MODEL TRAINING METHOD, DEVICE, COMPUTER EQUIPMENT, AND READABLE STORAGE MEDIUM

[0001] This application claims priority to the Chinese patent application filed on June 9, 2023, with application number 202310677518.8 entitled "Model Training Method, Device, Computer Equipment, and Readable Storage Medium", the entire disclosure of which is hereby incorporated by reference.

TECHNICAL FIELD

[0002] This application relates to the field of data processing technology, particularly to a model training method, device, computer equipment, and readable storage medium.

BACKGROUND

[0003] Federated learning is a new machine learning paradigm, whose core idea is to train the local model with local data on the client and send the model parameters to the server to aggregate the global model. However, in the whole learning process, the image data of different medical institutions are often very different, and the performance of the training model of the client is poor, which leads to the lower accuracy of the federated learning.

SUMMARY

[0004] For this reason, the present application provides a model training method, device, computer equipment and readable storage medium, which is mainly aimed at solving the problem that image data of different medical institutions are often very different, and the training model performance of clients is poor, thus resulting in lower federated learning accuracy.

[0005] A model training method is provided based on the first aspect of the present application, comprising:

[0006] obtaining multiple image data, multiple first labels, and multiple second labels from multiple captured images;

[0007] determining multiple feature vector sets for multiple captured images based on the feature extractor of the preset model;

[0008] training the first classifier of the preset model based on multiple feature vector sets

and multiple first labels;

[0009] training the second classifier of the preset model based on multiple image data and multiple second labels; and

[0010] training the feature extractor based on the multiple feature vector sets, the multiple first labels, the multiple second labels, the first classifier, and the second classifier.

[0011] Optionally, the step of determining multiple feature vector sets of multiple captured images based on the feature extractor of the preset model, further comprising:

[0012] inputting each captured image into the feature extractor in sequence to determine the feature vector set corresponding to each captured image.

[0013] Optionally, the step of training the first classifier of the preset model based on multiple feature vector sets and multiple first labels, further comprising:

[0014] adjusting the parameters of the first classifier by using multiple feature vector sets from multiple captured images as input items and multiple first labels as output items.

[0015] Optionally, the step of training the second classifier of the preset model based on multiple image data and multiple second labels, further comprising:

[0016] adjusting the parameters of the second classifier by using multiple image data from multiple captured images as input items and multiple second labels as output items.

[0017] Optionally, the step of training the feature extractor based on multiple feature vector sets, multiple first labels, multiple second labels, a first classifier, and a second classifier, further comprising:

[0018] inputting multiple feature vector sets from multiple captured images into the first classifier to generate multiple first recognition results;

[0019] comparing multiple first recognition results with multiple first labels to determine whether the multiple first recognition results are the same as the multiple first labels;

[0020] if not, obtaining at least one first target recognition result that is different from the first label among multiple first recognition results, as well as the first target feature vector set corresponding to each first target recognition result;

[0021] training the feature extractor based on at least one first target recognition result and at least one first target feature vector set;

[0022] inputting multiple feature vector sets from multiple captured images into a second classifier to generate multiple second recognition results;

[0023] comparing multiple second recognition results with multiple second labels to determine whether the multiple second recognition results contain the same recognition results as the second labels;

[0024] or, if yes, obtaining at least one second target recognition result that is identical to the second label among multiple second recognition results, as well as the second target feature vector set corresponding to each second target recognition result;

[0025] training the feature extractor based on at least one second target recognition result and at least one second target feature vector set.

[0026] Optionally, the method is applied to multiple clients, each of which is in communication with each server, and the method further comprises:

[0027] determining multiple target clients among multiple clients based on the preset quantity;

[0028] obtaining multiple first model parameters, multiple second model parameters and multiple third model parameters of the preset models of multiple target clients; and

[0029] sending multiple first model parameters, multiple second model parameters and multiple third model parameters to the server.

[0030] Optionally, after sending multiple first model parameters, multiple second model parameters and multiple third model parameters to the server, further comprising:

[0031] receiving the fourth, fifth, and sixth model parameters sent by the server;

[0032] updating the feature extractor based on the fourth model parameters;

[0033] updating the first classifier based on the fifth model parameters; and

[0034] updating the second classifier based on the parameters of the sixth model.

[0035] Based on the second aspect of the present application, there is provided a model training device comprising:

[0036] an acquisition module used to obtain multiple image data, multiple first labels, and multiple second labels from multiple captured images;

[0037] a determination module used to determine multiple feature vector sets for multiple captured images based on the feature extractor of the preset model; and

[0038] a training module used to train the first classifier of the preset model based on multiple feature vector sets and multiple first labels;

[0039] the training module is also used to train the second classifier of the preset model based on multiple image data and multiple second labels; and

[0040] the training module is also used to train the feature extractor based on multiple feature vector sets, multiple first labels, multiple second labels, a first classifier, and a second classifier.

[0041] Optionally, the determination module, specifically for:

[0042] inputting each captured image into the feature extractor in sequence to determine the feature vector set corresponding to each captured image.

[0043] Optionally, the training module, specifically for:

[0044] adjusting the parameters of the first classifier by using multiple feature vector sets from multiple captured images as input items and multiple first labels as output items.

[0045] Optionally, the training module, also specifically for:

[0046] adjusting the parameters of the second classifier by using multiple image data from multiple captured images as input items and multiple second labels as output items.

[0047] Optionally, the device further comprises:

[0048] a generation module used to input multiple feature vector sets from multiple captured images into a first classifier to generate multiple first recognition results; and

[0049] a judgment module used to compare multiple first recognition results with multiple first labels, and to determine whether multiple first recognition results are the same as multiple first labels.

