the palm comparison and identification stage, the access control server 902 compares the palm feature with the preset palm feature, to obtain an identification result of a current user.

[00208] FIG. 10 is a schematic diagram of a palm image processing method according to an exemplary embodiment of this application. The method is performed by a palm identification device. The method includes:

[00209] obtaining, by the palm identification device, a palm image 1001 corresponding to a palm, inputting the palm image 1001 into a focusing layer 1002 in a backbone network, and performing, by the focusing layer 1002, a slicing operation on the palm image 1001, to obtain image features of a plurality of sizes. The image features are input into a convolution layer 1003 for convolution and channel adjustment, and then are input into a bottleneck layer 1004 for feature extraction. A pooling layer 1005 is configured to pool the image features.

[00210] For example, the palm image 1001 is input into the focusing layer 1002, and the focusing layer 1002 performs a slicing operation on the palm image 1001, to obtain the image features of the plurality of sizes that are represented as (3, 64, 1, 1). The image features obtained by performing convolution and channel adjustment at the convolution layer 1003 are represented as (64, 128, 3, 2). After passing through the bottleneck layer 1004, the image

features are represented as $(128, 128)*3$. The image features obtained by performing convolution and channel adjustment at the convolution layer 1003 are represented as $(128, 256, 3, 2)$. The image features obtained by passing through the bottleneck layer 1004 are represented as $(256, 256)*3$. The image features $(256, 256)*3$ obtained by performing convolution and channel adjustment at the convolution layer 1003 are represented as $(256, 512, 3, 2)$. The image features obtained by passing through the bottleneck layer 1004 are represented as $(512, 512)*9$. The image features obtained by performing convolution and channel adjustment at the convolution layer 1003 again are represented as $(512, 1024, 3, 2)$. The image features obtained by pooling the image at the pooling layer 1005 and passing through the bottleneck layer 1004 are represented as $(1024, 1024)*3$. After the image features $(1024, 1024)*3$ are upsampled 1006 and spliced 1007 with the image features $(512, 512)*9$, the image features obtained by passing through the bottleneck layer 1004 are represented as $(1024, 512)*3$. The image features $(1024, 512)*3$ obtained by performing convolution and channel adjustment at the convolution layer 1003 are represented as $(512, 256, 1, 1)$. After the image features $(512, 256, 1, 1)$ are upsampled 1006 and spliced 1007 with the image features $(256, 512, 3, 2)$, the image features obtained by passing through the bottleneck layer 1004 are represented as $(512, 256)*3$. The image features $(512, 256)*3$ obtained by performing convolution and channel adjustment at the convolution layer 1003 are represented as $(256, 256, 3, 2)$. After the image features $(256, 256, 3, 2)$ are spliced 1007 with the image features $(512, 256, 1, 1)$, the image features obtained by passing through the bottleneck layer 1004 are represented as $(512, 512)*3$. The image features $(512, 512)*3$ obtained by performing convolution and channel adjustment at the convolution layer 1003 are represented as $(512, 512, 3, 2)$. After the image features $(512, 512, 3, 2)$ are spliced 1007 with the image features represented as $(1024, 512, 1, 1)$, the image features obtained by passing through the bottleneck layer 1004 are represented as $(1024, 1024)*3$. The finally output image features 1008 of the plurality of sizes are: $(512, 256)$, $(512, 512)$, and $(1024, 1024)$. (x, y, a, b) , where x and y may respectively represent two dimensions of a feature matrix, and a and b may respectively represent a quantity of channels and a size. $(x, y)*b$, where x and y may respectively represent two dimensions of the feature matrix, and b may represent the quantity of channels. (x, y) , where x and y may respectively represent two dimensions of the feature matrix.

[00211] The palm identification device performs feature fusion on the image features 1008 of the plurality of sizes, to obtain a fused image feature. The palm identification device

performs prediction based on the fused image feature, to obtain a palm area corresponding to a palm in the palm image 1001. The palm identification device obtains a palm part image corresponding to the palm by cropping the palm image 1001 based on the palm area. The palm identification device performs image encryption on the palm part image, and transmits the palm part image to the palm identification server through the network for image decryption and palm comparison and identification, to obtain a user identifier corresponding to the palm image 1001.

[00212] In summary, according to the method provided in this embodiment, the slicing operation is performed on the palm image through the backbone network and feature extraction is performed, to obtain the image features of the plurality of sizes. Based on the foregoing method, features of the plurality of sizes, under different receptive fields, and in different categories are obtained. This improves accuracy of prediction of the palm area.

[00213] The prediction of the palm area related in this application may be implemented based on a palm part detection model. The solution includes a palm part detection model generation stage and a palm area prediction stage. FIG. 11 is a framework diagram of a palm part detection model generation stage and a palm area prediction stage according to an exemplary embodiment of this application. As shown in FIG. 11, in the palm part detection model generation stage, after obtaining a palm part detection model by using a preset training sample dataset (including a sample palm image and a sample palm part image), a palm part detection model generation device 1110 generates a palm area based on the palm part detection model. In the palm area prediction stage, the palm area prediction processing device 1120 processes the input palm image based on the palm part detection model, to obtain a prediction result of the palm image, for example, determines an area and a position of a palm in the palm image.

[00214] The foregoing palm part detection model generation device 1110 and the palm area prediction processing device 1120 may be computer devices. For example, the computer devices may be fixed computer devices such as a personal computer or a server, or the computer devices may also be mobile computer devices such as a tablet computer or an e-book reader.

[00215] In some embodiments, the palm part detection model generation device 1110 and the palm area prediction processing device 1120 may be the same device, or the palm part detection model generation device 1110 and the palm area prediction processing device

1120 may be different devices. In addition, when the palm part detection model generation device 1110 and the palm area prediction processing device 1120 are different devices, the palm part detection model generation device 1110 and the palm area prediction processing device 1120 may be the same type of devices. For example, both the palm part detection model generation device 1110 and the palm area prediction processing device 1120 may be servers. Alternatively, the palm part detection model generation device 1110 and the palm area prediction processing device 1120 may be different types of devices. For example, the palm area prediction processing device 1120 may be a personal computer or a terminal, and the palm part detection model generation device 1110 may be a server or the like. Specific types of the palm part detection model generation device 1110 and the palm area prediction processing device 1120 are not limited in the embodiments of this application.

[00216] The foregoing embodiments describe the palm image processing method, and the following describes a method for training a palm part detection model.

[00217] FIG. 12 is a flowchart of a method for training a palm part detection model according to an exemplary embodiment of this application. The method is performed by a palm identification device. The method includes the following operations.

[00218] Operation 1201: Obtain a sample palm image and a sample palm part image.

[00219] The sample palm image is an image of a to-be-determined user identifier. The sample palm image includes a palm. The palm is an image of a physical palm of a user whose identity is to be verified. The sample palm image may further include other information, such as fingers of the user, a scene in which the palm of the user is captured, and the like. The sample palm image may be obtained by capturing, by the palm identification device, a physical palm of a user whose identity is to be verified, or may be transmitted by another device.

[00220] For example, the palm identification device is a cash register device in a store, and the cash register device in the store captures a physical palm of a user through a camera, to obtain the sample palm image. Alternatively, the palm identification device is a palm image identification server, and after capturing a sample palm image of a user, a cash register device in a store transmits the sample palm image to the palm image identification server.

[00221] The sample palm part image is an effective identification area labeled in the sample palm image, or the sample palm part image is an area in the sample palm image in

which the labeled palm is located, or the sample palm part image is an area in the sample palm image that is labeled and that can be configured for identification.

[00222] Operation 1202: Perform feature extraction on the sample palm image, to obtain sample image features of a plurality of sizes.

[00223] The sample palm image is an image obtained by capturing a physical palm. The sample palm image includes an image of all or a part of the physical palm. For ease of description, an image of the physical palm in the sample palm image is referred to as a palm subsequently. The palm included in the sample palm image may be a physical palm of a user whose identity identification is to be performed. The palm in the sample palm image includes at least a palm, may further include the fingers or a part of the arm, and may further include an image of an environment in which the physical palm is located when the physical palm is captured.

[00224] A size is a size of a sample image feature, and may be represented by a matrix dimension number, such as 100*100, 80*80, 60*60, or 40*40. The plurality of sizes refer to two or more sizes, and each size is different. A maximum size in the plurality of sizes may be the same as a size of the sample palm image, and is generally smaller than the size of the sample palm image.

