CLAIMS

1. An image processing method, performed by a computer device and comprising:

acquiring an identity feature of a source image and an initial attribute feature of at least one scale of a target image in response to a received face change request;

the face change request being used for requesting to replace a target face in the target image with a source face in the source image, the identity feature representing an object to which the source face belongs, and the initial attribute feature representing a three-dimensional attribute of the target face;

inputting the identity feature and the initial attribute feature of the at least one scale into a face change model;

iteratively performing feature fusion on the identity feature and the initial attribute feature of the at least one scale by using the face change model to obtain a post-fusion feature; and

generating a target face change image based on the post-fusion feature by using the face change model, and outputting the target face change image;

a face in the target face change image being fused with an identity feature of the source face and a target attribute feature of the target face;

wherein the face change model comprises at least one convolutional layer, each of which corresponds to one scale; and the iteratively performing feature fusion on the identity feature and the initial attribute feature of the at least one scale by using the face change model to obtain a post-fusion feature comprises:

separately performing the following processing on the identity feature and an initial attribute feature of a corresponding scale by using each convolutional layer of the face change model:

acquiring a first feature map outputted by a previous convolutional layer of a current convolutional layer;

generating a second feature map based on the identity feature and the first feature map, and determining a control mask of the target image at a corresponding scale based on the second feature map and the initial attribute feature; the control mask being used for representing a pixel that carries a feature other than the identity feature of the target face;

screening the initial attribute feature of the at least one scale based on the control mask to

obtain the target attribute feature; the target attribute feature being a feature other than an identity feature of the target face; and

generating a third feature map based on the target attribute feature and the second feature map, the third feature map being a first feature map of a next convolutional layer of the current convolutional layer; and

determining, as the post-fusion feature, a third feature map outputted by the last convolutional layer in the at least one convolutional layer.

2. The method according to claim 1, wherein the determining a control mask of the target image at a corresponding scale based on the second feature map and the initial attribute feature comprises:

performing feature concatenation on the second feature map and the initial attribute feature to obtain a concatenated feature map; and

mapping the concatenated feature map to the control mask based on a preconfigured mapping convolution kernel and an activation function.

3. The method according to claim 1, wherein both the initial attribute feature and a quantity of convolutional layers are a target quantity, the convolutional layers of the target quantity are connected in series, different initial attribute features are corresponding to different scales, each convolutional layer is corresponding to an initial attribute feature of one scale, and the target quantity is greater than or equal to two; and

the acquiring a first feature map outputted by a previous convolutional layer of a current convolutional layer comprises:

acquiring an initial feature map in a case that the current convolutional layer is the first convolutional layer in the convolutional layers of the target quantity, and using the initial feature map as a first feature map inputted to the current convolutional layer.

4. The method according to claim 1, wherein the generating a second feature map based on the identity feature and the first feature map comprises:

performing affine transform on the identity feature to obtain a first control vector;

mapping a first convolution kernel of the current convolutional layer to a second convolution kernel based on the first control vector; and

performing a convolution operation on the first feature map based on the second convolution kernel, to generate the second feature map.

5. The method according to claim 1, wherein before the inputting the identity feature and the initial attribute feature of the at least one scale into a face change model, the method further comprises:

acquiring a sample identity feature of a sample source image in a sample image pair and a sample initial attribute feature of at least one scale of a sample target image in the sample image pair;

iteratively performing feature fusion on the sample identity feature and the sample initial attribute feature of the at least one scale by using a generator of an initial face change model to obtain a sample post-fusion feature;

generating a sample generation image based on the sample post-fusion feature by using the generator of the initial face change model;

discriminating, by using a discriminator of the initial face change model, the sample generation image and the sample source image to obtain a discrimination result; and

determining a loss of the initial face change model based on the discrimination result, and training the initial face change model based on the loss to obtain the face change model.

6. The method according to claim 5, wherein the discrimination result comprises a first discrimination result for the sample source image and a second discrimination result for the sample generation image; and the determining a loss of the initial face change model based on the discrimination result comprises:

acquiring a sample mask of at least one scale of the sample target image, and determining a first loss value based on the sample mask of the at least one scale;

determining a second loss value based on the first discrimination result and the second discrimination result;

obtaining a total training loss based on the first loss value and the second loss value; and

training the initial face change model based on the total training loss until a target condition is met and training stops, to obtain the face change model.