[0050] Optionally, the acquisition module is also used, if not, to acquire at least one first target recognition result different from the first label among the multiple first recognition results, and the first target feature vector set corresponding to each first target recognition result.

[0051] Optionally, the training module is also used to train the feature extractor based on at least one first target recognition result and at least one first target feature vector set.

[0052] Optionally, the generation module is also used to input multiple feature vector sets of multiple captured images into the second classifier to generate multiple second recognition results.

[0053] Optionally, the judgment module is also used to compare multiple second recognition results with multiple second labels to determine whether the multiple second recognition results contain the same recognition result as the second label.

[0054] Optionally, the acquisition module is also used to obtain, if yes, at least one second target recognition result that is the same as the second label among the multiple second recognition results, and the second target feature vector set corresponding to each second target recognition result.

[0055] Optionally, the training module is also used to train the feature extractor based on at least one second target recognition result and at least one second target feature vector set.

[0056] Optionally, the determination module is also used to determine multiple target clients among multiple clients based on the preset quantity.

[0057] The acquisition module is also used to acquire multiple first model parameters, multiple second model parameters and multiple third model parameters of the preset model of

multiple target clients.

[0058] Optionally, the device further comprises:

[0059] a sending module used to send multiple first model parameters, multiple second model parameters and multiple third model parameters to the server.

[0060] Optionally, the device further comprises:

[0061] a receiving module used to receive the fourth, fifth, and sixth model parameters sent by the server;

[0062] an update module, used to update the feature extractor based on the fourth model parameters;

[0063] the update module is also used to update the first classifier based on the fifth model parameters;

[0064] the update module is also used to update the second classifier based on the sixth model parameters.

[0065] A computer equipment is provided based on the third aspect of the present application, comprising a memory and a processor, wherein the said memory stores a computer program, and the said processor implements the step of any one of the methods said in the first aspect when executing the computer program.

[0066] A readable storage medium storing a computer program is provided based on the fourth aspect of the present application, wherein the computer program, when executed by a processor, implements the step of any one of the methods said in the first aspect.

[0067] By virtue of the above technical scheme, the present application provides a model training method, device, computer equipment and readable storage medium. In the process of federated learning in comparison with the prior art, each client uses all the features extracted in the sample image as model training data to conduct model training, so that the feature vectors are concentrated with image features unrelated to the lesion , resulting in poor performance of the training model of the client and low accuracy of the federated learning. The present application proposes to ensure the accuracy of the feature extractor in extracting features related to the lesion by combining the feature extractor with the first classifier for the same target training while conducting confrontation training between the feature extractor and the second classifier, Thus, the accuracy of model parameters of the lesion recognition model provided by each client is higher, and the accuracy of federated learning is effectively improved.

[0068] The above description is only an overview of the technical scheme of the present application. In order to better understand the technical means of the present application, it may

be implemented in accordance with the contents of the description, and in order to make the above and other purposes, features and advantages of the present application more obvious and easy to understand, the specific implementation methods of the application are described below.

BRIEF DESCRIPTION OF DRAWINGS

[0069] Various other advantages and benefits will become apparent to those of ordinary skill in the art by reading the detailed description of the preferred embodiment below. The appended drawings are used only for the purpose of showing the preferred embodiment and are not to be regarded as limiting the present application. Throughout the drawings, the same components are represented by the same reference symbols. In the attached drawings:

[0070] Fig. 1 illustrates a schematic flowchart of a model training method provided by the embodiment of the present application;

[0071] Fig. 2 illustrates a schematic diagram of the model training method provided in the embodiment of the present application;

[0072] Fig. 3 illustrates a schematic diagram of the structure of a model training device provided by the embodiment of the present application.

DETAILED DESCRIPTION OF THE EMBODIMENTS

[0073] The exemplary embodiments of the present application will be described in more detail below with reference to the attached drawings. Although exemplary embodiments of the present application are shown in the attached drawings, it should be understood that the present application may be implemented in various forms and should not be limited by the embodiments described herein. On the contrary, these embodiments are provided to enable a more thorough understanding of the present application and to fully convey the scope of the present application to those skilled in the art.

[0074] The present embodiment provides a model training method, as shown in Fig. 1, which includes:

[0075] S101: Obtaining multiple image data, multiple first labels, and multiple second labels from multiple captured images.

[0076] It is to be understood that the executing body of the invention may be multiple clients of multiple medical institutions, each of which is in communication with a server.

[0077] In this step, multiple captured images are obtained, along with multiple image data,

first labels, and second labels corresponding to each captured image. Wherein, the captured images can be medical images such as magnetic resonance imaging. The first label refers to the category of the lesion category. For example, if the captured image is a lung imaging image, the first label will be the lung health category, which can be lung health, viral pneumonia, bacterial pneumonia, etc.. The second label refers to the source category corresponding to each captured image. For example, if the captured image is a sample image from the first hospital, the second label of the captured image will be the first hospital.

[0078] S102: Determining multiple feature vector set for multiple captured images based on the feature extractor of the preset model.

[0079] In this step, the preset model refers to a neural network model used to perform classification tasks. Specifically, for any client, each captured image is sequentially input into the preset model after multiple captured images are obtained as samples. The feature extractor of the preset model is used to encode the pixel information of the captured images into lower dimensional vectors, which contain the feature vector set of the captured images, and then extract the deep feature vector set of the captured images.

[0080] S103: Training the first classifier of the preset model based on multiple feature vector sets and multiple first labels.

