

METHOD FOR UPDATING A DATA AUGMENTATION STRATEGY AND APPARATUS, DEVICE, AND STORAGE MEDIUM

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application is filed based upon and claims priority to Chinese Patent Application No. 202010519507.3, filed on June 09, 2020, the disclosure of which is hereby incorporated by reference in its entirety.

TECHNICAL FIELD

[0002] Embodiments of the disclosure relate to the field of machine learning, and particularly to a method, apparatus, and device for updating a data augmentation strategy, and a storage medium.

BACKGROUND

[0003] An application effect of a deep learning technology depends on a large amount of training data, and a data processing model obtained by training with a limited amount of training data may usually have an over-fitting phenomenon. In order to improve a training effect of a data processing model and reduce manpower needed by model training, an automatic data augmentation technology is gradually used to improve the data amount and diversity of training data.

[0004] The automatic data augmentation technology refers to an automatic data augmentation process based on an automatic machine learning technology, and thus it is critical to find a proper data augmentation strategy. Generally, a data augmentation strategy may be optimized by a reinforcement learning algorithm based on a training effect of a data processing model.

[0005] The amount of training data is usually relatively large, and relatively more time is needed by training a data processing model, so the efficiency of the generation of a data augmentation strategy needs to be improved.

SUMMARY

[0006] The embodiments of the disclosure provide a method, apparatus, and device for updating a data augmentation strategy, and a storage medium.

[0007] According to a first aspect, an embodiment of the disclosure provides a method for updating a data augmentation strategy, which may include the following operations.

[0008] An initial data augmentation strategy is acquired.

[0009] Second-stage training is performed on a preset data processing model subjected to first-stage training according to the data augmentation strategy and preset training data.

[0010] The data augmentation strategy is updated according to the data processing model subjected to second-stage training, to obtain an updated data augmentation strategy.

[0011] In a possible implementation mode, the method may further include the following operations.

[0012] A data augmentation strategy subjected to M updates is acquired, M being more than or equal to 1.

[0013] Second-stage training is performed on the data processing model subjected to first-stage training according to the data augmentation strategy obtained subjected to M updates and the training data.

[0014] An (M+1)th update is performed on the data augmentation strategy according to the data processing model subjected to second-stage training.

- [0015] In a possible implementation mode, the initial data augmentation strategy may include multiple initial data augmentation strategies, which are concurrently updated. The method may further include the following operations.
- [0016] After every preset number of updates, an optimal strategy is selected from the updated multiple data augmentation strategies according to the data processing model subjected to second-stage training.
- [0017] Each of the updated data augmentation strategies other than the optimal data augmentation strategy is replaced with the optimal data augmentation strategy.
- [0018] In a possible implementation mode, the data augmentation strategy may include multiple preset data augmentation operations. The operation that second-stage training is performed on a preset data processing model subjected to first-stage training according to the data augmentation strategy and preset training data may include the following operations.
- [0019] Data augmentation is sequentially performed on the training data according to each of the data augmentation operations.
- [0020] Second-stage training is performed on the data processing model subjected to first-stage training according to the training data subjected to data augmentation.
- [0021] In a possible implementation mode, the operation that the data augmentation strategy is updated according to the data processing model subjected to second-stage training may include the following operations.
- [0022] A preset strategy model is updated according to the data processing model subjected to second-stage training.
- [0023] A selection probability of each of preset strategies is determined by using the updated strategy model.
- [0024] The updated data augmentation strategy is selected from the preset strategies according to the selection probabilities of the preset strategies.

- [0025] In a possible implementation mode, in the case that the data augmentation strategy is updated for multiple times, the operation that the preset strategy model is updated according to the data processing model subjected to second-stage training may include the following operations.
- [0026] The data processing model subjected to second-stage training is checked according to preset verification data to obtain a checking result.
- [0027] Historical checking results of the data processing model subjected to second-stage training in first N-1 updates of the data augmentation strategy are acquired, N being a present total number of times for which the data augmentation strategy has been updated.
- [0028] The strategy model is updated according to the historical checking results and the checking result.
- [0029] In a possible implementation mode, the operation that the strategy model is updated according to the historical checking results and the checking result may include the following operations.
- [0030] An average value of the historical checking results is determined.
- [0031] A difference between the checking result and the average value is determined.
- [0032] A strategy parameter in the strategy model is updated according to the difference.
- [0033] In a possible implementation mode, before the operation that the initial data augmentation strategy is acquired, the method may further include the following operations.
- [0034] A data augmentation strategy in the first-stage training is uniformly and randomly selected from preset strategies.
- [0035] First-stage training is performed on the data processing model according to the data augmentation strategy in the first-stage training and the training data.

[0036] According to a second aspect, an embodiment of the disclosure provides a data processing method, which may include the following operations.

[0037] Data to be processed is acquired.

[0038] The data to be processed is processed by using a pre-trained data processing model. Herein, the data processing model may be sequentially subjected to first-stage training and second-stage training, the data processing model may be trained by using a preset data augmentation strategy and preset training data in the second-stage training, and the data augmentation strategy may be generated by the method for updating the data augmentation strategy as described in the first aspect or each possible implementation mode of the first aspect.

[0039] In a possible implementation mode, the method may further include the following operations.

[0040] First-stage training is performed on the data processing model according to the training data.

[0041] Data augmentation is performed on the training data by using the data augmentation strategy.

[0042] Second-stage training is performed on the data processing model subjected to first-stage training according to the training data subjected to data augmentation.

[0043] In a possible implementation mode, the operation that first-stage training is performed on the data processing model according to the training data may include the following operations.

[0044] A data augmentation strategy in first-stage training is uniformly and randomly selected from preset strategies.