[00225] The sample palm image may be obtained by capturing, by the palm identification device, a physical palm of a user whose identity is to be verified, or may be transmitted by another device. The palm identification device may directly collect an image of the physical palm, to obtain the sample palm image. The palm identification device may also collect the image of the physical palm through an image collector connected to the sample palm image, to obtain the sample palm image. The palm identification device can perform feature extraction on the obtained sample palm image, to obtain sample image features of a plurality of sizes corresponding to the sample palm image.

[00226] For example, the palm identification device is a cash register device in a store, and the cash register device in the store captures a physical palm of a user through a camera, to obtain the sample palm image. Alternatively, the palm identification device is a sample palm image identification server, and after capturing a sample palm image of a user, a cash register device in a store transmits the sample palm image to the sample palm image identification server.

[00227] For example, the sample image features of the plurality of sizes may be respectively extracted from low-resolution images of the plurality of sizes by downsampling a high-resolution sample palm image into the low-resolution images of the plurality of sizes. The high resolution and the low resolution are relative concepts, and the high resolution is a resolution higher than the low resolution.

[00228] For example, a size of the input sample palm image is 640*640, and sizes of the output sample image features of the plurality of sizes are: 80*80, 40*40, and 20*20.

[00229] Operation 1203: Fuse the sample image features of the plurality of sizes, to obtain a sample fused image feature.

[00230] The palm identification device may perform feature fusion on the sample image features of the plurality of sizes, to obtain a sample fused image feature.

[00231] The sample image features of the plurality of sizes refer to features obtained by performing feature extraction on images of the plurality of sizes. Sample image features at different sizes may express different information. In the sample image features of the plurality of sizes, sample image features of larger sizes express more position information, and sample image features of smaller sizes express more semantic information.

[00232] For example, the palm identification device extracts features of the plurality of sizes, under different receptive fields, and in different categories, and fuses the sample image features of the plurality of sizes, to obtain the sample fused image feature corresponding to the sample palm image.

[00233] Operation 1204: Determine, based on the sample fused image feature, a sample prediction palm area identifying a palm in the sample palm image.

[00234] The sample prediction palm area is an identifier configured for representing a position of the palm in the sample palm image. The sample prediction palm area includes the palm in the sample palm image. A boundary of the sample prediction palm area encloses the palm in the sample palm image.

[00235] In some embodiments, a shape of the sample prediction palm area may be a regular shape, for example, may be at least one of a rectangle, a square, a rhombus, a hexagon, a circle, or a triangle. The shape of the sample prediction palm area may alternatively be an irregular shape. The regular shape is a shape that can be drawn by using a fixed parameter and based on a fixed rule. The irregular shape cannot be drawn by using the

fixed parameter and based on the fixed rule, and can generally only be represented by a set of points that constitute the irregular shape.

[00236] The palm identification device can perform prediction based on the sample fused image feature by using artificial intelligence, to determine the sample prediction palm area identifying the palm in the sample palm image.

[00237] In some embodiments, a computer device can determine, based on the sample fused image feature, the position of the sample prediction palm area in the sample palm image that identifies the palm in the sample palm image. When the sample prediction palm area is of the regular shape, the position may be represented by a value of the fixed parameter required for drawing the regular shape, for example, may be represented by a position of a fixed point of the sample prediction palm area and a size parameter of the sample prediction palm area. The fixed point is a point whose position is fixed relative to the sample prediction palm area. The size parameter is a parameter representing a size of the sample prediction palm area. For example, when the sample prediction palm area is of the rectangle shape, the fixed point may be any corner point or a center point of the rectangle, and the size parameter is a side length of an adjacent side of the sample prediction palm area.

[00238] Operation 1205: Extract out a sample prediction palm part image from the sample palm image based on the sample prediction palm area.

[00239] The sample prediction palm part image is an effective identification area in the sample palm image, or the sample prediction palm part image is an area in the sample palm image in which the palm is located, or the sample prediction palm part image is an area in the sample palm image that can be configured for palm identification.

[00240] For example, the palm identification device extracts, based on the position of the sample prediction palm area, an image area located outside the sample prediction palm area in the sample palm image, and uses a remaining image as the sample prediction palm part image.

[00241] Operation 1206: Calculate a loss function value based on the sample palm part image and the sample prediction palm part image.

[00242] For example, the palm identification device calculates the loss function value based on the sample palm part image and the sample prediction palm part image.

[00243] Operation 1207: Update a model parameter of the palm part detection model based on the loss function value.

[00244] For example, the palm identification device updates the model parameter of the palm part detection model based on the loss function value.

[00245] Updating of the model parameter refers to updating a network parameter in the palm part detection model, or updating a network parameter of each network module in the model, or updating a network parameter of each network layer in the model, but is not limited thereto. This is not limited in the embodiments of this application.

[00246] The model parameter of the palm part detection model includes at least one of a network parameter of a backbone network, a network parameter of a neck network, and a network parameter of a prediction network in the palm part detection model.

[00247] In summary, according to the method provided in this embodiment, the sample image features of the plurality of sizes corresponding to the sample palm image are obtained by performing feature extraction on the sample palm image; feature fusion is performed on the sample image features of the plurality of sizes, to obtain the sample fused image feature; prediction is performed based on the sample fused image feature, to obtain the sample prediction palm area identifying the palm in the sample palm image; the sample palm image is cropped based on the sample prediction palm area to obtain the sample prediction palm part image corresponding to the palm; a loss function value is calculated based on the sample prediction palm part image and the sample prediction palm part image; and the model parameter of the palm part detection model is updated based on the loss function value, so that the trained palm part detection model can have higher prediction accuracy of the sample prediction palm area, to obtain a more accurate sample prediction palm area.

[00248] FIG. 13 is a flowchart of a method for training a palm part detection model according to an exemplary embodiment of this application. The method is performed by a palm identification device. The method includes the following operations.

[00249] Operation 1301: Obtain a sample palm image and a sample palm part image.

[00250] The sample palm image is an image of a to-be-determined user identifier. The sample palm image includes a palm. The palm is an image of a physical palm of a user whose identity is to be verified. The sample palm image may further include other information, such as fingers of the user, a scene in which the palm of the user is captured, and the like. The sample palm image may be obtained by capturing, by the palm identification device, a

physical palm of a user whose identity is to be verified, or may be transmitted by another device.

[00251] The sample palm part image is an effective identification area labeled in the sample palm image, or the sample palm part image is an area in the sample palm image in which the labeled palm is located, or the sample palm part image is an area in the sample palm image that is labeled and that can be configured for identification.

[00252] For example, the sample palm image is collected in a manner in which 50 to 100 sample palm images are obtained. The sample palm image is cleaned up by finding a duplicate value, a missing value, and an outlier, so that a valid sample palm image is obtained. Code for cleaning up the sample palm image may be represented as:

```
allocate output Pixel Value [image width][image height]

allocate window [window width×window height]

edge x=(window width/2) rounded down
edge y=(window height/2) rounded down

for x from edge x to image width−edge x do
  for y from edge y to image height−edge y do
    i=0
    for fx from 0 to window width do
      for fy from 0 to window height do
        window[i]=input Pixel Value[x+fx−edge x][y+fy−edge y]
        i=i+1
      sort entries in window []
    output Pixel Value[x][y]=window [window width*window height/2]
```

[00253] A collection distance requirement of the sample palm image is that the palm is placed at a position of 5 cm, 8 cm, 10 cm, 12 cm, or 15 cm from the palm identification device for collection, and collection is performed at each distance for 10 to 15 seconds.

[00254] The sample palm part image is obtained in a manner in which a palm area in the sample palm image is labeled by using a labeling tool Labeling. Code for labeling the palm area may be represented as:

```

<object>

  <name>plam</name>

  <pose>Unspecified</pose>

  <truncated>0</truncated>

  <difficult>0</difficult>

  <bndbox>

    <xmin>45</xmin>

    <ymin>34</ymin>

    <xmax>104</xmax>

    <ymax>84</ymax>

  </bndbox>

</object>

```

[00255] In the foregoing code, 45 represents a minimum value in the palm area on an x-axis, 104 represents a maximum value in the palm area on the x-axis, 34 represents a minimum value in the palm area on a y-axis, and 84 represents a minimum value in the palm area on the y-axis.

[00256] The labeling tool Labeling replaces the conventional one-hot encoded label vector y_{hot} with the updated label vector $\hat{y}_i$ with reference to the uniform distribution theorem.

$$\hat{y}_i = y_{hot}(1 - \alpha) + \alpha/k$$

[00257] In the formula, k is a total quantity of categories of the multi-classification, and α is a small hyperparameter (generally 0.1).