[0081] In this step, the first classifier of the preset model is used to classify different types of lesions in the captured images. When training the first classifier, the feature vector set of each captured image is used as input, and the corresponding first label of the captured image is used as output to train the first classifier and adjust its parameters, so as to enable the first classifier to more accurately identify the type of a lesion.

[0082] S104: Training the second classifier of the preset model based on multiple image data and multiple second labels.

[0083] In this step, the second classifier of the preset model is used to distinguish the source of the feature vector set based on the image features in the feature vector set, in order to identify which client the feature vector set belongs to. Practically, due to the difference of image acquisition sensors for different medical institutions, pictures taken for the same lesion will have different pixels, colors and other differences, which are image features unrelated to the lesion. In order to ensure that the feature extractor only extracts the features related to the lesion, the second classification is used to identify the image features. When the second classifier is trained, multiple image data of each captured image are used as input, and the second label of the captured image is used as output to train the second classifier and adjust its parameters, so as to enable the second classifier to more accurately identify lesion related

features.

[0084] S105: Training the feature extractor based on multiple feature vector sets, multiple first labels, multiple second labels, a first classifier, and a second classifier.

[0085] In this step, the feature vector set of each captured image is sequentially input into the trained first classifier, and the recognition result of whether the feature vector set contains lesion types is obtained based on the output of the first classifier. Afterwards, each recognition result is sequentially compared with the first label of its corresponding feature vector set to determine whether the recognition result is the same as the first label. If the recognition result is different from the first label, it will mean that the first classifier cannot recognize the lesion type based on the feature vector set. At this time, this recognition result needs to be fed back to the feature extractor, so that the feature extractor can optimize and train based on the recognition result and its corresponding feature vector set to achieve parameter adjustment of the feature classifier. Furthermore, the feature vector set corresponding to each captured image is sequentially input into the trained second classifier. The recognition result of whether the feature vector set contains image features is obtained based on the output of the second classifier, such as whether the recognition result is the first hospital or the feature source cannot be identified. Afterwards, each recognition result is sequentially compared with the second label of its corresponding feature vector set. If the recognition result is the same as the second label, it will indicate that the second classifier can identify which client the captured image comes from based on the feature vector set, that is, the feature vector set of the captured image contains image features unrelated to the lesion. At this point, the recognition result needs to be fed back to the feature extractor, so that the feature extractor can optimize and train based on the feature vector set and recognition results to achieve parameter adjustment of the feature extractor.

[0086] Practically, the captured image is input into a feature extractor to obtain the feature vector set of the captured image. The feature vector set is input into the first classifier. If the recognition result output by the first classifier is "No Lesion", it will indicate that the first classifier cannot determine the lesion type based on the feature vector set. At this time, the recognition result is fed back to the feature extractor for parameter adjustment. At the same time, the feature vector set is input into the second classifier. If the recognition result output by the second classifier is "First Hospital" and the second label of the captured image is also "First Hospital", it will indicate that the feature vector set contains image features unrelated to the lesion. At this point, the recognition result is fed back to the feature extractor for parameter adjustment.

[0087] This embodiment of the present application provides a model training method. Compared with the existing technology of federated learning process, each client uses all the features extracted in the sample image as the model training data for model training, so that the feature vectors are concentrated with image features unrelated to the lesion, resulting in poor performance of the training model of the client and low accuracy of federated learning. the present application proposes to conduct confrontational training between the feature extractor and the second classifiers by combining the feature extractor with the first classifiers for the same target training at the same time, so as to ensure the accuracy of the feature extractor in extracting the features related to the lesion, thus making the model parameters of the lesion recognition model provided by each client more accurate, and effectively improving the accuracy of federated learning.

[0088] Furthermore, as a refinement and extension of the specific implementation method of the above embodiment, in order to fully illustrate the specific implementation process of this embodiment, the present application provides another model training method, which includes:

[0089] S201: Obtaining multiple image data, multiple first labels, and multiple second labels from multiple captured images.

[0090] In this step, multiple captured images are obtained, along with multiple image data, first labels, and second labels corresponding to each captured image. Wherein, the captured images can be taken as medical images such as magnetic resonance imaging, etc.. The first label refers to the category of lesions in the captured image. For example, if the captured image is a lung imaging image, the first label will refer to the lung health category, which can be: lung health, viral pneumonia, bacterial pneumonia, etc.. The second label refers to the source category corresponding to each captured image. For example, if the captured image is a sample image from the first hospital, the second label of the captured image will be the first hospital.

[0091] S202: Inputting each captured image into the feature extractor in sequence to determine the feature vector set corresponding to each captured image.

[0092] In this step, after collecting a large number of captured images as samples from each medical institution, each captured image is sequentially input into the feature extractor. The feature extractor encodes the pixel information in each captured image into a lower dimensional vector, which contains the feature vectors of the captured images. Subsequently, the feature extractor outputs the feature vector set of each captured image.

[0093] Practically, a neural network model suitable for the classification task of the model can be adopted as the preset model based on the specific classification task to be performed, and this application does not make specific limitations here. For example, ResNet model,

EfficientNet model, etc. can be used. It should be noted that in the process of federated learning, the preset model used by multiple clients can be the same neural network model or different neural network models. But the set of classification tasks, labels, and output recognition results used by different preset models is the same.

[0094] S203: Adjusting the parameters of the first classifier by using multiple feature vector sets from multiple captured images as input items and multiple first labels as output items.