[0045] First-stage training is performed on the data processing model according to the data augmentation strategy in first-stage training and the training data.

[0046] In a possible implementation mode, the data to be processed and the training

data may be image data or text data.

[0047] According to a third aspect, an embodiment of the disclosure provides an apparatus for updating a data augmentation strategy, which may include an acquisition part, a training part and a processing part.

[0048] The acquisition part may be configured to acquire an initial data augmentation strategy.

[0049] The training part may be configured to perform second-stage training on a preset data processing model subjected to first-stage training according to the data augmentation strategy and preset training data.

[0050] The updating part may be configured to update the data augmentation strategy according to the data processing model subjected to second-stage training to obtain an updated data augmentation strategy.

[0051] According to a fourth aspect, an embodiment of the disclosure provides a data processing apparatus, which may include an acquisition part and a processing part.

[0052] The acquisition part may be configured to acquire data to be processed.

[0053] The processing part may be configured to process the data to be processed by using a pre-trained data processing model. Herein, the data processing model may be sequentially subjected to first-stage training and second-stage training, the data processing model may be trained by a preset data augmentation strategy and preset training data in a second training stage, and the data augmentation strategy may be generated by the method for updating the data augmentation strategy as described in the first aspect or each possible implementation mode of the first aspect.

[0054] According to a fifth aspect, an embodiment of the disclosure provides an electronic device, which may include:

[0055] a memory and a processor.

[0056] The memory may be configured to store program instructions.

[0057] The processor may be configured to call the program instructions in the memory to execute the method as described in the first aspect, each possible implementation mode of the first aspect, the second aspect or each possible implementation mode of the second aspect.

[0058] According to a sixth aspect, an embodiment of the disclosure provides a computer-readable storage medium having stored therein a computer program which, when executed, may implement the method as described in the first aspect, each possible implementation mode of the first aspect, the second aspect or each possible implementation mode of the second aspect.

[0059] According to a seventh aspect, an embodiment of the disclosure provides a computer program, which may include a computer-readable code. When the computer-readable code runs in an electronic device, a processor in the electronic device may execute the method as described in the first aspect, each possible implementation mode of the first aspect, the second aspect or each possible implementation mode of the second aspect.

[0060] According to the method for updating a data augmentation strategy provided in the embodiments of the disclosure, a training stage of the data processing model is divided into two stages, i.e., a first stage and a second stage. When the data augmentation strategy is updated, second-stage training is performed on the data processing model subjected to first-stage training based on the data augmentation strategy and the training data, and then the data augmentation strategy is updated based on the data processing model subjected to second-stage training. Therefore, the data processing model needs not to be totally retrained in an updating process of the data augmentation strategy, the quality of the data augmentation strategy is ensured, and meanwhile, the efficiency of the generation of the data augmentation strategy is improved. In addition, the generated data augmentation strategy may be applied to similar data of the training data and is transferrable.

[0061] It is to be understood that the above general description and the following detailed description are only exemplary and explanatory and not intended to limit the disclosure.

BRIEF DESCRIPTION OF THE DRAWINGS

- [0062] The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate embodiments consistent with the disclosure and, together with the specification, serve to explain the principle of the disclosure.
- [0063] FIG. 1 is an example diagram of a relationship between data augmentation and a training effect of an image classification model.
- [0064] FIG. 2 is a schematic diagram of a network architecture according to an embodiment of the disclosure.
- [0065] FIG. 3 is a flowchart of a method for updating a data augmentation strategy according to an embodiment of the disclosure.
- [0066] FIG. 4 is a flowchart of a method for updating a data augmentation strategy according to another embodiment of the disclosure.
- [0067] FIG. 5 is a flowchart of a method for updating a data augmentation strategy according to another embodiment of the disclosure.
- [0068] FIG. 6 is a flowchart of a method for updating a data augmentation strategy according to another embodiment of the disclosure.
- [0069] FIG. 7 is an example diagram of concurrently updating multiple data augmentation strategies according to another embodiment of the disclosure.
- [0070] FIG. 8 is a flowchart of a data processing method according to an embodiment of the disclosure.
- [0071] FIG. 9 is a structure diagram of an apparatus for updating a data augmentation strategy according to an embodiment of the disclosure.
- [0072] FIG. 10 is a structure diagram of a data processing apparatus according to an embodiment of the disclosure.

[0073] FIG. 11 is a structure diagram of an electronic device according to an embodiment of the disclosure.

[0074] FIG. 12 is a block diagram of an apparatus for updating a data augmentation strategy according to the present embodiment.

[0075] Specific embodiments of the disclosure have been shown in the drawings, and more detailed descriptions will be made hereinafter. These drawings and text descriptions are not for limiting the scope of the concept of the disclosure in any manner but for explaining the concept of the disclosure to those skilled in the art with reference to the specific embodiments.

DETAILED DESCRIPTION

[0076] Exemplary embodiments will now be described in detail, and examples thereof are represented in the drawings. When the following descriptions involve the drawings, the same numerals in different drawings represent the same or similar elements, unless otherwise indicated. Implementation modes described in the following exemplary embodiments do not represent all implementation modes consistent with the disclosure. Instead, they are merely examples of an apparatus and method consistent with some aspects of the disclosure described in detail in the appended claims.

[0077] First of all, terms involved in the embodiments of the disclosure are explained.

[0078] First-stage training and second-stage training refer to that training of a data processing model is divided into first-stage training and second-stage training in a sequential order according to a total training count of the data processing model. For example, if the total training count of the data processing model is preset to be 300, the first 100 trainings may be called first-stage training, and the following 200 trainings are called second-stage training. There is no limitation on the training count in first-stage training and the training count in second-stage training.