7. The method according to claim 6, wherein the sample source image and the sample target image are corresponding to a same object; and

the obtaining a total training loss based on the first loss value and the second loss value comprises:

acquiring a third loss value based on the sample generation image and the sample target image;

and

obtaining the total training loss based on the third loss value, the first loss value, and the second loss value.

8. The method according to claim 6, wherein the discriminator comprises at least one convolutional layer; and the obtaining a total training loss based on the first loss value and the second loss value comprises:

determining a first similarity between a non-face region of a first discrimination feature map and a non-face region of a second discrimination feature map, the first discrimination feature map being a feature map of a sample target image outputted by a first part of the convolutional layer, and the second discrimination feature map being a feature map of a sample generation image outputted by the first part of the convolutional layer;

determining a second similarity between a third discrimination feature map and a fourth discrimination feature map, the third discrimination feature map being a feature map of a sample target image outputted by a second part of the convolutional layer, and the fourth discrimination feature map being a feature map of a sample generation image outputted by the second part of the convolutional layer;

determining a fourth loss value based on the first similarity and the second similarity; and

obtaining the total training loss based on the first loss value, the second loss value, and the fourth loss value.

9. The method according to claim 6, wherein the obtaining a total training loss based on the first loss value and the second loss value comprises:

extracting a first identity feature of the sample source image, a second identity feature of the sample target image, and a third identity feature of the sample generation image;

determining a first identity similarity between the sample source image and the sample generation image based on the first identity feature and the third identity feature;

determining a first identity distance between the sample generation image and the sample target image based on the second identity feature and the third identity feature;

determining a second identity distance between the sample source image and the sample target image based on the first identity feature and the second identity feature;

determining a distance difference based on the first identity distance and the second identity distance;

determining a fifth loss value based on the first identity similarity and the distance difference; and

obtaining the total training loss based on the first loss value, the second loss value, and the fifth loss value.

10. An image processing apparatus, comprising:

a feature acquiring module, configured to acquire an identity feature of a source image and an initial attribute feature of at least one scale of a target image in response to a received face change request;

the face change request being used for requesting to replace a target face in the target image with a source face in the source image, the identity feature representing an object to which the source face belongs, and the initial attribute feature representing a three-dimensional attribute of the target face; and

a face change module, configured to input the identity feature and the initial attribute feature of the at least one scale into a face change model;

iteratively perform feature fusion on the identity feature and the initial attribute feature of the at least one scale by using the face change model to obtain a post-fusion feature; and

generate a target face change image based on the post-fusion feature by using the face change model, and output the target face change image;

a face in the target face change image being fused with an identity feature of the source face and a target attribute feature of the target face;

wherein the face change model comprises at least one convolutional layer, each of which corresponds to one scale; and the face change module is further configured to:

separately perform the following processing on the identity feature and an initial attribute feature of a corresponding scale by using each convolutional layer of the face change model:

acquire a first feature map outputted by a previous convolutional layer of a current convolutional layer;

generate a second feature map based on the identity feature and the first feature map, and determine a control mask of the target image at a corresponding scale based on the second feature map and the initial attribute feature; the control mask being used for representing a pixel that carries a feature other than the identity feature of the target face;

screen the initial attribute feature of the at least one scale based on the control mask to obtain

the target attribute feature; the target attribute feature being a feature other than an identity feature of the target face;

generate a third feature map based on the target attribute feature and the second feature map, the third feature map being a first feature map of a next convolutional layer of the current convolutional layer; and

determine, as the post-fusion feature, a third feature map outputted by the last convolutional layer in the at least one convolutional layer.

11. A computer device, comprising a memory and a processor,

the memory being configured to store a computer program; and

the processor being configured to execute the computer program stored in the memory, performing the image processing method according to any one of claims 1 to 9.

12. A computer-readable storage medium, storing a computer program, the computer program, when executed by a processor, implementing the image processing method according to any one of claims 1 to 9.

13. A computer program product, comprising a computer program, the computer program, when executed by a processor, performing the image processing method according to any one of claims 1 to 9.