[0095] In this step, when training the first classifier of the preset model, the feature vector set of each captured image is used as input, and the first label of the captured image is used as output to train the first classifier, and then adjust the parameters of the first classifier, so as to enable the first classifier to accurately identify the type of lesion in the captured image.

[0096] S204: Multiple image data from multiple captured images are used as input items, and multiple second labels are used as output items to adjust the parameters of the second classifier.

[0097] In this step, due to the difference of image acquisition sensors, there are differences in pixels, colors and other features in the pictures taken by each medical institution for the same lesion, which are image features unrelated to the lesion. In order to ensure that the feature extractor only extracts the features related to the lesion, the second classification is used to identify the image features. When the second classifier is trained, multiple image data of each captured image are taken as input, and the second label of the captured image is taken as output to train the second classifier, and then adjust the parameters of the second classifier. This enables the second classifier to more accurately identify lesions related features.

[0098] S205: Inputting multiple feature vector sets from multiple captured images into the first classifier to generate multiple first recognition results.

[0099] In this step, the purpose of federated learning is to jointly train a lesion recognition model with data from different medical institutions. During each round of training, each medical institution conducts model training based on its own local data, sends the trained model parameters to the server in a unified manner, aggregates them through the server, and distributes them to each medical institution. That is, the local models of the clients of various medical institutions should be able to accurately identify the lesions. However, the feature vector set extracted by the feature extractor contains not only the features related to the lesion, but also the image features unrelated to the lesion. In order to improve the accuracy of the model parameters uploaded by each client, the feature extractor needs to be optimized so that the feature extractor only extracts the features related to the lesion. Specifically, the feature vector set of each captured image is input to the first trained classifier, and the first recognition

result of whether the feature vector set contains the type of lesion is obtained based on the output of the first classifier.

[00100] S206: Comparing multiple first recognition results with multiple first labels to determine whether they are all the same. If so, proceed to step S209. If not, proceed to step S207.

[00101] In this step, the first recognition result of each captured image is compared with the first label of the captured image to determine whether the first recognition result is the same as the first label. If all the first recognition results are the same as the first label, it will mean that the first classifier can recognize the type of lesions in the captured image based on the feature vector set, that is, the feature vectors related to the lesions in the feature vector set extracted by the feature extractor are accurate. If any of the first recognition results is different from the first label, it will mean that the first classifier is unable to identify the type of lesions in the captured image based on the feature vector set, or the type of lesions identified is inconsistent with the correct type of lesions, that is, the feature vectors related to the lesions in the feature vector set extracted by the feature extractor are wrong or incomplete. At this time, the feature extractor needs to be optimized to adjust the parameters of the feature extractor.

[00102] S207: Obtaining at least one first target recognition result that is different from the first label among multiple first recognition results, as well as the first target feature vector set corresponding to each first target recognition result.

[00103] S208: Training the feature extractor based on at least one first target recognition result and at least one first target feature vector set.

[00104] In the steps S207 and S208, the multiple first recognition results recognized by the first classifier include the recognition results different from the first label, that is, if there are errors in multiple first recognition results, at least one wrong first target recognition result will be called out in multiple first recognition results, and the corresponding first target feature vector set will be determined based on each first target recognition result. Afterwards, each first target recognition result and its corresponding first target feature vector set are fed back to the feature extractor, so that the feature extractor optimizes the model based on at least one first target recognition result and at least one first target feature vector set which are fed back.

[00105] Alternatively, when the feature extraction optimizes at least one first target recognition result and at least one first target feature vector set based on feedback, the optimized feature extractor is used to extract the feature vector set of each captured image again, inputting the feature vector set into the first classification, obtaining the first recognition

result, and judging the accuracy of the first recognition result. If the first recognition result is correct, the feature extractor will stop to be optimized; If the first recognition result is still wrong, the feature extractor will continue to be optimized until the first recognition result is correct.

[00106] S209: No need to optimize the feature extractor.

[00107] In this step, it is determined that all the first recognition results identified by the first classifier are the same as multiple first labels, indicating that the multiple first recognition results are accurate and correct. That is, the feature extractor extracts comprehensive and accurate feature data related to the lesion, so there is no need to optimize the feature extractor.

[00108] In the above manner, the feature extractor is optimized based on the recognition results related to the type of lesions identified by the first classifier, so as to achieve the same target training of the feature extractor and the first classifier, and improve the accuracy of the features related to the lesions extracted by the feature extractor.

[00109] S210: Inputting multiple feature vector sets from multiple captured images into the second classifier to generate multiple second recognition results.

[00110] In this step, the feature vectors extracted by the feature extractor not only include the features related to the lesion, but also may include the image features unrelated to the lesion, such as image pixels, image yellowing and other color difference features. In order to improve the accuracy of the lesion recognition model of the federated learning, it is necessary to ensure the accuracy of the model parameters provided by each client, and it is necessary to verify the feature vectors extracted from the feature extractor to determine whether the feature vector set contains image features unrelated to the lesion, and then optimize the feature extractor based on the verification results. Specifically, the feature vector set of each captured image is input to the second classifier, and according to the output of the second classifier, the second recognition result of whether to identify which client the feature vector set is from is obtained.

[00111] S211: Comparing multiple second recognition results with multiple second labels to determine whether they contain the same recognition result as the second label. If so, proceed to step S212. If not, proceed to step S214.

[00112] In this step, each second recognition result is compared with its corresponding second label to determine whether the second recognition result is the same as the second label. If there is the same recognition result as the second label in the multiple second recognition results, which indicates that the second classifiers recognize which client the feature vector set comes from, it will be determined that the feature vector set contains the image features of the client, and the feature extractor will need to be optimized. Further, if the multiple second

recognition results are different from the second label, it will mean that the second classifiers cannot identify which client the feature vector set comes from, and at this time it will be determined that the feature vector set does not contain the image feature of the client.