[0079] A data augmentation operation refers to an operation of finely adjusting

training data to improve the data amount and diversity of the training data. Taking image data as an example, a size and color of the image data are adjusted.

[0080] A data augmentation strategy refers to a solution to the data augmentation of training data. The data augmentation strategy includes a data augmentation operation. For example, the data augmentation operation in the data augmentation strategy is horizontal image cropping corresponding to a cropping amplitude of $0.1 \times \text{width}$, namely a width involved in each horizontal image cropping is 10% of an original image width.

[0081] The deep learning technology is applied extensively to many fields with remarkable achievements. Taking the field of image vision as an example, the deep learning technology may accomplish image classification, target detection, image segmentation, human pose estimation, and other tasks. In order to accomplish these tasks well, data processing models using the deep learning technology usually needs to be trained based on a large amount of training data, otherwise a model obtained by training may have an over-fitting phenomenon. Therefore, data augmentation becomes a common method for improving the data amount and diversity of training data, and designing a proper data augmentation strategy becomes a key factor of improving a training effect of the data processing model.

[0082] Generally, a data augmentation strategy may be manually designed by a professional. However, this manner is relatively high in time cost and personnel cost and low in reusability of the data augmentation strategy, and thus is usually suitable for training a specific data processing model only. Compared with manually designing a data augmentation strategy by a professional, automatically generating a data augmentation strategy may not only improve the efficiency of the generation of the data augmentation strategy but also generate a better data augmentation strategy.

[0083] Generally, in the case that a data augmentation strategy is generated automatically, the data augmentation strategy may be optimized by a reinforcement learning algorithm according to a training effect of a data processing model. The inventor finds that, in this case, the whole training process of the data processing model needs to be repeated again and again, and in addition, training data is massive and wholly needs a relatively large amount of calculation and relatively long time, so the efficiency of the

generation of the data augmentation strategy is not so high.

[0084] In the deep learning technology, over-fitting of the data processing model usually occurs in a later training stage. Therefore, the inventor conjectures that the improvement of the training effect of the data processing model by data augmentation mainly occurs in the later training stage of the data processing model. In order to improve the efficiency of the generation of the data augmentation strategy and ensure a training effect of a data processing model trained based on the data augmentation strategy, the inventor deeply researched the model training process based on the data augmentation strategy to verify the above-mentioned conjecture.

[0085] For example, the data processing model is an image classification model, and a total training count of the image classification model is 300. In such case, the inventor obtains a result shown in FIG. 1. FIG. 1 shows a relationship between data augmentation and a training effect of the image classification model. The horizontal ordinate represents a data augmentation round count in 300 trainings of the image classification model, and the vertical ordinate represents a classification accuracy of the image classification model subjected to the 300 trainings. The dotted line is a relationship between data augmentation round counts in a later training stage and classification accuracies of the image classification model, and the solid line is a relationship between data augmentation round counts in an earlier training stage and classification accuracies of the image classification model.

[0086] The data augmentation round count in the later training stage is continuously calculated backwards from the last training of the image classification model. For example, if the data augmentation round count in the later training stage is 50, it indicates that data augmentation is performed in the last 50 trainings of the image classification model. The data augmentation round count in the earlier training stage is continuously calculated forwards from the first training of the image classification model. For example, if the data augmentation round count in the earlier training stage is 50, it indicates that data augmentation is performed in the first 50 trainings of the image classification model.

[0087] The following results may be obtained based on FIG. 1. First, the dotted line is always above the solid line in the case that the data augmentation round counts are the

same. Therefore, the classification accuracy of the image classification model obtained by performing data augmentation in the later training stage is higher than that of the image classification model obtained by performing data augmentation in the earlier training stage in the case that the data augmentation round counts are the same. Second, the dotted line is always on the left side of the solid line in the case that the classification accuracies of the image classification model are the same. Therefore, the data augmentation round count needed by performing data augmentation in the later training stage is smaller than that needed by performing data augmentation in the earlier training stage in the case that the classification accuracies of the image classification model are the same. It is to be noted that the first point on each of the solid line and the dotted line represents that the data augmentation round count is 0 and the last point on each of the solid line and the dotted line represents that the data augmentation round count is 300, so the four points are not considered in the above-mentioned comparison process.

[0088] Based on the above-mentioned findings of the inventor, a method for updating a data augmentation strategy according to the embodiments of the disclosure includes the following operations. An initial data augmentation strategy is acquired. Second-stage training is performed on a preset data processing model subjected to first-stage training according to the data augmentation strategy and preset training data. The data augmentation strategy is updated according to the data processing model subjected to second-stage training. Therefore, in the process of updating the data augmentation strategy, merely the second-stage training on the data processing model is needed, not only is the quality of the data augmentation strategy ensured, but also the efficiency of the generation of the data augmentation strategy is improved.

[0089] The method for updating a data augmentation strategy provided in the embodiments of the disclosure may be applied to a network architecture shown in FIG. 2. As shown in FIG. 2, the network architecture at least includes a terminal device 201 or a server 202. The terminal device 201 may store the data processing model subjected to first-stage training, perform second-stage training on the data processing module and update the data augmentation strategy. Alternatively, the server 202 may store the data processing model subjected to first-stage training, perform second-stage training on the data processing model and update the data augmentation strategy. Alternatively, the

terminal device 201 may store the data processing model subjected to first-stage training, and the server 202 may perform second-stage training on the data processing model and update the data augmentation strategy. Alternatively, the server 202 stores the data processing model subjected to first-stage training, and the terminal device 201 performs second-stage training on the data processing model and update the data augmentation strategy.

[0090] The terminal device may be a computer, a tablet computer, a smart phone, or another device. The server may be a single server or a server group consisting of multiple servers.