[00258] That is, the label vector $\hat{y}_i$ may be represented as:

$$\hat{y}_i = \begin{cases} 1 - \alpha, & i = a \text{ quantity of categories} \\ \frac{\alpha}{k}, & i \neq a \text{ quantity of categories} \end{cases}$$

[00259] The label vector $\hat{y}_i$ is equivalent to adding noise to a real distribution, to prevent the model from being too confident about a correct label, so that output values of

predicted positive and negative samples do not differ as greatly. This avoids overfitting and improves a generalization capability of the model.

[00260] Operation 1302: Perform feature extraction on the sample palm image, through a backbone network in a palm part detection model, to obtain sample image features of a plurality of sizes.

[00261] For example, the palm identification device captures a physical palm of the user, to obtain the sample palm image. The sample palm image includes the palm. The palm may be a left palm of the user or a right palm of the user.

[00262] For example, the palm identification device is an internet of things device. The internet of things device captures the left palm of the user through a camera, to obtain the sample palm image. The internet of things device may be a cash register device in a store. For another example, when the user shops for a transaction in the store, the user extends a palm toward a camera of a payment terminal in the store, and the payment terminal in the store captures the physical palm of the user through the camera, to obtain the sample palm image.

[00263] In some embodiments, the palm identification device establishes a communication connection with another device, and receives, through the communication connection, a sample palm image transmitted by the another device. For example, the palm identification device is a payment application server, and the another device may be a payment terminal. After the payment terminal captures the physical palm of the user, to obtain the sample palm image, the sample palm image is transmitted to the payment application server through the communication connection between the payment terminal and the payment application server, so that the payment application server can determine a user identifier of the sample palm image.

[00264] In some embodiments, the sample prediction palm area is obtained by prediction by a palm part detection model, and the palm part detection model includes a backbone network; and the performing feature extraction on the sample palm image, to obtain sample image features at a plurality of sizes includes: inputting the sample palm image into the backbone network; performing a slicing operation on the sample palm image through the backbone network, to obtain slice images of the plurality of sizes; and respectively performing feature extraction on the slice images of the plurality of sizes through the backbone network, to obtain the sample image features of the plurality of sizes.

[00265] The palm part detection model is an artificial intelligence model, may be a deep learning model, or may be a neural network model. The palm part detection model has at least a function of determining, based on a sample fused image feature, a sample prediction palm area identifying a palm in the sample palm image. The backbone network is a structure configured for extracting sample image features at a plurality of sizes from the sample palm image in the palm part detection model. The slicing operation is an operation of segmenting the sample palm image to generate images of smaller sizes. The smaller size means that a size of the generated image is smaller than the size of the sample palm image. Pixels in the slice image are derived from the sample palm image.

[00266] The palm identification device may input the sample palm image into the backbone network, and the backbone network performs a slicing operation, to obtain the slice images of the plurality of sizes. The palm identification device may perform feature extraction on the slice images of the plurality of sizes, to obtain the sample image features of the plurality of sizes corresponding to the sample palm image.

[00267] In some embodiments, the performing a slicing operation on the sample palm image through the backbone network, to obtain slice images of the plurality of sizes includes: determining, through the backbone network and based on the sample palm image, a slice image of a maximum size in the plurality of sizes; and using, through the backbone network, the slice image of the maximum size as a first layer, and starting downsampling layer by layer, to obtain the slice images of the plurality of sizes including the slice image at the first layer.

[00268] A computer device may first determine, based on the sample palm image, the slice image of the maximum size in the plurality of sizes. The size of the sample image feature of the maximum size may be smaller than or equal to the size of the sample palm image. The computer device may downsample the sample palm image once, to generate the slice image of the maximum size in the plurality of sizes.

[00269] Downsampling the sample palm image is a process of sampling at least one pixel in the sample palm image. Pixels obtained through downsampling are spliced in sequence in the sample palm image, so that a slice image at a smaller size can be obtained.

[00270] If a quantity of the plurality of sizes is N , the plurality of sizes are N sizes. The N sizes are all different sizes, and therefore, there is a maximum size in the N sizes. The slice image of the maximum size is used as a first layer, downsampling layer by layer is started

from the first layer, an n^{th} layer is downsampled to obtain a slice image at an $(n+1)^{\text{th}}$ layer, and an $(N-1)^{\text{th}}$ layer is downsampled to obtain a slice image at the N^{th} layer. n is a positive integer from 1 to $N-1$. Finally, a total of N slice images at 1 to N layers may be obtained.

[00271] In this embodiment, the slice images of the plurality of sizes are obtained by downsampling layer by layer, which provides a new way to obtain the slice images of the plurality of sizes by segmenting the sample palm image.

[00272] In some embodiments, the performing a slicing operation on the sample palm image through the backbone network, to obtain slice images of the plurality of sizes includes: downsampling and splicing, through the backbone network, pixels in the sample palm image of the plurality of sizes, to obtain the slice images of the plurality of sizes. There are two adjacent pixels in slice images of different sizes, and a quantity of pixels between the two pixels at which sampling occurs in the sample palm image is different.

[00273] In this embodiment, the palm identification device may sample the pixels from the sample palm image of each size by using a sampling interval at the size, and splice the sampled pixels according to a relative position relationship in the sample palm image, to obtain a slice image of each size. The sampling interval is a quantity of pixels that are spaced, including intervals of pixels in two directions: a row and a column. The size is negatively correlated with the sampling interval. A larger size indicates a smaller sampling interval. A smaller size indicates a larger sampling interval. Pixels that are not sampled in the sample palm image may be concentrated to a channel of a sample image feature.

[00274] For example, for a maximum size in the plurality of sizes, in the sample palm image, sampling is performed on every other column. That is, a pixel is selected every other pixel in the row direction, and the selected pixels are spliced, to obtain the slice image. Information in the spliced slice image is not lost, but size information in the sample palm image is concentrated on the channel, and a convolution operation is performed on an obtained new image, to obtain a downsampling feature map without information loss.

[00275] For another example, for a second largest size in the plurality of sizes, in the sample palm image, sampling is performed in every two columns. That is, pixels are selected every two pixels, and the selected pixels are spliced, to obtain the slice image. Information in the spliced slice image is not lost, but size information in the sample palm image is concentrated on the channel, and a convolution operation is performed on an obtained new image, to obtain a downsampling feature map without information loss.

[00276] For example, a size of an original image is: $640 \times 640 \times 3$. Through the slicing operation, a feature map of $320 \times 320 \times 12$ is obtained. 640×640 in the original image is configured for representing width*height, and 3 in the original image is configured for representing a length (also referred to as a quantity of channels) of a feature vector corresponding to each pixel.

[00277] Operation 1303: Perform, through a neck network of the palm part detection model, feature fusion on the sample image features of the plurality of sizes, to obtain the sample fused image feature.

[00278] The sample prediction palm area is obtained by prediction by the palm part detection model, and the palm part detection model includes a neck network; and the fusing the sample image features of the plurality of sizes, to obtain a sample fused image feature includes: performing, through the neck network, feature fusion on the sample image features of the plurality of sizes, to obtain the sample fused image feature.

[00279] The neck network belongs to the palm part detection model and is a structure configured for performing feature fusion on the sample image features of the plurality of sizes to obtain the sample fused image feature. The neck network is connected to the backbone network, and is configured to receive the sample image features of the plurality of sizes output by the backbone network.

[00280] The sample image features of the plurality of sizes refer to features obtained by performing feature extraction on images of the plurality of sizes. Sample image features at different sizes may express different information. In the sample image features of the plurality of sizes, sample image features of larger sizes express more position information, and sample image features of smaller sizes express more semantic information.

[00281] For example, the palm identification device extracts features of the plurality of sizes, under different receptive fields, and in different categories, and fuses the sample image features of the plurality of sizes, to obtain the sample fused image feature corresponding to the sample palm image.

[00282] In some embodiments, the performing, through the neck network, feature fusion on the sample image features of the plurality of sizes, to obtain the sample fused image feature includes: performing, through the neck network, fusion at the same size on the sample image features of the plurality of sizes, to obtain the sample fused image feature. The fusion of the sample image features of the same size may be summation or averaging.

[00283] The palm identification device fuses the sample image features of the plurality of sizes, which increases detection accuracy of a small target, so that the sample prediction palm area can be detected more accurately. In this way, the sample prediction palm part image can be extracted more accurately. This further reduces a data amount of the sample prediction encrypted palm part image data, improves identity identification efficiency, and ensures security of identity identification.

