WHAT IS CLAIMED IS:

1. A feature extraction model training method, comprising:

for any obtained first sample image, obtaining a first feature vector of a first sample based on a trained convolutional neural network; processing the first sample image into a one-dimensional sequence based on a preset image processing manner, inputting the one-dimensional sequence into a feature extraction model to-be-trained, and obtaining a second feature vector of the first sample image based on the feature extraction model; and based on the second feature vector, calibrating the first feature vector to obtain a calibrated first baseline feature vector, wherein the feature extraction model is a model trained based on a framework structure of a MPCFormer (Multi-Party Computation Transformer) model and a training logic of a semantic vision transformer model SETR (Segmentation Transformer); and

training the feature extraction model based on the first baseline feature vector, a label feature vector carried by the first sample image, and a configured first loss function.

2. The method according to claim 1, after obtaining the second feature vector of the first sample image, and before calibrating the first feature vector based on the second feature vector, the method further including:

performing global average pooling on the second feature vector to obtain global feature vector information corresponding to each dimension feature of the second feature vector, and obtaining an importance weight coefficient corresponding to each dimension feature of the global feature vector information based on a probability distribution function; based on the importance weight coefficient corresponding to each dimension feature,

calibrating the second feature vector to obtain a calibrated second feature vector; and based on the calibrated second feature vector, performing a subsequent step of calibrating the first feature vector based on the second feature vector.

3. The method according to claim 2, wherein the first loss function includes the importance weight coefficient corresponding to each dimension feature.

4. The method according to claim 1, wherein calibrating the first feature vector based on the second feature vector to obtain the calibrated first baseline feature vector includes:

performing an XOR operation on each dimension feature of the second feature vector and each dimension feature of the first feature vector and obtaining the calibrated first baseline feature vector.

5. The method according to claim 4, after performing the XOR operation on each dimension feature of the second feature vector and each dimension feature of the first feature vector and before obtaining the calibrated first baseline feature vector, the method further including:

performing a normalization processing on vectors obtained after the XOR operation; and based on normalized vectors, determining the calibrated first baseline feature vector.

6. The method according to claim 1, wherein activation functions used in the feature extraction model includes a 2quad activation function.

7. The method according to claim 1, wherein the feature extraction model is a model obtained based on knowledge distillation.

8. A convolutional neural network training method, comprising:

for any obtained second sample image, inputting a second sample image into a convolutional neural network to-be-trained, and based on the convolutional neural network, obtaining a third feature vector of the second sample image; processing the second sample image into a one-dimensional sequence based on a preset image processing manner, inputting the one-dimensional sequence into a trained feature extraction model, and obtaining a fourth feature vector of the second sample image based on the feature extraction model; and based on the fourth feature vector, calibrating the third feature vector to obtain a calibrated second baseline feature vector, wherein the feature extraction model is a model trained based on a framework structure of a MPCFormer model and a training logic of a semantic vision transformer model SETR; and

training the convolutional neural network based on the second baseline feature vector, a label feature vector carried by the second sample image, and a configured second loss function.

9. The method according to claim 8, wherein based on the fourth feature vector, calibrating the third feature vector to obtain the calibrated second baseline feature vector includes:

performing an XOR operation on each dimension feature of the fourth feature vector and each dimension feature of the third feature vector and obtaining the calibrated second baseline feature vector.

10. The method according to claim 9, after performing the XOR operation on each dimension feature of the fourth feature vector and each dimension feature of the third feature

vector and before obtaining the calibrated second baseline feature vector, the method further including:

performing a normalization processing on vectors obtained after the XOR operation; and based on normalized vectors, determining the calibrated second baseline feature vector.