[0091] The technical solutions of the embodiments of the disclosure and how to solve the foregoing technical problems through the technical solutions of the disclosure will be described below in detail. The following embodiments may be combined. The same or similar concepts or processes will not be elaborated in some embodiments. The embodiments of the disclosure will be described below in combination with the drawings.

[0092] FIG. 3 is a flowchart of a method for updating a data augmentation strategy according to an embodiment of the disclosure. As shown in FIG. 3, the method includes the following operations.

[0093] In S301, an initial data augmentation strategy is acquired.

[0094] In a possible implementation mode, the initial data augmentation strategy may be acquired from preset data augmentation strategies in S301. In order to distinguish each preset data augmentation strategy from a data augmentation strategy presently in use, in terms of description, each preset data augmentation strategy is called a preset strategy, and the data augmentation strategy presently in use is called a data augmentation strategy.

[0095] In a possible implementation mode, as an alternative of acquiring the initial data augmentation strategy from preset strategies, a user may preset the initial data augmentation strategy, and the set data augmentation strategy is directly acquired. Alternatively, an initial data augmentation operation may be acquired from preset data augmentation operations to further obtain the initial data augmentation strategy.

- [0096] In S302, second-stage training is performed on a preset data processing model subjected to first-stage training according to the data augmentation strategy and preset training data.
- [0097] First-stage training may be performed in advance on the data processing model to obtain the data processing model subjected to first-stage training. The training data may be acquired in advance. The training data may be stored in form of a database.
- [0098] In a possible implementation mode, after the initial data augmentation strategy is obtained, data augmentation may be performed on the training data by using the data augmentation strategy. Second-stage training is performed on the data processing model subjected to first-stage training according to the training data subjected to data augmentation to obtain the data processing model subjected to second-stage training. Therefore, data augmentation is performed on the training data in later training of the data processing model by full utilization of the characteristic that data augmentation has a greater influence on the later training of the data processing model.
- [0099] There is no limitation on the training algorithm for the data processing model.
- [00100] In S303, the data augmentation strategy is updated according to the data processing model subjected to second-stage training.
- [00101] In a possible implementation mode, after the data processing model is subjected to first-stage training and second-stage training, a training process for the data processing model is completed, and a trained processing model is obtained. Therefore, a training effect of the data processing model subjected to second-stage training may be checked to obtain a checking result. For example, in the case that a task of the data processing model is an image classification task, the checking result of the data processing model is the image classification accuracy of the data processing model.
- [00102] In a possible implementation mode, after the checking result of the data processing model is obtained, the training effect of the data processing model obtained by training based on the training data subjected to data augmentation in the case that data augmentation is performed on the training data by using the data augmentation strategy may be obtained. It can thus be seen that the checking result of the data processing model

reflects the quality of the data augmentation strategy. For example, higher image classification accuracy of the data processing model indicates higher quality of the data augmentation strategy. Therefore, the data augmentation strategy may be updated according to the checking result of the data processing model. A preset strategy in a strategy updating space may be acquired as an updated data augmentation strategy in a process of updating the data augmentation strategy.

[00103] In the embodiment of the disclosure, by using the initial data augmentation strategy and the training data, second-stage training is performed on the data processing model subjected to first-stage training, and the data augmentation strategy is updated according to the data processing model subjected to second-stage training. The characteristic that the data augmentation strategy has a greater influence on the later training of the data processing model is fully utilized, the quality of the data augmentation strategy is ensured, and meanwhile, the efficiency of the generation of the data augmentation strategy is improved.

[00104] FIG. 4 is a flowchart of a method for updating a data augmentation strategy according to another embodiment of the disclosure. As shown in FIG. 4, the method includes the following operations.

[00105] In S401, an initial data augmentation strategy is acquired.

[00106] In a possible implementation mode, the initial data augmentation strategy may be acquired from preset strategies. As an alternative of acquiring the initial data augmentation strategy from the preset strategies, a user may preset the initial data augmentation strategy, and the set data augmentation strategy is directly acquired. Alternatively, an initial data augmentation operation may be acquired from preset data augmentation operations to further obtain the initial data augmentation strategy.

[00107] In a possible implementation mode, the data augmentation strategy includes multiple preset data augmentation operations to improve the quality of the data augmentation strategy. In the case that second-stage training is subsequently performed on the data processing model subjected to first-stage training, data augmentation may be sequentially performed on training data according to each of the data augmentation

operations in the data augmentation strategy, and second-stage training is performed on the data processing model subjected to first-stage training according to the training data subjected to data augmentation.

[00108] For example, the training data is image data. Data augmentation operations shown in Table 1 and operation amplitudes corresponding to the data augmentation operations may be preset. There are totally 14 types of data augmentation operations in FIG. 1. Three operation amplitudes are set for each of the 11 types of data augmentation operations respectively. No operation amplitudes need to be set for the other three types of data augmentation operations. The same type of data augmentation operations with different operation amplitudes may be considered as different data augmentation operations. Therefore, there are totally 36 data augmentation operations in Table 1. In the case that a data augmentation strategy includes two data augmentation operations, the data augmentation operations in Table 1 may be combined into 36×36 data augmentation strategies. Therefore, 36×36 preset strategies may be set according to Table 1.