[00284] In some embodiments, the performing, through the neck network, feature fusion on the sample image features of the plurality of sizes, to obtain the sample fused image feature includes: performing, through the neck network, feature fusion based on the sample image features of the plurality of sizes, to obtain first intermediate features of the plurality of sizes; performing, through the neck network, feature fusion of the plurality of sizes based on the first intermediate features of the plurality of sizes, to obtain second intermediate features of the plurality of sizes; and performing, through the neck network, feature fusion on the second intermediate features of the plurality of sizes, to obtain the sample fused image feature.

[00285] The feature fusion of the plurality of sizes refers to performing feature fusion of different sizes respectively. The first intermediate features of the plurality of sizes may be in a one-to-one correspondence with the sample image features of the plurality of sizes. The first intermediate features of the plurality of sizes may be in a one-to-one correspondence with the second intermediate features of the plurality of sizes. The sample image features of the plurality of sizes may be consistent with the first intermediate features of the plurality of sizes. The first intermediate features of the plurality of sizes may be consistent with the second intermediate features of the plurality of sizes. The performing feature fusion on the second intermediate features of the plurality of sizes may be fusing, for the same size, the second intermediate features of the plurality of sizes. The fusion may be summation or averaging.

[00286] In this embodiment, features of different sizes are fused and are finally fused into the sample fused image feature, which can fully express detailed features in the sample palm image, so that a sample prediction palm area is detected more accurately. In this way, the sample prediction palm part image can be extracted more accurately. This further reduces a data amount of the sample prediction encrypted palm part image data, improves identity identification efficiency, and ensures security of identity identification.

[00287] In some embodiments, the plurality of sizes are N sizes, N is a positive integer greater than 1, and the extracting, based on the sample image features of the plurality of sizes, first intermediate features of the plurality of sizes includes: determining, based on a sample image feature ranking first in a positive order of size, a first intermediate feature ranking first in the positive order of size; upsampling layer by layer starting from the first intermediate feature ranking first in the positive order of size, and fusing a result of upsampling a first intermediate feature ranking n^{th} in the positive order of size with a sample image feature ranking $(n+1)^{\text{th}}$ in the positive order of size, to obtain a first intermediate feature ranking $(n+1)^{\text{th}}$ in the positive order of size; and determining, when obtaining a first intermediate feature at an N^{th} layer, the first intermediate features of the plurality of sizes from a first layer to the N^{th} layer. n is a positive integer from 1 to $N-1$; and the positive order of size is in ascending order of size.

[00288] The ascending order is referred to as the positive order, and the positive order of size is in ascending order of size. Upsampling is a process of increasing the size. The upsampling may use a size at a next layer as a target, and adjacent feature elements in a feature at each layer are interpolated, to obtain a feature at the next layer.

[00289] In some embodiments, the determining, based on a sample image feature ranking first in a positive order of size, a first intermediate feature ranking first in the positive order of size includes: performing at least one time of convolution on the sample image feature ranking first in the positive order of size, to obtain the first intermediate feature ranking first in the positive order of size. In addition to performing convolution, channel adjustment may be further performed. A quantity of channels of the first intermediate features after the channel adjustment is performed may be consistent.

[00290] In some embodiments, a size of a result of upsampling the first intermediate feature ranking n^{th} in the positive order of size may be the same as a size of a sample image feature ranking $(n+1)^{\text{th}}$ in the positive order of size. The result of upsampling the first intermediate feature ranking n^{th} in the positive order of size and the sample image feature ranking $(n+1)^{\text{th}}$ in the positive order of size are fused, and the two may be spliced, or summation or averaging may be performed on the two at the same size.

[00291] For example, FIG. 5 is a schematic diagram of sample image feature fusion. A palm identification device performs feature extraction on a sample palm image 501, to obtain sample image features 502 at a plurality of sizes corresponding to the sample palm image

501. The sample image features 502 include a feature layer T1, a feature layer T2, a feature layer T3, and a feature layer T4. After the sample image feature 502 is obtained, one time of convolution is performed on the feature layer T4 and channel adjustment is performed to obtain a feature layer P4. The feature layer P4 is upsampled and is combined with the feature layer T3, and then convolution is performed and channel adjustment is performed to obtain a feature layer P3. The feature layer P3 is upsampled and is combined with the feature layer T2, and then convolution is performed and channel adjustment is performed to obtain a feature layer P2. The feature layer P2 is upsampled and is combined with the feature layer T1, and then 1*1 convolution is performed and channel adjustment is performed to obtain a feature layer P1.

[00292] In some embodiments, the plurality of sizes are N sizes, N is a positive integer greater than 1, and the performing feature fusion of the plurality of sizes based on the first intermediate features of the plurality of sizes, to obtain second intermediate features of the plurality of sizes includes: determining, based on a first intermediate feature ranking first in a reverse order of size, a second intermediate feature ranking first in the reverse order of size; downsampling layer by layer starting from the second intermediate feature ranking first in the reverse order of size, and fusing a result of downsampling a second intermediate feature ranking m^{th} in the reverse order of size with a first intermediate feature ranking $(m+1)^{\text{th}}$ in the reverse order of size, to obtain a second intermediate feature ranking $(n+1)^{\text{th}}$ in the reverse order of size, where m is an integer from 1 to N-1; and the reverse order of size is in descending order of size; and determining, when obtaining the first intermediate feature at the N^{th} layer, the first intermediate features of the plurality of sizes from the first layer to the N^{th} layer.

[00293] The descending order is referred to as the reverse order, and the reverse order of size is in descending order of size. Downsampling is a process of reducing a size. The downsampling may use a size at a next layer as a target, and feature elements in a feature at each layer are used at intervals, to obtain a feature at the next layer. The first intermediate feature may be referred to as a semantic sample image feature. The second intermediate feature may be referred to as a position sample image feature. The semantic sample image features represent more semantic information, and the semantic sample image features represent more position information.

[00294] In some embodiments, the determining, based on a first intermediate feature ranking first in a reverse order of size, a second intermediate feature ranking first in the

reverse order of size includes: performing at least one time of convolution on the first intermediate feature ranking first in the reverse order of size, to obtain the second intermediate feature ranking first in the reverse order of size. In addition to performing convolution, channel adjustment may be further performed. A quantity of channels of the first intermediate features after the channel adjustment is performed may be consistent.

[00295] A size of a result of downsampling the second intermediate feature ranking m^{th} in the reverse order of size may be consistent with a size of a first intermediate feature ranking $(m+1)^{\text{th}}$ in the reverse order of size. The result of downsampling the second intermediate feature ranking m^{th} in the reverse order of size and the first intermediate feature ranking $(m+1)^{\text{th}}$ in the reverse order of size are fused, and the two may be spliced, or summation or averaging may be performed on the two at the same size.

[00296] For example, the palm identification device may perform one time of convolution on the feature layer P1 and perform channel adjustment to obtain a feature layer N1. The feature layer N1 is downsampled and is combined with the feature layer P2, and then convolution is performed and channel adjustment is performed to obtain a feature layer N2. The feature layer N2 is downsampled and is combined with the feature layer P3, and then convolution is performed and channel adjustment is performed to obtain a feature layer N3. The feature layer N3 is downsampled and is combined with the feature layer P4, and then convolution is performed and channel adjustment is performed to obtain a feature layer N4. The palm identification device fuses features at a feature layer N1–4, to obtain a fused image feature 503.

[00297] In some embodiments, an activation function in the convolution is a logistic activation function.

[00298] In some embodiments, the palm identification device upsamples the sample image features of the plurality of sizes in ascending order of size, to obtain the first intermediate feature; the palm identification device downsamples the sample image features of the plurality of sizes in descending order of size, to obtain the second intermediate feature; and the palm identification device performs feature fusion on the first intermediate feature and the second intermediate feature, to obtain the sample fused image feature.

[00299] Operation 1304: Determine, through a prediction network of the palm part detection model based on the sample fused image feature, a sample prediction palm area identifying a palm in the sample palm image.

[00300] The sample prediction palm area is an identifier configured for representing a position of the palm in the sample palm image.

[00301] In some embodiments, the sample prediction palm area is determined by the palm part detection model, and the palm part detection model includes a prediction network; and the determining, based on the sample fused image feature, a sample prediction palm area identifying a palm in the sample palm image includes: performing grid division on the sample fused image feature, to obtain a plurality of grid features of the sample fused image feature; and inputting the plurality of grid features into the prediction network, and performing prediction on the plurality of grid features, to obtain the sample prediction palm area.