[00113] S212: If so, obtaining at least one second target recognition result that is identical to the second label among multiple second recognition results, as well as the second target feature vector set corresponding to each second target recognition result.

[00114] S213: Training the feature extractor based on at least one second target recognition result and at least one second target feature vector set.

[00115] In steps S212 and S213, when the second recognition result recognized by the second classifier is determined to include the same recognition result as the second label, that is, when the feature vector set contains the image features of the client, at least one second target recognition result identical to the second label is called out in the multiple second recognition results, and the corresponding second target feature vector set is determined based on each second target recognition result. Thereafter, each second target recognition result and its corresponding second target feature vector set are fed back to the feature extractor, so that the feature extractor is optimized based on the feedback of at least one second target recognition result and at least one second target feature vector set.

[00116] Alternatively, when the feature extractor optimizes at least one second target recognition result based on feedback and the set of at least one second target feature vector, the optimized feature extractor is used to extract the feature vector set of each captured image again, input the feature vector set into the second classifier to obtain the second recognition result, and compare the second recognition result with the second label. If the second recognition results are different from the second label, the optimization of the feature extractor shall be stopped; If any of the second recognition results is the same as the second label, the feature extraction will continue to be optimized until the second recognition results are different from the second label.

[00117] S214: No need to optimize the feature extractor.

[00118] In this step, it is determined that the second recognition results identified by the second classifier are different from those of the second label, that is, the feature vector set extracted by the feature extractor does not contain image features unrelated to the lesions, so it is not necessary to optimize the feature extractor.

[00119] In the above manner, the feature extractor is optimized based on the recognition results related to the source of the feature identified by the second classifier to realize the confrontation training between the feature extractor and the first classifier, so that the feature

extractor will not extract the features unrelated to the lesion. On the one hand, the accuracy of the feature extractor in extracting the features related to the lesion is improved; On the other hand, the image features of the client will not flow out of the client, thus ensuring the privacy of the data.

[00120] S215: Determining multiple target clients among multiple clients based on the preset quantity.

[00121] In this step, federated learning can jointly train a lesion recognition model using data from different medical institutions. Specifically, federated learning includes clients from multiple medical institutions, with each client communicating with the server. The entire federated learning process is divided into multiple communication rounds. In any communication round, each client trains a local preset model based on local data. Afterwards, the parameters of the trained model are sent to the server for aggregation and distribution to clients of various medical institutions for the next round of training. However, throughout the entire federated learning process, if each client interacts with the server for data in each round of training, the client will need to frequently communicate with the server, which also limits the efficiency of training the lesion recognition model of federated learning. And the communication cost is relatively high. Therefore, in order to improve the efficiency of federated learning, during each round of training, some clients are selected from multiple clients for interaction, reducing the cost of single communication and overall communication frequency. Specifically, for any communication round, multiple target clients are randomly selected from multiple clients based on a preset number as the clients for transmitting model parameters in this round.

[00122] Alternatively, the preset number ranges from 10% to 50% of the total number of clients, and the specific number can be determined according to the current round of communication signals, and the present application does not make specific moves here. Further, when the current round of training is completed, the multiple target clients selected for the current round are marked, and the remaining clients other than the marked multiple target clients are randomly selected again at the next round of training.

[00123] S216: Obtaining multiple first model parameters, multiple second model parameters and multiple third model parameters of the preset model of multiple target clients.

[00124] S217: Sending multiple first model parameters, multiple second model parameters and multiple third model parameters to the server.

[00125] In steps S216 and S217, after screening multiple target clients from multiple clients, the first model parameters, the second model parameters and the third model parameters of

the preset model of each target client are obtained. Wherein the first model parameter refers to the parameter of the trained feature extractor, the second model parameter refers to the parameter of the trained first classifiers, and the third model parameter refers to the parameter of the trained second classifiers. Subsequently, multiple first model parameters, multiple second model parameters and multiple third model parameters of multiple target clients are sent to the server for the server to aggregate and distribute the received model parameters.

[00126] Alternatively, after receiving multiple first model parameters, multiple second model parameters and multiple third model parameters, the server calculates the multiple first model parameters, multiple second model parameters and multiple third model parameters by means of weighted average to obtain new fourth model parameters, fifth model parameters and sixth model parameters.

[00127] In the above way, model parameters of some clients are selected for transmission, which reduces the communication time, improves the transmission efficiency, and improves the efficiency of federated learning of lesion recognition model training.

[00128] S218: Receiving the fourth, fifth, and sixth model parameters sent by the server.

[00129] S219: Updating the feature extractor based on the fourth model parameters.

[00130] S220: Updating the first classifier based on the fifth model parameters.

[00131] S221: Updating the second classifier based on the sixth model parameters.

[00132] In steps S218 to S221, the server calculates the weighted average based on the received multiple first model parameters, multiple second model parameters, and multiple third model parameters to obtain the fourth model parameters, the fifth model parameters, and the sixth model parameters. Afterwards, the fourth model parameters, the fifth model parameters and the sixth model parameters are sent to each client respectively. After receiving the fourth model parameters, the fifth model parameters and the sixth model parameters sent by the server, each client updates the feature extractor based on the fourth model parameters, updates the first classifier based on the fifth model parameters, and updates the second classifier based on the sixth model parameters.