[00109] Table 1 Image Data Augmentation Operations and Amplitudes of the Augmentation Operations

Data augmentation operation	Operation amplitude	Amplitude unit
Horizontal cropping	{0.1, 0.2, 0.3}	Width scale
Vertical cropping	{0.1, 0.2, 0.3}	Height scale
Horizontal translation	{0.15, 0.3, 0.45}	Width scale
Vertical translation	{0.15, 0.2, 0.45}	Height scale
Image rotation	{10, 20, 30}	Angle value
Color adjustment	{0.3, 0.6, 0.9}	Color balance
Posterization	{4.4, 5.6, 6.8}	Pixel bit number value

Daylight-based processing	{26, 102, 179}	Pixel threshold
Contrast adjustment	{1.3, 1.6, 1.9}	Contrast scale
Sharpness adjustment	{1.3, 1.6, 1.9}	Sharpness scale
Brightness adjustment	{1.3, 1.6, 1.9}	Brightness scale
Auto contrast	None	None
Equalization	None	None
Color inversion	None	None

[00110] In S402, second-stage training is performed on a preset data processing model subjected to first-stage training according to the data augmentation strategy and preset training data.

[00111] In a possible implementation mode, after the initial data augmentation strategy is obtained, data augmentation may be performed on the training data by using the data augmentation strategy. Second-stage training is performed on the data processing model subjected to first-stage training according to the training data subjected to data augmentation, to obtain the data processing model subjected to second-stage training. Therefore, data augmentation is performed on the training data in later training of the data processing model by full utilization of the characteristic that data augmentation has a greater influence on the later training of the data processing model.

[00112] In S403, the data augmentation strategy is updated according to the data processing model subjected to second-stage training.

[00113] In a possible implementation mode, after the data processing model is subjected to first-stage training and second-stage training, a training process thereof is completed to obtain a trained processing model. Therefore, a training effect of the data processing model subjected to second-stage training may be checked to obtain a checking

result.

[00114] In a possible implementation mode, after the checking result of the data processing model is obtained, the training effect of the data processing model obtained by training based on the training data subjected to data augmentation in the case that data augmentation is performed on the training data by using the data augmentation strategy may be obtained. It can thus be seen that the checking result of the data processing model reflects the quality of the data augmentation strategy presently in use. Therefore, the data augmentation strategy may be updated according to the checking result of the data processing model. In a process of updating the data augmentation strategy, a preset strategy in a strategy updating space may be acquired as an updated data augmentation strategy.

[00115] In S404, whether an updated data augmentation strategy satisfies a preset condition is determined.

[00116] In a possible implementation mode, S406 is executed in the case that the updated data augmentation strategy satisfies the preset condition. S405 is executed in the case that the updated data augmentation strategy does not satisfy the preset condition.

[00117] In S405, the initial data augmentation strategy is updated to the updated data augmentation strategy.

[00118] In a possible implementation mode, the initial data augmentation strategy is updated to the updated data augmentation strategy, namely the data augmentation strategy presently in use is updated to the updated data augmentation strategy, and S402 is re-executed to update the data augmentation strategy for multiple times.

[00119] In S406, a final data augmentation strategy is obtained.

[00120] In a possible implementation mode, in the case that the updated data augmentation strategy satisfies the preset condition, the updating of the data augmentation strategy is stopped, and a data augmentation strategy used when the checking result of the data processing model subjected to second-stage training is maximum among all updating processes is selected as the final data augmentation

strategy. Therefore, the quality of the data augmentation strategy is effectively improved.

[00121] In a possible implementation mode, a data augmentation strategy subjected to an M^{th} update is acquired, M being more than or equal to 1. Second-stage training is performed on the data processing model subjected to first-stage training according to the data augmentation strategy obtained by the M^{th} update and the training data. An $(M+1)^{\text{th}}$ update is performed on the data augmentation strategy according to the data processing model subjected to second-stage training.

[00122] In a possible implementation mode, whether the number of times for which the data augmentation strategy has been updated (update count) reaches a preset count threshold may be determined to determine whether the updated data augmentation strategy satisfies the preset condition. It is determined that the updated data augmentation strategy satisfies the preset condition in the case that the update count reaches the count threshold. It is determined that the updated data augmentation strategy does not satisfy the preset condition in the case that the update count does not reach the count threshold. Therefore, whether to continue updating the data augmentation strategy is controlled according to the update count to avoid the data augmentation strategy from being constantly updated.

[00123] In a possible implementation mode, whether to stop constantly updating the data augmentation strategy may be determined by determining whether the checking result of the data processing model subjected to second-stage training satisfies a preset condition, besides determining whether the number of times for which the data augmentation strategy has been updated reaches the preset count threshold.

[00124] The checking result of the data processing model may be compared with a preset checking threshold. In the case that the checking result of the data processing model is greater than the checking threshold, it is determined that the data processing model subjected to second-stage training satisfies the preset condition, and the data augmentation strategy is set as the final data augmentation strategy. In the case that the checking result of the data processing model is less than or equal to the checking threshold, it is determined that the data processing model subjected to second-stage training does not satisfy the preset condition, and the data augmentation strategy

continues to be updated.

[00125] In a possible implementation mode, multiple data augmentation strategies are updated in each updating process. The data augmentation strategies are concurrently updated. Therefore, the efficiency of the generation of the data augmentation strategy is effectively improved.

[00126] In a possible implementation mode, after every preset number of updates, an optimal data augmentation strategy is selected from the updated data augmentation strategies according to the data processing model subjected to second-stage training. Each of the updated data augmentation strategies other than the optimal strategy is replaced with the optimal data augmentation strategy. Therefore, the convergence of the updating process and the efficiency of the generation of the data augmentation strategy are improved. The optimal data augmentation strategy is selected in a selection process according to the checking result obtained by detecting the training effect of the data processing model subjected to second-stage training.

[00127] In a possible implementation mode, the training data is image data or text data. The data processing model is an image processing model in the case that the training data is image data. The data processing model is a natural language processing model in the case that the training data is text data. Therefore, the method for updating a data augmentation strategy provided in the embodiment of the disclosure may be applied to the generation of a data augmentation strategy in the field of image processing and the generation of a data augmentation strategy in the field of natural languages.

