CLAIMS

1. An image conversion method, comprising:

determining a first domain corresponding to a first image using an image conversion network;

determining a feature mapping relationship between the first image and a preset image in each layer of the image conversion network based on the first domain and a second domain corresponding to the preset image; and

converting a picture of the first image from the first domain to the second domain based on the feature mapping relationship to obtain a second image~~.~~;

~~2. The method of claim 1,~~ wherein the determining a feature mapping relationship between the first image and a preset image in each layer of the image conversion network based on the first domain and a second domain corresponding to the preset image comprises:

determining a feature class shared by the first image and the preset image in each layer of the image conversion network based on the first domain and the second domain; and

executing the following steps taking the each layer as a target layer:

determining a first output feature of the first image in the target layer and a second output feature of the preset image in the target layer based on the feature class shared in the target layer respectively; and

determining the feature mapping relationship between the first image and the preset image in the target layer based on the first output feature and the second output feature~~.~~;

~~3. The method of claim 2,~~ wherein the determining a feature class shared by the first image and the preset image in each layer of the image conversion network based on the first domain

and the second domain comprises:

determining an abstract content feature class shared by the first image and the preset image based on the first domain and the second domain;

determining a concrete content feature class shared by the first image and the preset image in each layer; and

obtaining the feature class in the each layer based on the abstract content feature class and the concrete content feature class shared in the each layer.

~~4~~2. The method of claim ~~2 or 3~~1, wherein the determining a first output feature of the first image in the target layer based on a feature class shared in the target layer comprises:

determining an encoder feature of the feature class shared in the target layer in the first image using a content encoder of the image conversion network;

inputting the encoder feature in the target layer to a generator of the image conversion network to determine a generator feature in the target layer; and

fusing the generator feature in the target layer and the encoder feature in the target layer to obtain the first output feature of the first image in the target layer.

~~5~~3. The method of claim ~~4~~2, wherein the determining an encoder feature of the feature class shared in the target layer in the first image using a content encoder of the image conversion network comprises:

determining an abstract content feature belonging to an abstract content feature class in the feature class shared in the target layer and a concrete content feature belonging to a concrete content feature class in the feature class shared in the target layer in the first image using the content encoder;

determining an initial hidden state based on the abstract content feature; and

updating the concrete content feature in the target layer based on the initial hidden state to obtain the encoder feature in the target layer.

~~6~~4. The method of claim ~~4 or 5~~2 or 3, wherein the inputting the encoder feature in the target layer to a generator of the image conversion network to determine a generator feature in the target layer comprises:

determining mask information of the first image in the target layer based on a hidden state in the target layer and the encoder feature in the target layer; and

inputting the concrete content feature in the target layer and the mask information in the target layer to the generator to generate the generator feature in the target layer; and

wherein the fusing the generator feature in the target layer and the encoder feature in the target layer to obtain the first output feature of the first image in the target layer comprises:

fusing the encoder feature in the target layer and the generator feature in the target layer based on the mask information in the target layer to obtain the first output feature of the first image in the target layer.

~~7~~5. The method of claim ~~6~~4, wherein the determining mask information of the first image in the target layer based on a hidden state in the target layer and the encoder feature in the target layer comprises:

in a case where the target layer is a first layer, determining the abstract content feature as a hidden state in the first layer; or in a case where the target layer is an i-th layer, determining a hidden state in the i-th layer based on a hidden state in an (i-1)th layer and a concrete content feature in the i-th layer, wherein i is an integer greater than 1; and

determining mask information of the first image in a (j+1)th layer based on a hidden state in a j-th layer and a concrete content feature in the (j+1)th layer, wherein j is an integer greater than or equal to 1.

~~8~~6. The method of claim ~~6 or 7~~4 or 5, wherein the inputting the concrete content feature in the target layer and the mask information in the target layer to the generator to generate the generator feature in the target layer comprises:

performing regularization processing on the mask information in the target layer to obtain sparse mask information in the target layer;

acquiring a target content feature matched with the sparse mask information in the target layer from the concrete content feature in the target layer; and

inputting the target content feature in the target layer to the generator to generate the generator feature in the target layer.

~~9~~7. The method of any one of claims 4-6~~6-8~~, wherein the fusing the encoder feature in the target layer and the generator feature in the target layer based on the mask information in the target layer to obtain the first output feature of the first image in the target layer comprises:

adjusting the encoder feature in the target layer based on the mask information in the target layer to obtain a first adjusted feature in the target layer;

adjusting the generator feature in the target layer based on adversarial information of the mask information in the target layer to obtain a second adjusted feature in the target layer; and

fusing the first adjusted feature in the target layer and the second adjusted feature in the target layer to obtain the first output feature in the target layer.