[00133] In the above way, the server is introduced to be responsible for the aggregation and distribution of the model while each client uses local data to train the in-depth learning model. This realizes the training of the lesion recognition model and ensures the data privacy of each medical institution.

[00134] As an embodiment, the embodiment of the present application provides a model training method, as shown in Fig. 2 which illustrates a schematic block diagram of the model training method. Wherein each client carries out the same target training for the feature

extractor and the first classifiers respectively, and at the same time, the feature extractor and the second classifiers are trained against each other to obtain the first model parameters, the second model parameters and the third model parameters after training. Afterwards, the selected multiple first model parameters, multiple second model parameters and multiple third model parameters of multiple target clients are sent to the server separately for aggregating the model parameters in the form of weighted average. Subsequently, the calculated fourth model parameters, fifth model parameters and sixth model parameters sent by the receiving server update the feature extractor, the first classifier and the second classifier respectively. In the above methods, the accuracy of model training in federated learning scenarios with heterogeneous features can be ensured, and the degradation of model performance caused by heterogeneous features can be effectively avoided.

[00135] Furthermore, as a specific implementation of the method described in Fig. 1, the present embodiment provides a model training device 400, as shown in Fig. 3, which includes:

[00136] An acquisition module 401, used to obtain multiple image data, multiple first labels, and multiple second labels of multiple captured images;

[00137] A determination module 402, used to determine multiple feature vector sets for multiple captured images based on the feature extractor of the preset model;

[00138] A training module 403, used to train the first classifier of the preset model based on multiple feature vector sets and multiple first labels;

[00139] The training module 403 is also used to train the second classifier of the preset model based on multiple image data and multiple second labels;

[00140] The training module 403 is also used to train the feature extractor based on multiple feature vector sets, multiple first labels, multiple second labels, a first classifier, and a second classifier.

[00141] Optionally, the determine module 402 is specifically for:

[00142] Input each captured image into the feature extractor in sequence to determine the feature vector set corresponding to each captured image.

[00143] Optionally, the training module 403 is specifically used for:

[00144] Adjusting the parameters of the first classifier by using multiple feature vector sets from multiple captured images as input items and multiple first labels as output items.

[00145] Optionally, the training module 403 is also specifically used for:

[00146] Adjusting the parameters of the second classifier by using multiple image data from multiple captured images as input items and multiple second labels as output items.

[00147] Optionally, the device also includes:

[00148] A generation module 404, used to input multiple feature vector sets of multiple captured images into the first classifier to generate multiple first recognition results;

[00149] A judgment module 405, used to compare multiple first recognition results with multiple first labels, and determine whether multiple first recognition results are the same as multiple first labels.

[00150] Optionally, the acquisition module 401 is also used to, if not, obtain at least one first target recognition result that is different from the first label among multiple first recognition results, as well as the first target feature vector set corresponding to each first target recognition result.

[00151] Alternatively, the training module 403 is also used to train the feature extractor based on at least one first target recognition result and at least one first target feature vector set.

[00152] Alternatively, the generation module 404 is also used to input multiple feature vector sets of multiple captured images into the second classifier, generating multiple second recognition results.

[00153] Optionally, the judgment module 405 is also used to compare multiple second recognition results with multiple second labels, and determine whether the multiple second recognition results contain the same recognition results as the second labels.

[00154] Alternatively, the acquisition module 401 is also used to obtain at least one second target recognition result that is identical to the second label among multiple second recognition results, as well as a second target feature vector set corresponding to each second target recognition result.

[00155] Optionally, the training module 403 is also used to train the feature extractor based on at least one second target recognition result and at least one second target feature vector set.

[00156] Alternatively, the determination module 402 is also used to determine multiple target clients among multiple clients based on a preset quantity;

[00157] The acquisition module 401 is also used to acquire multiple first model parameters, multiple second model parameters and multiple third model parameters of the preset model of multiple target clients.

[00158] Optionally, the device also includes:

[00159] A sending module 406, used to send multiple first model parameters, multiple second model parameters and multiple third model parameters to the server.

[00160] Optionally, the device also includes:

[00161] A receiving module 407, used to receive the fourth, fifth, and sixth model parameters sent by the server;

[00162] An update module 408, used to update the feature extractor based on the fourth model parameters;

[00163] The update module 408 is also used to update the first classifier based on the fifth model parameters;

[00164] The update module 408 is also used to update the second classifier based on the sixth model parameters.

[00165] The embodiments of the present application provides a model training device 400. Compared with the federated learning process of the prior art, each client uses all the features extracted in the sample image as the model training data to conduct model training, so that the feature vectors are concentrated with image features unrelated to the lesion, resulting in poor performance of the training model of the client and low precision of the federated learning. the present application proposes to conduct confrontational training between the feature extractor and the second classifiers by combining the feature extractor with the first classifiers for the same target training, so as to improve the accuracy of the feature extractor to extract the features related to the lesion, thus making the model parameters of the lesion recognition model provided by each client more accurate, and effectively improving the accuracy of the federated learning.

[00166] In the exemplary embodiment, the present application also provides a computer device, including a memory and a processor. The memory stores a computer program and a processor, which are used to execute the program stored on the memory and execute the model training method in the above embodiments.

[00167] In the exemplary embodiments, the present application also provides a readable storage medium on which a computer program is stored, as well as the steps of the model training method being implemented when the computer program is executed by a processor.