[00128] In the embodiment of the disclosure, second-stage training is performed on the data processing model subjected to first-stage training, by using the initial data augmentation strategy and the training data, and the data augmentation strategy is updated for multiple times according to the data processing model subjected to second-stage training. The characteristic that the data augmentation strategy has a greater influence on the later training of the data processing model is fully utilized, the quality of the data augmentation strategy is ensured, and meanwhile, the efficiency of the generation of the data augmentation strategy is improved.

[00129] FIG. 5 is a flowchart of a method for updating a data augmentation strategy according to another embodiment of the disclosure. As shown in FIG. 5, the method includes the following operations.

[00130] In S501, an initial data augmentation strategy is acquired.

[00131] In a possible implementation mode, the initial data augmentation strategy may be acquired from preset strategies. As an alternative of acquiring the initial data augmentation strategy from the preset strategies, the initial data augmentation strategy preset by a user may be directly acquired. Alternatively, an initial data augmentation operation may be acquired from multiple preset data augmentation operations, to further obtain the initial data augmentation strategy.

[00132] In a possible implementation mode, in the case that the initial data augmentation strategy is acquired from multiple preset strategies, one or more preset strategies are uniformly and randomly selected from the multiple preset strategies as initial data augmentation strategies. Therefore, the fairness of selecting the initial data augmentation strategy is improved. Uniformly and randomly selecting one or more preset strategies from multiple preset strategies indicates that each preset strategy may be selected at an equal probability.

[00133] In a possible implementation mode, if multiple preset strategies are uniformly and randomly selected from the multiple preset strategies as initial data augmentation strategies, it indicates that there are multiple initial data augmentation strategies. In such a case, the data augmentation strategies are synchronously updated in a subsequent updating process. Therefore, the efficiency of the generation of the data augmentation strategies is improved.

[00134] In S502, second-stage training is performed on a preset data processing model subjected to first-stage training according to the data augmentation strategy and preset training data.

[00135] In a possible implementation mode, data augmentation is performed on the training data by using a data augmentation operation in the data augmentation strategy. In the case that the data augmentation strategy includes multiple data augmentation

operations, data augmentation is sequentially performed on the training data by using each of the data augmentation operations in the data augmentation strategy to obtain training data subjected to data augmentation. Second-stage training is performed on the data processing model subjected to first-stage training according to the training data subjected to data augmentation, to obtain a data processing model subjected to second-stage training.

[00136] In S503, a preset strategy model is updated according to the data processing model subjected to second-stage training.

[00137] The strategy model is a parameterized model, and a parameter thereof is a preset strategy parameter. An output of the strategy model may be adjusted by a strategy adjustment operation. The output of the strategy model is a selection probability of each preset strategy, i.e., a probability that each preset strategy is selected as an updated data augmentation strategy in the case that the data augmentation strategies are updated. Therefore, the strategy model may be understood as a multinomial distribution.

[00138] In a possible implementation mode, preset verification data may be acquired. The verification data includes input data and label data corresponding to the input data. Taking image data as an example, in the case that the verification data is image data and a task of the data processing model is an image classification task, the verification data includes an input image and a classification label corresponding to the input image. The classification label is a class of input data.

[00139] In a possible implementation mode, the input data in the verification data is input to the data processing model subjected to second-stage training to obtain an output result of the data processing model. The output result of the data processing model is compared with the label data corresponding to the input data to obtain a checking result of checking the data processing model. Here, checking the data processing model refers to checking a training effect of the data processing model. Taking image data as an example, in the case that the verification data is image data and a task of the data processing model is an image classification task, an input image is input to the data processing model, and an output of the data processing model is compared with a classification label corresponding to the input image to obtain the classification accuracy

of the data processing model.

[00140] In a possible implementation mode, after the checking result of the data processing model is obtained, the strategy parameter of the strategy model may be updated according to the checking result to obtain an updated strategy model.

[00141] In S504, a selection probability of each preset strategy is determined by using the updated strategy model.

[00142] In S505, an updated data augmentation strategy is selected from the preset strategies according to the selection probabilities of the preset strategies.

[00143] In a possible implementation mode, the selection probability of each preset strategy may be redetermined according to the updated strategy model. One preset strategy is selected from the preset strategies as the updated data augmentation strategy according to the selection probabilities of the preset strategies.

[00144] In a possible implementation mode, the strategy parameter includes a weight corresponding to each preset strategy. The weight corresponding to each preset strategy is updated when the strategy parameter is updated. In the process of acquiring the initial data augmentation strategy, a same weight may be set for all the preset strategies to implement the uniform and random selection of the initial data augmentation strategy from the preset strategies. The selection probabilities of the preset strategies gradually differs from each other as the weights of preset strategies changes differently in the process of updating the strategy parameter. Therefore, the strategy parameter is adjusted according to the training effect of the data model subjected to second-stage training, then the selection probability of each preset strategy is redetermined according to the strategy model, and data augmentation strategies with higher quality are constantly selected from the preset strategies. As such, not only is the efficiency of the generation of the data augmentation strategy improved, but also the quality of the data augmentation strategy is ensured.