~~10~~8. The method of any one of claims 1-~~7~~9, wherein the converting a picture of the first image from the first domain to the second domain based on the feature mapping relationship to obtain a second image further comprises:

determining a style feature of the preset image in each layer in the image conversion network using a style encoder in the image conversion network;

determining the style feature and corresponding feature mapping relationship in any layer in the image conversion network as a feature group; and

inputting a feature group in the each layer in the image conversion network to the generator of the image conversion network to obtain the second image.

~~11~~9. The method of any one of claims 1-~~8~~10, wherein the image conversion network is obtained by training a network to be trained based on a first sample image of which a picture belongs to a first sample domain and a second sample image of which a picture belongs to a second sample domain, and a training process of the network to be trained comprises:

acquiring the first sample image and second sample image which are correlated with a picture;

determining, using a trained content encoder in the network to be trained, a first sample feature of the first sample image in each layer of the trained content encoder and a second sample feature of the second sample image in the each layer;

determining a sample feature mapping relationship between the first sample image and the second sample image in each layer of the trained content encoder based on the first sample feature in the each layer and the second sample feature in the each layer;

inputting the sample feature mapping relationship in the each layer and a sample style feature of the second sample image to a first generator of the network to be trained to obtain a converted sample image after the picture is converted from the first sample domain to the second sample domain;

inputting the sample style feature of the second sample image and the second sample feature in the each layer to a second generator of the network to be trained to obtain a first reconstructed sample image after the picture of the second sample image is reconstructed; and

adjusting a network parameter of the network to be trained based on a loss of the converted sample image and a loss of the first reconstructed sample image to obtain the image conversion network.

~~12~~10. The method of claim 9~~11~~, wherein the acquiring the first sample image and second sample image which are correlated with a picture comprises:

acquiring a preset noise and a domain label set at least including the first sample domain and the second sample domain; and

inputting the preset noise and the domain label set to a trained adversarial network to generate the first sample image and second sample image which have matched semantic information.

~~13~~11. The method of claim ~~11~~ 9 or 10~~12~~, before the determining, using a trained content encoder in the network to be trained, a first sample feature of the first sample image in each layer of the trained content encoder and a second sample feature of the second sample image in the each layer, further comprising:

training a content encoder to be trained based on the first sample image and the second sample image to obtain the trained content encoder.

~~14~~12. The method of claim ~~13~~11, wherein the training a content encoder to be trained based on the first sample image and the second sample image to obtain the trained content encoder comprises:

determining a first predicted content feature of the first sample image and a second predicted content feature of the second sample image using the content encoder to be trained;

inputting the first predicted content feature, a sample style feature of the first sample image, and the first sample domain to a decoder to reconstruct the first sample image to obtain

a second reconstructed sample image;

determining a first loss between the second reconstructed sample image and the first sample image;

determining a second loss between the second predicted content feature and the first predicted content feature; and

adjusting a network parameter of the content encoder to be trained based on the first loss and the second loss to obtain the trained content encoder.

~~15~~13. The method of any one of claims 9-12~~11-14~~, wherein the adjusting a network parameter of the network to be trained based on a loss of the converted sample image and a loss of the first reconstructed sample image to obtain the image conversion network comprises:

determining the loss of the converted sample image based on a content feature of the converted sample image and a content feature of the first sample image;

determining the loss of the first reconstructed sample image based on a style feature of the first reconstructed sample image and the style feature of the second sample image;

determining an adversarial loss between the first reconstructed sample image and the second sample image; and

adjusting the network parameter of the network to be trained based on the loss of the converted sample image, the loss of the first reconstructed sample image, and the adversarial loss to obtain the image conversion network.

~~16~~14. An image conversion apparatus, comprising:

a first determination module, configured to determine a first domain corresponding to a first image using an image conversion network;

a second determination module, configured to determine a feature mapping relationship between the first image and a preset image in each layer of the image conversion network based on the first domain and a second domain corresponding to the preset image; and

a first conversion module, configured to convert a picture of the first image from the first domain to the second domain based on the feature mapping relationship to obtain a second image.;

~~17. The apparatus of claim 16,~~ wherein the second determination module comprises:

a first determination submodule, configured to determine a feature class shared by the first image and the preset image in each layer of the image conversion network based on the first domain and the second domain;

a second determination submodule, configured to, taking each layer as a target layer, determine a first output feature of the first image in the target layer and a second output feature of the preset image in the target layer based on the feature class shared in the target layer respectively; and

a third determination submodule, configured to determine the feature mapping relationship between the first image and the preset image in the target layer based on the first output feature and the second output feature.;

wherein the first determination submodule comprises:

a first determination unit, configured to determine an abstract content feature class shared by the first image and the preset image based on the first domain and the second domain;

a second determination unit, configured to determine a concrete content feature class shared by the first image and the preset image in each layer; and

a third determination unit, configured to obtain the feature class in the each layer

based on the abstract content feature class and the concrete content feature class shared in the each layer.