[00302] For example, the palm identification device performs grid division on the sample fused image feature, to obtain gridded sample fused image features; and the palm identification device inputs the gridded fusion features into the prediction network, and the prediction network performs prediction on each grid, to obtain the sample prediction palm area. Prediction is performed on the plurality of grid features, to obtain the sample prediction palm area. Specifically, a position of the sample prediction palm area may be determined based on prediction results of the plurality of grid features.

[00303] In some embodiments, the palm identification device is equipped with an infrared camera, and the sample palm image processing method further includes: obtaining an infrared image collected of a palm the same as the sample palm image, where the infrared image is an image obtained by the infrared camera imaging, based on infrared light, the palm; identifying a palm area from the infrared image; and the extracting a sample prediction palm part image from the sample palm image based on the sample prediction palm area includes: determining an overlapping area between the sample prediction palm area in the sample palm image and the palm area in the infrared image; and extracting a sample prediction palm part image from the sample palm image based on the overlapping area.

[00304] In this embodiment, the palm identification device is further equipped with an infrared camera. The palm identification device obtains an infrared image corresponding to the same physical palm. The palm identification device performs area identification on the infrared image, to determine a palm area in the infrared image. The infrared image is an image obtained by the infrared camera imaging, based on infrared light, the palm.

[00305] In some embodiments, the identifying a palm area from the infrared image includes: detecting a finger gap point in the infrared image; and determining the palm area in the infrared image based on the finger gap point.

[00306] For example, the palm identification device performs detection on the finger gap point in the infrared image, and determines the palm area in the infrared image based on the finger gap point.

[00307] FIG. 6 is a schematic diagram of a finger gap point in a palm. The finger gap point is a first finger gap point 601 between an index finger and a middle finger, or the finger gap point is a second finger gap point 602 between the middle finger and a ring finger, or the finger gap point is a third finger gap point 603 between the ring finger and a little finger.

[00308] Because a palm area in a sample palm image may exist in any area in the sample palm image, to determine a position of the palm area in the sample palm image, finger gap point detection is performed on the sample palm image, to obtain at least one finger gap point in the sample palm image, so that the palm area can be determined subsequently according to the at least one finger gap point.

[00309] Operation 1305: Extract a sample prediction palm part image from the sample palm image based on the sample prediction palm area.

[00310] The sample prediction palm part image is an effective identification area predicted in the sample palm image, or the sample prediction palm part image is an area in the sample palm image in which the predicted palm is located, or the sample prediction palm part image is an area in the sample palm image that is predicted and that can be configured for identification.

[00311] In some embodiments, the palm identification device obtains at least three finger gap points in the infrared image; the finger gap points are sequentially connected, to obtain a finger gap point connecting line; and the palm identification device extracts the infrared image based on the finger gap point connecting line to obtain the palm area.

[00312] In some embodiments, the determining an overlapping area between the sample prediction palm area in the sample palm image and the palm area in the infrared image includes: obtaining a coordinate parameter of the sample prediction palm area in the sample palm image, and obtaining a coordinate parameter of the palm area in the infrared image; and determining an overlapping area between the sample prediction palm area and the

palm area based on the coordinate parameter of the sample prediction palm area and the coordinate parameter of the palm area.

[00313] The palm identification device obtains a coordinate parameter of the sample prediction palm area in the sample palm image, and obtains a coordinate parameter of the palm area in the infrared image; and the palm identification device extracts, based on the overlapping area between the coordinate parameter of the sample prediction palm area and the coordinate parameter of the palm area, the sample palm image to obtain the sample prediction palm part image corresponding to the palm.

[00314] Operation 1306: Calculate a loss function value according to the sample palm part image and the sample prediction palm part image.

[00315] For example, the palm identification device calculates the loss function value based on the sample palm part image and the sample prediction palm part image.

[00316] Operation 1307: Update a model parameter of the palm part detection model based on the loss function value.

[00317] Updating of the model parameter refers to updating a network parameter in the palm part detection model, or updating a network parameter of each network module in the model, or updating a network parameter of each network layer in the model, but is not limited thereto. This is not limited in the embodiments of this application.

[00318] The model parameter of the palm part detection model includes at least one of a network parameter of a backbone network, a network parameter of a neck network, and a network parameter of a prediction network.

[00319] In some embodiments, updating the model parameter of the palm part detection model includes updating network parameters of all network modules in the palm part detection model, or fixing network parameters of some network modules in the palm part detection model, and updating only network parameters of the remaining part of network modules. For example, when the model parameter of the palm part detection model is updated, the network parameter of the backbone network in the palm part detection model is fixed, and only the network parameter of the neck network and the network parameter of the prediction network are updated.

[00320] The network parameter of the backbone network, the network parameter of the neck network, and the network parameter of the prediction network in the palm part detection

model are updated based on the loss function value by using the loss function value as a training indicator until the loss function value converges, to obtain the palm part detection model whose training is completed.

[00321] That the loss function value converges refers to at least one of the following cases: The loss function value no longer changes, or an error difference between two adjacent iterations during training of the palm part detection model is less than a preset value, or a quantity of times of training of the palm part detection model reaches a preset quantity of times, but is not limited thereto. This is not limited in the embodiments of this application.

[00322] In some embodiments, a target condition satisfied by the training may be that a quantity of times of iterations of training of an initial model reaches a target quantity of times. A technician may preset the quantity of times of iterations of training. Alternatively, the target condition satisfied by the training may be that the loss value satisfies a target threshold condition, but is not limited thereto. This is not limited in the embodiments of this application.

[00323] In summary, according to the method provided in this embodiment, the sample image features of the plurality of sizes corresponding to the sample palm image are obtained by performing feature extraction on the sample palm image; feature fusion is performed on the sample image features of the plurality of sizes, to obtain the sample fused image feature; prediction is performed based on the sample fused image feature, to obtain the sample prediction palm area identifying the palm in the sample palm image; the sample palm image is extracted based on the sample prediction palm area to obtain the sample prediction palm part image corresponding to the palm; a loss function value is calculated based on the sample palm part image and the sample prediction palm part image; and the model parameter of the palm part detection model is updated based on the loss function value, so that the trained palm part detection model can have higher prediction accuracy of the palm area, to obtain a more accurate palm area.

[00324] For example, application scenarios of the palm image processing method provided in the embodiments of this application include but are not limited to the following scenarios.

[00325] For example, in a palm image identification payment scenario:

[00326] A palm identification device of a merchant captures a physical palm of a user, to obtain a palm image of the user, determines a target user identifier of the palm image by

using the palm image processing method provided in the embodiments of this application, and transfers some resources in a resource account corresponding to the target user identifier to a merchant resource account, to implement automatic payment through the palm.

[00327] For another example, in a cross-device payment scenario:

[00328] A user may complete identity registration at home or another private space by using a personal mobile phone, and bind an account of the user to a palm image of the user. Then the palm image of the user may be identified on an in-store device, to determine the account of the user, and the user directly performs payment through the account.

[00329] For another example, in a check scenario at work:

[00330] A palm identification device obtains a palm image of a user by capturing a physical palm of the user. The palm image processing method provided in the embodiments of this application is used, so that a target user identifier of the palm image is determined, a check in label is established for the target user identifier, and it is determined that the target user identifier has completed check in at work at a current time point.

[00331] Certainly, in addition to being applied to the foregoing scenarios, the method provided in this embodiment of this application may be alternatively applied to other scenarios requiring palm image processing. A specific application scenario is not limited in the embodiments of this application.

[00332] FIG. 14 is a schematic diagram of a structure of a palm image processing apparatus according to an exemplary embodiment of this application. The apparatus may be implemented as the entire or a part of a palm identification device by using software, hardware, or a combination of thereof. The apparatus includes:

[00333] a feature extraction module 1401, configured to obtain a palm image, and perform feature extraction on the palm image, to obtain image features at a plurality of sizes;

[00334] a feature fusion module 1402, configured to fuse the image features of the plurality of sizes, to obtain a fused image feature;

[00335] a prediction module 1403, configured to determine, based on the fused image feature, a palm area identifying a palm in the palm image;

[00336] a cropping module 1404, configured to extract a palm part image from the palm image based on the palm area; and

[00337] an encryption module 1405, configured to perform image encryption on the palm part image, to obtain encrypted palm part image data, the encrypted palm part image data being configured for being transmitted to a server, and the server performing identity verification based on the encrypted palm part image data.