[00145] In a possible implementation mode, the strategy model may be represented as formula (1):

$$[00146] \quad p_{\theta}(O^{(k)}) = \frac{\frac{1}{1+e^{-\theta_k}}}{\sum_{i=1}^K (\frac{1}{1+e^{-\theta_i}})} \quad (1)$$

[00147] where e represents a base number of a natural logarithm. θ_k represents a k^{th} weight in the strategy parameter θ , i.e., a weight corresponding to a k^{th} preset strategy. K represents the total number of the preset strategies. $O^{(k)}$ represents the k^{th} preset strategy. $p_{\theta}(O^{(k)})$ represents a selection probability of the k^{th} preset strategy. Therefore, the selection probability of each preset strategy may be determined according to the strategy model and the strategy parameter including the weight corresponding to each preset strategy, and the strategy parameter may be adjusted to effectively adjust the selection probability of each preset strategy. Not only is the efficiency of the generation of the data augmentation strategy improved, but also the quality of the data augmentation strategy is ensured.

[00148] In a possible implementation mode, the updating of the strategy parameter may be represented as formula (2):

$$[00149] \quad \theta^* = \underset{\theta}{\operatorname{argmax}} ACC(\bar{\omega}_{\theta}^*, D_{val}) \quad (2)$$

[00150] $ACC(\bar{\omega}_{\theta}^*, D_{val})$ represents the checking result of the data processing model subjected to second-stage training. $\bar{\omega}_{\theta}^*$ represents a model parameter of the data processing model subjected to second-stage training. D_{val} represents the verification data.

[00151] In a possible implementation mode, in a process of updating the strategy parameter according to the checking result of the data processing model subjected to second-stage training, the strategy parameter may be updated by a preset heuristic search algorithm, to improve an updating effect of the strategy parameter.

[00152] In a possible implementation mode, the updating of the strategy parameter may be represented as formula (3) in the case that the heuristic search algorithm configured to update the strategy parameter is a reinforcement learning algorithm:

$$[00153] \quad \nabla_{\theta} J(\theta, \bar{\omega}_{\theta}^*) = \nabla_{\theta} E[ACC(\bar{\omega}_{\theta}^*, D_{val})] \approx \frac{1}{N} \sum_1^N \nabla_{\theta} \log(p(T_n)) ACC(\bar{\omega}_{\theta}^*, D_{val}) \quad (3).$$

[00154] where $\nabla_{\theta}J(\theta, \bar{\omega}_{\theta}^*)$ represents a gradient value of the strategy parameter. T_n represents an n^{th} search trajectory in the reinforcement learning algorithm. $p(T_n)$ represents a probability that search trajectory T_n in the reinforcement learning algorithm is found. N represents the number of search trajectories in the reinforcement learning algorithm. $E[ACC(\bar{\omega}_{\theta}^*, D_{val})]$ represents an expected value of the checking result of the data processing model subjected to second-stage training.

[00155] In the process of updating the strategy parameter by the reinforcement learning algorithm, the gradient value $\nabla_{\theta}J(\theta, \bar{\omega}_{\theta}^*)$ of the strategy parameter may be multiplied by a preset learning rate in the reinforcement learning algorithm to obtain a product, and then the product and the strategy parameter are added to obtain an updated strategy parameter. For example, in the case that an adaptive moment estimation (Adam) algorithm is used as the reinforcement learning algorithm, a learning rate of the Adam may be set such that $\eta_{\theta}=0.1$, $\beta_1=0.5$ and $\beta_2=0.999$.

[00156] In a possible implementation mode, in a process of performing second-stage training on the data processing model subjected to first-stage training according to the data augmentation strategy and the training data, the model parameter of the data processing model subjected to second-stage training may be represented as formula (4):

$$[00157] \quad \bar{\omega}_{\theta}^* = \operatorname{argmin}_{\omega} \frac{1}{Z} \sum_{(x,y) \in D_{tr}} E_{O \sim p_{\theta}} L(M_{\omega}(O(x)), y) \quad (4)$$

[00158] x represents input data in the training data. y represents label data corresponding to x in the training data. $O(x)$ represents performing data augmentation on x . $L(\cdot)$ represents a preset loss function. $E_{O \sim p_{\theta}}$ represents selecting a data augmentation strategy from the preset strategies in compliance with a probability distribution obtained by the strategy model. D_{tr} represents the training data. Z represents the amount of the input data x in the training data.

[00159] In a possible implementation mode, in a process of performing first-stage training in advance on the data processing model, a data augmentation strategy in first-stage training is uniformly and randomly selected from the preset strategies, data augmentation is performed on the training data according to the data augmentation

strategy in first-stage training, and first-stage training is performed on the data processing model according to the training data subjected to data augmentation. Data augmentation is also performed on the training data in first-stage training, so that the training effect of the data processing model subjected to first-stage training is improved.

[00160] In a possible implementation mode, the model parameter of the data processing model obtained by first-stage training may be represented as formula (5):

$$[00161] \quad \omega_{share} = \underset{\omega}{\operatorname{argmin}} \frac{1}{Z} \sum_{(x,y) \in D_{tr}} E_{O \sim U} L(M_{\omega}(O(x)), y) \quad (5)$$

[00162] where ω_{share} represents the model parameter of the data processing model obtained by first-stage training. $E_{O \sim p_{\theta}}$ represents selecting a data augmentation strategy from the preset strategies in compliance with a uniform probability distribution.

[00163] In the embodiment of the disclosure, by full use of the characteristic that the data augmentation strategy has a greater influence on the later training of the data processing model, second-stage training is performed on the data processing model subjected to first-stage training according to the data augmentation strategy and the training data, the strategy model is updated according to the data processing model subjected to second-stage training, the selection probability of each preset strategy is determined according to the updated strategy model, and the probability of each preset strategy is adjusted to improve the quality of the updated data augmentation strategy. Therefore, not only is the quality of the data augmentation strategy improved, but also the efficiency of the generation of the data augmentation strategy is improved.

[00164] FIG. 6 is a flowchart of a method for updating a data augmentation strategy according to another embodiment of the disclosure. As shown in FIG. 6, the method includes the following operations.