[00168] From the above description of the implementation mode, those skilled in the art can clearly understand that the present application can be realized by means of hardware, or by means of software plus the necessary general hardware platform. Based on such understanding, the technical scheme of the present application may be embodied in the form of a software product, which may be stored in a non-volatile storage medium (may be a CD-ROM, a U-disk, a mobile hard disk, etc.), or a volatile storage medium, including a number of instructions to enable a computer device (may be a personal computer, a server, or a network device, etc.) to perform the methods described in the various implementation scenarios of the present application.

[00169] Those skilled in the art may understand that the drawings are merely schematic

diagrams of a preferred implementation scenario, and that the modules or processes in the drawings are not necessarily necessary for the implementation of the present application.

[00170] Those skilled in the art may understand that the modules in the devices in the implementation scenarios may be distributed among the devices in the implementation scenarios according to the description of the implementation scenarios, and may also be located in one or more devices different from the implementation scenarios based on the corresponding changes. The modules of the above implementation scenarios may be combined into one module or further divided into multiple sub-modules.

[00171] The above serial numbers of the present application are only for description and do not represent the advantages and disadvantages of the implementation scenarios.

[00172] The above disclosure is only a few specific implementation scenarios of the present application, but the present application is not limited to this, and any change that a person skilled in the art can contemplate shall fall within the scope of protection of the present application.

What is claimed is:

1. A model training method, comprising:

obtaining multiple image data, multiple first labels, and multiple second labels from multiple captured images;

determining multiple feature vector sets for the said multiple captured images based on the feature extractor of the preset model;

training the first classifier of the preset model based on the said multiple feature vector sets and the said multiple first labels;

training the second classifier of the preset model based on the said multiple image data and the multiple second labels; and

training the said feature extractor based on the said multiple feature vector sets, the said multiple first labels, the said multiple second labels, the said first classifier, and the said second classifier.

2. The method of claim 1, wherein the said step of determining multiple feature vector sets of the said multiple captured images based on the feature extractor of the preset model further comprises:

inputting each captured image into the said feature extractor sequentially, and determining the said feature vector set corresponding to each captured image.

3. The method of claim 1, wherein the said step of training the said first classifier of the preset model based on the said multiple feature vector sets and the said multiple first labels further comprises:

adjusting the parameters of the said first classifier by using the said multiple feature vector sets of the said captured images as input items and the said multiple first labels as output items.

4. The method of claim 3, wherein the step of training the second classifier of the preset model based on the said multiple image data and the said multiple second labels further comprises:

adjusting the parameters of the said second classifier by using the said multiple image data of the said captured images as input items and the said multiple second labels as output items.

5. The method of claim 4, wherein the step of training the said feature extractor based on the said multiple feature vector sets, the said multiple first labels, the said multiple second labels, the said first classifier, and the said second classifier further comprises:

inputting the said multiple feature vector sets of the said captured images into the said

first classifier to generate multiple first recognition results;

comparing the said multiple first recognition results with the said multiple first labels, and determining whether the said multiple first recognition results are the same as the said multiple first labels;

if not, obtaining at least one first target recognition result that is different from the first label among the said multiple first recognition results, and the first target feature vector set corresponding to each first target recognition result;

training the said feature extractor based on at least one first target recognition result and at least one first target feature vector set;

inputting the said multiple feature vector sets of the said captured images into the second classifier to generate multiple second recognition results; and

comparing the said multiple second recognition results with the said multiple second labels to determine whether the said multiple second recognition results contain the same recognition results as the second labels;

or, if yes, obtaining at least one second target recognition result as well as the second target feature vector set corresponding to each second target recognition result that are identical to the second label among the multiple second recognition results; and

training the feature extractor based on at least one said second target recognition result and at least one second target feature vector set.

6. The method of any one of claims 1 to 5, wherein it is applied to multiple clients, each client is in communication connection with the server, and the said method further comprises:

determining multiple target clients among the said multiple clients based on the preset quantity;

obtaining multiple first model parameters, multiple second model parameters and multiple third model parameters of preset models of multiple target clients; and

sending the said multiple first model parameters, the said multiple second model parameters and the said multiple third model parameters to the said server.

7. The method of claim 6, wherein after the said multiple first model parameters, the said multiple second model parameters and the said multiple third model parameters are sent to the said server, the said steps further comprise:

receiving the fourth, fifth, and sixth model parameters sent by the said server;

updating the feature extractor based on the said fourth model parameters;

updating the said first classifier based on the fifth model parameters; and

updating the second classifier based on the said sixth model parameters.

8. A model training device, comprising:

an acquisition module used to obtain multiple image data, multiple first labels, and multiple second labels from multiple captured images;

a determination module used to determine multiple feature vector sets of the said captured images based on the feature extractor of the preset model;

a training module used to train the said first classifier of the preset model based on the said multiple feature vector sets and the said multiple first labels;

the said training module is also used to train the second classifier of the preset model based on the said multiple image data and the said multiple second labels; and

the said training module is also used to train the said feature extractor based on the said multiple feature vector sets, the said multiple first labels, the said multiple second labels, the said first classifier, and the said second classifier.

9. A computer equipment comprises a memory and a processor, the memory stores a computer program, wherein the processor implements a model training method when executing the computer program, comprising:

obtaining multiple image data, multiple first labels, and multiple second labels from multiple captured images; determining multiple feature vector sets for the said multiple captured images based on the feature extractor of the preset model; training the first classifier of the preset model based on the said multiple feature vector sets and the said multiple first labels; training the second classifier of the preset model based on the said multiple image data and the said multiple second labels; and training the said feature extractor based on the said multiple feature vector sets, the said multiple first labels, the said multiple second labels, the said first classifier, and the said second classifier.