[00338] The feature extraction module 1401 may be a backbone network of the palm part detection model, the feature fusion module 1402 may be a neck network of the palm part detection model, the prediction module 1403 may be at least a part of a prediction network of the palm part detection model, and the cropping module 1404 may be a part of the prediction network, or may be an independent module in the palm part detection model, or may be an independent module outside the palm part detection model.

[00339] In some embodiments, the feature extraction module 1401 is the backbone network of the palm part detection model, and is configured to perform a slicing operation on the palm image, to obtain slice images of the plurality of sizes; and respectively perform feature extraction on the slice images of the plurality of sizes, to obtain the image features of the plurality of sizes.

[00340] In some embodiments, the feature extraction module 1401 is the backbone network of the palm part detection model, and is configured to determine, based on the palm image, a slice image of a maximum size in the plurality of sizes; and use the slice image of the maximum size as a first layer, and start downsampling layer by layer, to obtain the slice images of the plurality of sizes including the slice image at the first layer.

[00341] In some embodiments, the feature extraction module 1401 is the backbone network of the palm part detection model, and is configured to downsample and splice, through the backbone network, pixels in the palm image of the plurality of sizes, to obtain the slice images of the plurality of sizes. There are two adjacent pixels in slice images at different sizes, and a quantity of pixels between the two pixels at which sampling occurs in the palm image is different.

[00342] In some embodiments, the feature fusion module 1402 is the neck network of the palm part detection model, and is configured to perform, through the neck network, feature fusion on the image features of the plurality of sizes, to obtain the fused image feature.

[00343] In some embodiments, the feature fusion module 1402 is the neck network of the palm part detection model, and is configured to perform feature fusion based on the image

features of the plurality of sizes, to obtain first intermediate features of the plurality of sizes; perform feature fusion of the plurality of sizes based on the first intermediate features of the plurality of sizes, to obtain second intermediate features of the plurality of sizes; and perform feature fusion on the second intermediate features of the plurality of sizes, to obtain the fused image feature.

[00344] In some embodiments, the plurality of sizes are N sizes, N is a positive integer greater than 1, and the feature fusion module 1402 is the neck network of the palm part detection model, and is configured to determine, based on an image feature ranking first in a positive order of size, a first intermediate feature ranking first in the positive order of size; upsample layer by layer starting from the first intermediate feature ranking first in the positive order of size, and fuse a result of upsampling a first intermediate feature ranking n^{th} in the positive order of size with an image feature ranking $(n+1)^{\text{th}}$ in the positive order of size, to obtain a first intermediate feature ranking $(n+1)^{\text{th}}$ in the positive order of size, where n is a positive integer from 1 to $N-1$; and the positive order of size is in ascending order of size; and determine, when obtaining a first intermediate feature at an N^{th} layer, the first intermediate features of the plurality of sizes from a first layer to the N^{th} layer.

[00345] In some embodiments, the plurality of sizes are N sizes, N is a positive integer greater than 1, and the feature fusion module 1402 is the neck network of the palm part detection model, and is configured to determine, based on a first intermediate feature ranking first in a reverse order of size, a second intermediate feature ranking first in the reverse order of size; downsample layer by layer starting from the second intermediate feature ranking first in the reverse order of size, and fuse a result of downsampling a second intermediate feature ranking m^{th} in the reverse order of size with a first intermediate feature ranking $(m+1)^{\text{th}}$ in the reverse order of size, to obtain a second intermediate feature ranking $(n+1)^{\text{th}}$ in the reverse order of size, where m is an integer from 1 to $N-1$; and the reverse order of size is in descending order of size; and determine, when obtaining the first intermediate feature at the N^{th} layer, the first intermediate features of the plurality of sizes from the first layer to the N^{th} layer.

[00346] In some embodiments, the prediction module 1403 is at least a part of the prediction network of the palm part detection model, and is configured to perform grid division on the fused image feature, to obtain a plurality of grid features of the fused image feature; and perform prediction based on the plurality of grid features, to obtain the palm area.

[00347] In some embodiments, the palm identification device is equipped with an infrared camera, and the obtaining module 1406 is configured to obtain an infrared image collected of a palm the same as the palm image, where the infrared image is an image obtained by the infrared camera imaging, based on infrared light, the palm.

[00348] The area identification module 1407 is configured to identify a palm area from the infrared image.

[00349] In some embodiments, the cropping module 1404 is configured to determine an overlapping area between the palm area in the palm image and the palm area in the infrared image; and extract the palm part image from the palm image based on the overlapping area.

[00350] In some embodiments, the area identification module 1407 is configured to detect a finger gap point in the infrared image; and determine the palm area in the infrared image based on the finger gap point.

[00351] In some embodiments, the area identification module 1407 is configured to obtain at least three finger gap points in the infrared image; sequentially connect the finger gap points, to obtain a finger gap point connecting line; and extract the infrared image based on the finger gap point connecting line to obtain the palm area.

[00352] In some embodiments, the cropping module 1404 is configured to obtain a coordinate parameter of the palm area in the palm image, and obtain a coordinate parameter of the palm area in the infrared image; and determine the overlapping area between the palm area and the palm area based on the coordinate parameter of the palm area and the coordinate parameter of the palm area.

[00353] FIG. 15 is a schematic diagram of a structure of an apparatus for training a palm part detection model according to an exemplary embodiment of this application. The apparatus may be implemented as all or a part of a palm identification device by using software, hardware, or a combination of the two. The apparatus includes a palm part detection model, an obtaining module 1501, a calculation module 1506, and an update module 1507. The palm part detection model includes a feature extraction module 1502, a feature fusion module 1503, and a prediction module 1504, and may further include a cropping module 1505. The feature extraction module 1502 may be a backbone network of the palm part detection model, the feature fusion module 1503 may be a neck network of the palm part detection model, the prediction module 1504 may be at least a part of a prediction

network, and the cropping module 1505 may be a part of the prediction network, or may be an independent module in the palm part detection model, or may be an independent module outside the palm part detection model.

[00354] The obtaining module 1501 is configured to obtain a sample palm image and a sample palm part image;

[00355] The feature extraction module 1502 is configured to perform feature extraction on the sample palm image, to obtain sample image features at a plurality of sizes;

[00356] The feature fusion module 1503 is configured to fuse the sample image features of the plurality of sizes, to obtain a sample fused image feature;

[00357] The prediction module 1504 is configured to determine, based on the sample fused image feature, a sample prediction palm area identifying a palm in the sample palm image;

[00358] The cropping module 1505 is configured to extract a sample prediction palm part image from the sample palm image based on the sample prediction palm area;

[00359] The calculation module 1506 is configured to calculate a loss function value according to the sample palm part image and the sample prediction palm part image.

[00360] The update module 1507 is configured to update a model parameter of the palm part detection model based on the loss function value.

[00361] In some embodiments, the feature extraction module 1502 is the backbone network of the palm part detection model, and is configured to perform a slicing operation on the sample palm image, to obtain slice images of the plurality of sizes; and respectively perform feature extraction on the slice images of the plurality of sizes, to obtain the sample image features of the plurality of sizes.

[00362] In some embodiments, the feature extraction module 1502 is the backbone network of the palm part detection model, and is configured to determine, based on the sample palm image, a slice image of a maximum size in the plurality of sizes; and use the slice image of the maximum size as a first layer, and start downsampling layer by layer, to obtain the slice images of the plurality of sizes including the slice image at the first layer.

[00363] In some embodiments, the feature extraction module 1502 is the backbone network of the palm part detection model, and is configured to sample and splice pixels in the sample palm image of the plurality of sizes, to obtain the slice images of the plurality of

sizes. There are two adjacent pixels in slice images of different sizes, and a quantity of pixels between the two pixels at which sampling occurs in the sample palm image is different.

[00364] In some embodiments, the feature fusion module 1503 is the neck network of the palm part detection model, and is configured to perform feature fusion on the sample image features of the plurality of sizes, to obtain the sample fused image feature.

[00365] In some embodiments, the feature fusion module 1503 is the neck network of the palm part detection model, and is configured to input the sample image features of the plurality of sizes into the neck network; perform feature fusion based on the sample image features of the plurality of sizes, to obtain first intermediate features of the plurality of sizes; perform feature fusion of the plurality of sizes based on the first intermediate features of the plurality of sizes, to obtain second intermediate features of the plurality of sizes; and perform feature fusion on the second intermediate features of the plurality of sizes, to obtain the sample fused image feature.