~~18~~15. A computer storage medium having a computer-executable instruction stored therein, wherein when executed, the computer-executable instruction is capable of implementing an image conversion method, wherein the method comprises:

determining a first domain corresponding to a first image using an image conversion network;

determining a feature mapping relationship between the first image and a preset image in each layer of the image conversion network based on the first domain and a second domain corresponding to the preset image; and

converting a picture of the first image from the first domain to the second domain based on the feature mapping relationship to obtain a second image~~.~~;

wherein the determining a feature mapping relationship between the first image and a preset image in each layer of the image conversion network based on the first domain and a second domain corresponding to the preset image comprises:

determining a feature class shared by the first image and the preset image in each layer of the image conversion network based on the first domain and the second domain; and

executing the following steps taking the each layer as a target layer:

determining a first output feature of the first image in the target layer and a second output feature of the preset image in the target layer based on the feature class shared in the target layer respectively; and

determining the feature mapping relationship between the first image and the preset image in the target layer based on the first output feature and the second output feature;

wherein the determining a feature class shared by the first image and the preset image in each

layer of the image conversion network based on the first domain and the second domain comprises:

determining an abstract content feature class shared by the first image and the preset image based on the first domain and the second domain;

determining a concrete content feature class shared by the first image and the preset image in each layer; and

obtaining the feature class in the each layer based on the abstract content feature class and the concrete content feature class shared in the each layer.

~~19~~16. An electronic device, comprising a memory and a processor, wherein a computer-executable instruction is stored in the memory, wherein when executing the computer-executable instruction in the memory, the processor is capable of implementing an image conversion method, wherein the method comprises:

determining a first domain corresponding to a first image using an image conversion network;

determining a feature mapping relationship between the first image and a preset image in each layer of the image conversion network based on the first domain and a second domain corresponding to the preset image; and

converting a picture of the first image from the first domain to the second domain based on the feature mapping relationship to obtain a second image~~.~~;

wherein the determining a feature mapping relationship between the first image and a preset image in each layer of the image conversion network based on the first domain and a second domain corresponding to the preset image comprises:

determining a feature class shared by the first image and the preset image in each layer of the image conversion network based on the first domain and the second domain; and

executing the following steps taking the each layer as a target layer:

determining a first output feature of the first image in the target layer and a second output feature of the preset image in the target layer based on the feature class shared in the target layer respectively; and

determining the feature mapping relationship between the first image and the preset image in the target layer based on the first output feature and the second output feature;

wherein the determining a feature class shared by the first image and the preset image in each layer of the image conversion network based on the first domain and the second domain comprises:

determining an abstract content feature class shared by the first image and the preset image based on the first domain and the second domain;

determining a concrete content feature class shared by the first image and the preset image in each layer; and

obtaining the feature class in the each layer based on the abstract content feature class and the concrete content feature class shared in the each layer.

~~20~~17. A computer program, comprising a computer-readable code, wherein when executed in an electronic device, a processor in the electronic device is capable of implementing an image conversion method, wherein the method comprises:

determining a first domain corresponding to a first image using an image conversion network;

determining a feature mapping relationship between the first image and a preset image in each layer of the image conversion network based on the first domain and a second domain corresponding to the preset image; and

converting a picture of the first image from the first domain to the second domain based on the feature mapping relationship to obtain a second image~~.~~;

wherein the determining a feature mapping relationship between the first image and a preset image in each layer of the image conversion network based on the first domain and a second domain corresponding to the preset image comprises:

determining a feature class shared by the first image and the preset image in each layer of the image conversion network based on the first domain and the second domain; and

executing the following steps taking the each layer as a target layer:

determining a first output feature of the first image in the target layer and a second output feature of the preset image in the target layer based on the feature class shared in the target layer respectively; and

determining the feature mapping relationship between the first image and the preset image in the target layer based on the first output feature and the second output feature;

wherein the determining a feature class shared by the first image and the preset image in each layer of the image conversion network based on the first domain and the second domain comprises:

determining an abstract content feature class shared by the first image and the preset image based on the first domain and the second domain;

determining a concrete content feature class shared by the first image and the preset image in each layer; and

obtaining the feature class in the each layer based on the abstract content feature class and the concrete content feature class shared in the each layer.