10. The computer equipment of claim 9, wherein the said processor, when executing a computer program, implements the said feature extractor based on the preset model to determine multiple feature vector sets of the said multiple captured images, further comprising:

inputting each captured image into the said feature extractor sequentially, and determining the feature vector set corresponding to each said captured image.

11. The computer equipment of claim 9, wherein the computer-readable instructions, when executed by the processor, implement the training of the said first classifier of the preset model based on the said multiple feature vector sets and the said multiple first labels, further comprising:

adjusting the parameters of the said first classifier by using the said multiple feature vector sets of the said captured images as input items and the said multiple first labels as output

items.

12. The computer equipment of claim 11, wherein the said processor, when executing a computer program, realizes the training of the second classifier of the preset model based on the said multiple image data and the said multiple second labels, further comprising:

adjusting the parameters of the said second classifier by using the said multiple image data of the said captured images as input items and the said multiple second labels as output items.

13. The computer equipment of claim 12, wherein the processor, when executing a computer program, implements the training of the said feature extractor based on the said multiple feature vector sets, the said multiple first labels, the said multiple second labels, the said first classifier and the said second classifier, further comprising:

inputting the said multiple feature vector sets of the said captured images into the said first classifier to generate multiple first recognition results; comparing the said multiple first recognition results with the said multiple first labels, and determining whether the said multiple first recognition results are the same as the said multiple first labels; if not, obtaining at least one first target recognition result that is different from the first label among the multiple first recognition results, and the first target feature vector set corresponding to each first target recognition result; training the said feature extractor based on the said at least one first target recognition result and at least one first target feature vector set; inputting the said multiple feature vector sets of the said captured images into the said second classifier to generate multiple second recognition results; comparing the said multiple second recognition results with the said multiple second labels to determine whether the said multiple second recognition results contain the same recognition results as the second labels; or, if yes, obtaining at least one second target recognition result as well as the second target feature vector set corresponding to each second target recognition result that are identical to the second label among the multiple second recognition results, and training the said feature extractor based on the said at least one second target recognition result and at least one second target feature vector set.

14.The computer equipment of claim 9, wherein, applied to multiple clients, each client is in communication with the server, and the processor implements the determination of multiple target clients among the multiple clients based on a predetermined number when executing a computer program; obtaining multiple first model parameters, multiple second model parameters and multiple third model parameters of the preset model of multiple target clients; and sending the said multiple first model parameters, the said multiple second model

parameters and the said multiple third model parameters to the said server.

15. A computer-readable storage medium storing a computer program, wherein the computer program, when executed by a processor, implements a model training method, comprising:

obtaining multiple image data, multiple first labels, and multiple second labels from multiple captured images; determining multiple feature vector sets for the said multiple captured images based on the feature extractor of the preset model; training the first classifier of the preset model based on the said multiple feature vector sets and the said multiple first labels; training the second classifier of the preset model based on the said multiple image data and the multiple second labels; and training the said feature extractor based on the said multiple feature vector sets, the said multiple first labels, the said multiple second labels, the said first classifier, and the said second classifier.

16. The computer-readable storage medium of claim 15, wherein the said computer-readable instructions, when executed by a processor, implement the said feature extractor based on the preset model to determine multiple feature vector sets of the said multiple captured images, further comprising:

inputting each captured image into the said feature extractor sequentially, and determining the feature vector set corresponding to the said each captured image.

17. The computer-readable storage medium of claim 15, wherein the computer-readable instructions, when executed by the processor, implement the training of the said first classifier of the preset model based on the said multiple feature vector sets and the said multiple first labels, further comprising:

adjusting the parameters of the said first classifier by using the said multiple feature vector sets of the captured images as input items and the said multiple first labels as output items.

18. The computer-readable storage medium of claim 17, wherein the computer-readable instructions, when executed by a processor, implement the training of a second classifier of the preset model based on the said multiple image data and the multiple second labels, further comprising:

adjusting the parameters of the second classifier by using the said multiple image data of the said captured images as input items and the said multiple second labels as output items.

19. The computer-readable storage medium of claim 18, wherein the computer-readable instructions, when executed by a processor, implement the training of the feature extractor based on the said multiple feature vector sets, the said multiple first labels, the multiple second

labels, the said first classifier, and the second classifier, further comprising:

inputting the said multiple feature vector sets of the said captured images into the said first classifier to generate multiple first recognition results; comparing the said multiple first recognition results with the said multiple first labels, and determining whether the said multiple first recognition results are the same as the said multiple first labels; if not, obtaining at least one first target recognition result that is different from the first label among the multiple first recognition results, and the first target feature vector set corresponding to each first target recognition result; training the said feature extractor based on the said at least one first target recognition result and at least one first target feature vector set; inputting the said multiple feature vector sets of the said captured images into the said second classifier to generate multiple second recognition results; comparing the said multiple second recognition results with the said multiple second labels to determine whether the said multiple second recognition results contain the same recognition results as the second labels; or, if yes, obtaining at least one second target recognition result as well as the second target feature vector set corresponding to each second target recognition result that are identical to the second label among the multiple second recognition results; training the said feature extractor based on the said at least one second target recognition result and at least one second target feature vector set.

20. The computer-readable storage medium of claim 15, wherein, applied to multiple clients, each client is in communication with the server, and the said computer-readable instructions, when implemented by the processor, realizes the determination of multiple target clients among the multiple clients based on a predetermined number; obtaining multiple first model parameters, multiple second model parameters and multiple third model parameters of the preset model of multiple target clients; and sending the said multiple first model parameters, the said multiple second model parameters and the said multiple third model parameters to the said server.