[00366] In some embodiments, the plurality of sizes are N sizes, N is a positive integer greater than 1, and the feature fusion module 1503 is the neck network of the palm part detection model, and is configured to determine, based on a sample image feature ranking first in a positive order of size, a first intermediate feature ranking first in the positive order of size; upsample layer by layer starting from the first intermediate feature ranking first in the positive order of size, and fuse a result of upsampling a first intermediate feature ranking n^{th} in the positive order of size with a sample image feature ranking $(n+1)^{\text{th}}$ in the positive order of size, to obtain a first intermediate feature ranking $(n+1)^{\text{th}}$ in the positive order of size, where n is a positive integer from 1 to $N-1$; and the positive order of size is in ascending order of size; and determine, when obtaining a first intermediate feature at an N^{th} layer, the first intermediate features of the plurality of sizes from a first layer to the N^{th} layer.

[00367] In some embodiments, the plurality of sizes are N sizes, N is a positive integer greater than 1, and the feature fusion module 1503 is configured to determine, based on a first intermediate feature ranking first in a reverse order of size, a second intermediate feature ranking first in the reverse order of size; downsample layer by layer starting from the second intermediate feature ranking first in the reverse order of size, and fuse a result of downsampling a second intermediate feature ranking m^{th} in the reverse order of size with a first intermediate feature ranking $(m+1)^{\text{th}}$ in the reverse order of size, to obtain a second intermediate feature ranking $(n+1)^{\text{th}}$ in the reverse order of size, where m is an integer from 1

to N-1; and the reverse order of size is in descending order of size; and determine, when obtaining the first intermediate feature at the Nth layer, the first intermediate features of the plurality of sizes from the first layer to the Nth layer.

[00368] In some embodiments, the prediction module 1504 is at least a part of the prediction network of the palm part detection model, and is configured to perform grid division on the sample fused image feature, to obtain a plurality of grid features of the sample fused image feature; and perform prediction on the plurality of grid features, to obtain the sample prediction palm area.

[00369] In some embodiments, the palm identification device is equipped with an infrared camera, and the obtaining module 1501 is configured to obtain an infrared image collected of a palm the same as the sample palm image, where the infrared image is an image obtained by the infrared camera imaging, based on infrared light, the palm.

[00370] The area identification module 1508 is configured to identify a palm area from the infrared image.

[00371] The cropping module 1505 is configured to determine an overlapping area between the sample prediction palm area in the sample palm image and the palm area in the infrared image; and extract a palm part image from the sample palm image based on the overlapping area.

[00372] In some embodiments, the area identification module 1508 is configured to detect a finger gap point in the infrared image; and determine the palm area in the infrared image based on the finger gap point.

[00373] In some embodiments, the cropping module 1505 is configured to obtain a coordinate parameter of the sample prediction palm area in the sample palm image, and obtain a coordinate parameter of the palm area in the infrared image; and determine an overlapping area between the sample prediction palm area and the palm area based on the coordinate parameter of the sample prediction palm area and the coordinate parameter of the palm area.

[00374] In some embodiments, the area identification module 1508 is configured to perform detection on the finger gap point in the infrared image, and determine the palm area in the infrared image based on the finger gap point.

[00375] In some embodiments, the area identification module 1508 is configured to obtain at least three finger gap points in the infrared image; sequentially connect the finger gap points, to obtain a finger gap point connecting line; and extract the infrared image based on the finger gap point connecting line to obtain the palm area.

[00376] FIG. 16 is a structural block diagram of a computer device 1600 according to an exemplary embodiment of this application. The computer device can be implemented as the server in the foregoing solution of this application. The image computer device 1600 includes a central processing unit (CPU) 1601, a system memory 1604 including a random access memory (RAM) 1602 and a read-only memory (ROM) 1603, and a system bus 1605 connecting the system memory 1604 to the CPU 1601. The image computer device 1600 further includes a large-capacity storage device 1606 configured to store an operating system 1609, an application 1610 and another program module 1611.

[00377] The mass storage device 1606 is connected to the CPU 1601 by using a mass storage controller (not shown) connected to the system bus 1605. The large-capacity storage device 1606 and a computer-readable medium associated with the large-capacity storage device provide non-volatile storage to the image computer device 1600. In other words, the mass storage device 1606 may include a computer-readable medium (not shown) such as a hard disk or a compact disc read-only memory (CD-ROM) drive.

[00378] In general, the computer-readable medium may include a computer storage medium and a communication medium. The computer storage medium includes volatile and non-volatile media, and removable and non-removable media implemented by using any method or technology and configured to store information such as a computer-readable instruction, a data structure, a program module, or other data. The computer storage medium includes a RAM, an erasable programmable ROM (EPROM), an electrically-erasable programmable ROM (EEPROM), a flash memory or another solid-state memory technology, a CD-ROM, a digital versatile disc (DVD) or another optical memory, a magnetic cassette, a magnetic tape, a magnetic disk memory, or another magnetic storage device. Certainly, a person skilled in the art may know that the computer storage medium is not limited to the foregoing types. The system memory 1604 and the large-capacity storage device 1606 may be collectively referred to as a memory.

[00379] According to the embodiments of the present disclosure, the image computer device 1600 may be further connected, through a network such as the internet, to a remote

computer on the network and run. That is, the image computer device 1600 may be connected to a network 1608 by using a network interface unit 1607 connected to the system bus 1605, or may be connected to another type of network or a remote computer system (not shown) by using a network interface unit 1607.

[00380] The memory further includes at least one piece of computer-readable instruction. The at least one piece of computer-readable instruction is stored in the memory. The central processing unit 1601 executes the at least one piece of computer-readable instruction to implement all or some of the operations in the palm image processing method shown in the foregoing embodiments or the method for training a palm part detection model.

[00381] The embodiments of this application further provide a computer device, the computer device includes a processor and a memory, the memory stores at least one program, and the at least one program is loaded and executed by the processor to implement the palm image processing method provided in the foregoing method embodiments or the method for training a palm part detection model.

[00382] The embodiments of this application further provide a computer-readable storage medium, the storage medium having at least one computer-readable instruction stored therein, and the at least one computer-readable instruction being loaded and executed by a processor to implement the palm image processing method provided in the foregoing method embodiments or the method for training a palm part detection model.

[00383] The embodiments of this application further provide a computer program product, the computer program product including computer-readable instructions, and the computer-readable instructions being stored in a computer-readable storage medium; and a processor of a computer device reading the computer-readable instructions from the computer-readable storage medium and executing the computer-readable instructions, to enable the computer device to implement the palm image processing method provided in the foregoing method embodiments or the method for training a palm part detection model.

[00384] In a specific implementation of this application, the related data, a palm image, historical data, and a portrait are user data processing-related data that is related to a user identity or user characteristics. When the embodiments of this application are applied to specific products or technologies, user permission or consent needs to be obtained, and the collection, use, and processing of related data comply with related laws, regulations, and standards of related countries and regions.

[00385] "Plurality of" mentioned in the specification means two or more. "And/or" describes an association relationship for describing associated objects and represents that three relationships may exist. For example, A and/or B may represent the following three cases: only A exists, both A and B exist, and only B exists. The character "/" in this specification generally indicates an "or" relationship between the associated objects.

[00386] A person of ordinary skill in the art may understand that all or some of the operations of the embodiments may be implemented by hardware or a program instructing related hardware. The program may be stored in a computer-readable storage medium. The storage medium may be a read-only memory, a magnetic disk, an optical disc, or the like.

[00387] Technical features of the foregoing embodiments may be combined in different manners to form other embodiments. For concise description, not all possible combinations of the technical features in the embodiment are described. However, provided that combinations of the technical features do not conflict with each other, the combinations of the technical features are considered as falling within the scope recorded in this specification.

[00388] The term "module" in this application refers to a computer program or part of the computer program that has a predefined function and works together with other related parts to achieve a predefined goal and may be all or partially implemented by using software, hardware (e.g., processing circuitry and/or memory configured to perform the predefined functions), or a combination thereof. Each module can be implemented using one or more processors (or processors and memory). Likewise, a processor (or processors and memory) can be used to implement one or more modules. Moreover, each module can be part of an overall module that includes the functionalities of the module. The foregoing embodiments only describe several implementations of this application, which are described specifically and in detail, but cannot be construed as a limitation to the patent scope of the present invention. For a person of ordinary skill in the art, several transformations and improvements can be made without departing from the idea of this application. These transformations and improvements belong to the protection scope of this application. Therefore, the protection scope of the patent of this application shall be subject to the appended claims.