11. A model training method, comprising:

for any obtained third sample image, obtaining a fifth feature vector of a third sample image based on a convolutional neural network to-be-trained; processing the third sample image into a one-dimensional sequence based on a preset image processing manner, inputting the one-dimensional sequence into a feature extraction model to-be-trained, and obtaining a sixth feature vector of the third sample image based on the feature extraction model; and based on the sixth feature vector, calibrating the fifth feature vector to obtain a calibrated third baseline feature vector, wherein the feature extraction model is a model trained based on a framework structure of a MPCFormer model and a training logic of a semantic vision transformer model SETR; and

training the feature extraction model based on the third baseline feature vector, a label feature vector carried by the third sample image, and a configured first loss function; and training the convolutional neural network based on the third baseline feature vector, the label feature vector carried by the third sample image, and a configured second loss function.

12. A feature extraction method, comprising:

receiving an image to-be-processed; and

obtaining a feature vector of the image to-be-processed based on the feature extraction model trained by any method according to claims 1-7 and 11 or based on the convolutional neural network trained by any method according to claims 8-11.

13. The method according to claim 12, wherein the method further includes:

performing secure multi-party computation based on the feature vector.

14. A feature extraction model training apparatus, comprising:

a first calibrating module, configured to, for any obtained first sample image, obtain a first feature vector of a first sample based on a trained convolutional neural network; process the first sample image into a one-dimensional sequence based on a preset image processing manner, input the one-dimensional sequence into a feature extraction model to-be-trained, and obtain a second feature vector of the first sample image based on the feature extraction model; and based on the second feature vector, calibrate the first feature vector to obtain a calibrated first baseline feature vector, wherein the feature extraction model is a model trained based on a framework structure of a MPCFormer model and a training logic of a semantic vision transformer model SETR; and

a first training module, configured to train the feature extraction model based on the first baseline feature vector, a label feature vector carried by the first sample image, and a configured first loss function.

15. A convolutional neural network training apparatus, comprising:

a second calibrating module, configured to, for any obtained second sample image, input a second sample image into a convolutional neural network to-be-trained, and obtain a third feature vector of the second sample image based on the convolutional neural network; process the second sample image into a one-dimensional sequence based on a preset image processing manner, input the one-dimensional sequence into a trained feature extraction

model, and obtain a fourth feature vector of the second sample image based on the feature extraction model; and based on the fourth feature vector, calibrate the third feature vector to obtain a calibrated second baseline feature vector, wherein the feature extraction model is a model trained based on a framework structure of a MPCFormer model and a training logic of a semantic vision transformer model SETR; and

a second training module, configured to train the convolutional neural network based on the second baseline feature vector, a label feature vector carried by the second sample image, and a configured second loss function.

16. A model training apparatus, comprising:

a third calibrating module, configured to, for any obtained third sample image, obtain a fifth feature vector of a third sample image based on a convolutional neural network to-be-trained; process the third sample image into a one-dimensional sequence based on a preset image processing manner, input the one-dimensional sequence into a feature extraction model to-be-trained, and obtain a sixth feature vector of the third sample image based on the feature extraction model; and based on the sixth feature vector, calibrate the fifth feature vector to obtain a calibrated third baseline feature vector, wherein the feature extraction model is a model trained based on a framework structure of a MPCFormer model and a training logic of a semantic vision transformer model SETR; and

a third training module, configured to train the feature extraction model based on the third baseline feature vector, a label feature vector carried by the third sample image, and a configured first loss function; and train the convolutional neural network based on the third baseline feature vector, the label feature vector carried by the third sample image, and a configured second loss function.

17. A feature extraction apparatus, comprising:

a receiving module, configured to receive an image to-be-processed; and

an obtaining module, configured to obtain a feature vector of the image to-be-processed based on the feature extraction model trained by any method according to claims 1-7 and 11 or based on the convolutional neural network trained by any method according to claims 8-11.

18. An electronic device, wherein the electronic device at least comprises a processor and a memory; and when executing a computer program stored in the memory, the processor is configured to implement steps of any method according to claims 1-13.

19. A computer-readable storage medium, storing a computer program, wherein when executed by a processor, the computer program implements steps of any method according to claims 1-13.