CLAIMS

What is claimed is:

1. A palm image processing method, performed by a palm identification device, comprising:

obtaining a palm image, and extracting image features of a plurality of sizes from the palm image;

fusing the image features of the plurality of sizes into a fused image feature;

determining, based on the fused image feature, a palm area defining a palm in the palm image;

extracting a palm part image from the palm image based on the palm area; and

encrypting the palm part image into encrypted palm part image data, wherein the encrypted palm part image data is for being transmitted to a server which performs identity identification using the encrypted palm part image data.

2. The method according to claim 1, wherein the palm area is obtained by prediction by a palm part detection model, and the palm part detection model comprises a backbone network; and

the performing feature extraction on the palm image, to obtain image features at a plurality of sizes comprises:

slicing the palm image through the backbone network into slice images of the plurality of sizes; and

performing feature extraction on the slice images of the plurality of sizes respectively through the backbone network, to obtain the image features of the plurality of sizes.

3. The method according to claim 2, wherein the performing a slicing operation on the palm image through the backbone network, to obtain slice images of the plurality of sizes comprises:

determining, through the backbone network and based on the palm image, a slice image of a maximum size in the plurality of sizes; and

using, through the backbone network, the slice image of the maximum size as a first layer, and starting downsampling layer by layer, to obtain the slice images of the plurality of

sizes including the slice image at the first layer.

4. The method according to claim 2, wherein the performing a slicing operation on the palm image through the backbone network, to obtain slice images of the plurality of sizes comprises:

downsampling and splicing, through the backbone network, pixels in the palm image of the plurality of sizes, to obtain the slice images of the plurality of sizes, wherein

there are two adjacent pixels in slice images of different sizes, and a quantity of pixels between the two pixels at which sampling occurs in the palm image is different.

5. The method according to any one of claims 1 to 3, wherein the palm area is obtained by prediction by the palm part detection model, and the palm part detection model comprises a neck network; and

the fusing the image features of the plurality of sizes into a fused image feature comprises:

inputting the image features of the plurality of sizes into the neck network for feature fusion, to obtain the fused image feature.

6. The method according to claim 5, wherein the inputting the image features of the plurality of sizes into the neck network for feature fusion, to obtain the fused image feature comprises:

inputting the image features of the plurality of sizes into the neck network;

performing, through the neck network, feature fusion based on the image features of the plurality of sizes, to obtain first intermediate features of the plurality of sizes;

performing, through the neck network, feature fusion of the plurality of sizes based on the first intermediate features of the plurality of sizes, to obtain second intermediate features of the plurality of sizes; and

performing, through the neck network, feature fusion on the second intermediate features of the plurality of sizes, to obtain the fused image feature.

7. The method according to claim 6, wherein the plurality of sizes are N sizes, N is a positive integer greater than 1, and the extracting, based on the image features of the plurality of sizes, first intermediate features of the plurality of sizes comprises:

determining, based on an image feature ranking first in a positive order of size, a first

intermediate feature ranking first in the positive order of size;

upsampling layer by layer starting from the first intermediate feature ranking first in the positive order of size, and fusing a result of upsampling a first intermediate feature ranking n^{th} in the positive order of size with an image feature ranking $(n+1)^{\text{th}}$ in the positive order of size, to obtain a first intermediate feature ranking $(n+1)^{\text{th}}$ in the positive order of size, wherein n is a positive integer from 1 to $N-1$; and the positive order of size is in ascending order of size; and

determining, when obtaining a first intermediate feature at an N^{th} layer, the first intermediate features of the plurality of sizes from a first layer to the N^{th} layer.

8. The method according to claim 6 or 7, wherein the plurality of sizes comprise N sizes, N is a positive integer greater than 1, and the performing feature fusion of the plurality of sizes based on the first intermediate features of the plurality of sizes, to obtain second intermediate features of the plurality of sizes comprises:

determining, based on a first intermediate feature ranking first in a reverse order of size, a second intermediate feature ranking first in the reverse order of size;

downsampling layer by layer starting from the second intermediate feature ranking first in the reverse order of size, and fusing a result of downsampling a second intermediate feature ranking m^{th} in the reverse order of size with a first intermediate feature ranking $(m+1)^{\text{th}}$ in the reverse order of size, to obtain a second intermediate feature ranking $(n+1)^{\text{th}}$ in the reverse order of size, wherein m is an integer from 1 to $N-1$; and the reverse order of size is in descending order of size; and

determining, when obtaining the first intermediate feature at the N^{th} layer, the first intermediate features of the plurality of sizes from the first layer to the N^{th} layer.

9. The method according to any one of claims 1 to 8, wherein the palm area is determined by the palm part detection model, and the palm part detection model comprises a prediction network; and

the determining, based on the fused image feature, a palm area defining a palm in the palm image comprises:

performing grid division on the fused image feature, to obtain a plurality of grid features of the fused image feature; and

inputting the plurality of grid features into the prediction network, and performing

prediction on the plurality of grid features, to obtain the palm area.

10. The method according to any one of claims 1 to 9, wherein the palm identification device is equipped with an infrared camera, and the method further comprises:

obtaining an infrared image collected of a same palm as the palm in the palm image, wherein the infrared image is an image obtained by the infrared camera imaging, based on infrared light, the palm;

identifying a palm area from the infrared image; and

the extracting a palm part image from the palm image based on the palm area comprises:

determining an overlapping area between the palm area in the palm image and the palm area in the infrared image; and

extracting the palm part image from the palm image based on the overlapping area.

11. The method according to claim 10, the identifying a palm area from the infrared image comprises:

detecting a finger gap point in the infrared image; and

determining the palm area in the infrared image based on the finger gap point.

12. The method according to claim 10 or 11, wherein the determining an overlapping area between the palm area in the palm image and the palm area in the infrared image comprises:

obtaining a coordinate parameter of the palm area in the palm image, and obtaining a coordinate parameter of the palm area in the infrared image; and

determining the overlapping area between the palm area and the palm area based on the coordinate parameter of the palm area and the coordinate parameter of the palm area.

13. A method for training a palm part detection model, performed by a computer device, the method comprising:

obtaining a sample palm image and a sample palm part image;

extracting sample image features of a plurality of sizes from the sample palm image;

fusing the sample image features of the plurality of sizes into obtain a sample fused image feature;

determining, based on the sample fused image feature, a sample prediction palm area

identifying a palm in the sample palm image;

extracting a sample prediction palm part image from the sample palm image based on the sample prediction palm area;

calculating a loss function value according to the sample palm part image and the sample prediction palm part image; and

updating a model parameter of the palm part detection model based on the loss function value.

14. A palm image processing apparatus, comprising:

a feature extraction module, configured to obtain a palm image, and extract image features of a plurality of sizes from the palm image;

a feature fusion module, configured to fuse the image features of the plurality of sizes into a fused image feature;

a prediction module, configured to determine, based on the fused image feature, a palm area defining a palm in the palm image;

a cropping module, configured to extract a palm part image from the palm image based on the palm area; and

an encryption module, configured to encrypt the palm part image into encrypted palm part image data, wherein the encrypted palm part image data is for being transmitted to a server which performing identity verification using the encrypted palm part image data.

15. An apparatus for training a palm part detection model, comprising a palm part detection model, and further comprising an obtaining module, a cropping module, a calculation module, and an update module, the palm part detection model comprising a feature extraction module, a feature fusion module, and a prediction module;

the obtaining module being configured to obtain a sample palm image and a sample palm part image;

the feature extraction module being configured to extract sample image features of a plurality of sizes from the sample palm image;

the feature fusion module being configured to fuse the sample image features of the plurality of sizes into a sample fused image feature;

the prediction module being configured to determine, based on the sample fused image feature, a sample prediction palm area defining a palm in the sample palm image;

the cropping module being configured to extract a sample prediction palm part image from the sample palm image based on the sample prediction palm area;

the calculation module being configured to calculate a loss function value according to the sample palm part image and the sample prediction palm part image; and

the update module being configured to update a model parameter of the palm part detection model based on the loss function value.

16. A computer device, comprising a processor and a memory, the memory storing at least one computer-readable instruction, the at least one computer-readable instruction being loaded and executed by the processor to implement the method according to any one of claims 1 to 13.

17. A computer storage medium, storing at least one computer-readable instruction, the at least one computer-readable instruction being loaded and executed by a processor to implement the method according to any one of claims 1 to 13.

18. A computer program product, comprising computer-readable instructions, the computer-readable instructions being loaded and executed by a processor to implement the method according to any one of claims 1 to 13.