[00165] In S601, an initial data augmentation strategy is acquired.

[00166] In S602, second-stage training is performed on a preset data processing model subjected to first-stage training according to the data augmentation strategy and preset training data.

[00167] In S603, a preset strategy model is updated according to the data processing model subjected to second-stage training.

[00168] In a possible implementation mode, in a process of updating the strategy model according to the data processing model subjected to second-stage training, a training effect of the data processing model subjected to second-stage training is checked with verification data to obtain a checking result, and checking results of the data processing model subjected to second-stage training in first N-1 updates of the data augmentation strategy are acquired. For the clarity of description, the checking results of the data processing model subjected to second-stage training in the first N-1 updates of the data augmentation strategy are called historical checking results. The checking result and the historical checking results may be combined to update the strategy model to ensure the updating stability of the strategy model in each updating process and further improve an updating effect of the data augmentation strategy. N is a present total update count of the data augmentation strategy. The Nth update refers to a present updating process.

[00169] In a possible implementation mode, in a process of combining the checking result and the historical checking results to update the strategy model, an average value of the historical checking results may be determined, a difference between the checking result and the average value is determined, and the strategy parameter in the strategy model is updated according to the difference. Therefore, the updating stability of the strategy model in each updating process is ensured, and the updating effect of the data augmentation strategy is further improved. A heuristic search algorithm may be used for the updating process of the strategy parameter. Elaborations are omitted.

[00170] In S604, a selection probability of each of preset strategies is determined by using the updated strategy model.

[00171] In S605, an updated data augmentation strategy is selected from the preset strategies according to the selection probabilities of the preset strategies.

[00172] In a possible implementation mode, S601 to S605 may refer to the detailed descriptions about S501 to S505, and will not be elaborated herein.

[00173] In S606, whether the updated data augmentation strategy satisfies a preset

condition is determined.

[00174] In a possible implementation mode, S608 is executed in the case that the updated data augmentation strategy satisfies the preset condition. S607 is executed in the case that the updated data augmentation strategy does not satisfy the preset condition.

[00175] In S607, the initial data augmentation strategy is updated to the updated data augmentation strategy.

[00176] In a possible implementation mode, the initial data augmentation strategy is updated to the updated data augmentation strategy, and S602 is re-executed to update the data augmentation strategy for multiple times to improve the quality of the data augmentation strategy.

[00177] In S608, a final data augmentation strategy is obtained.

[00178] In a possible implementation mode, the updated data augmentation strategy is set as the final data augmentation strategy.

[00179] In a possible implementation mode, it may be determined whether a number of times for which the data augmentation strategy has been updated (update count) reaches a preset count threshold to determine whether the updated data augmentation strategy satisfies the preset condition. It is determined that the updated data augmentation strategy satisfies the preset condition in the case that the update count reaches the count threshold. It is determined that the updated data augmentation strategy does not satisfy the preset condition in the case that the update count does not reach the count threshold. Therefore, whether to continue updating the data augmentation strategy is controlled according to the update count, to avoid the data augmentation strategy from being constantly updated.

[00180] In a possible implementation mode, whether to stop constantly updating the data augmentation strategy may be determined by determining whether the checking result of the data processing model subjected to second-stage training satisfies a preset condition, besides determining whether the number of times for which the data augmentation strategy has been updated (update count) reaches the preset count threshold.

[00181] In a possible implementation mode, the checking result of the data processing model may be compared with a preset checking threshold. In the case that the checking result of the data processing model is greater than the checking threshold, it indicates that the data processing model subjected to second-stage training satisfies the preset condition, and the data augmentation strategy is set as the final data augmentation strategy. In the case that the checking result of the data processing model is less than or equal to the checking threshold, the data augmentation strategy continues to be updated.

[00182] In a possible implementation mode, after every preset number of updates, an optimal data augmentation strategy is selected from the updated data augmentation strategies according to the data processing model subjected to second-stage training. Each of the updated data augmentation strategies other than the optimal strategy is replaced with the optimal data augmentation strategy. Therefore, the convergence of the updating process and the efficiency of the generation of the data augmentation strategy are improved. The optimal data augmentation strategy may be selected in a selection process according to the checking result obtained by detecting the training effect of the data processing model subjected to second-stage training.

[00183] For example, FIG. 7 provides a process of concurrently updating multiple data augmentation strategies. As shown in FIG. 7, each cuboid represents a data augmentation strategy. Each cube represents a data processing model. Accuracy (ACC) represents a checking result of a data processing model obtained by second-stage training. Each row represents an updating process of a data augmentation strategy. Each column represents an update of each data augmentation strategy.

[00184] As shown in FIG. 7, an initial data augmentation strategy may be uniformly and randomly selected from preset strategies. The initial data augmentation strategy is copied for multiple times to obtain multiple identical initial data augmentation strategies. The multiple data augmentation strategies are concurrently updated. After every preset number of updates, an optimal data augmentation strategy is selected from updated data augmentation strategies. The optimal data augmentation strategy is copied, like strategy copying shown by the dotted arrowhead. Here, strategy copying means replacing the other data augmentation strategies in each of the updated data augmentation strategies

other than the optimal data augmentation strategy with the optimal data augmentation strategy. Therefore, the convergence of the data augmentation strategy after multiple updates may be effectively improved, and a data augmentation strategy with relatively high quality may be obtained.

[00185] As shown in FIG. 7, in a single updating process, a model parameter ω_{share} of a data processing model subjected to first-stage training is loaded to a data processing model to obtain the data processing model subjected to first-stage training. Second-stage training is performed on the data processing model subjected to first-stage training according to a data augmentation strategy and training data. Then, ACC, i.e., a checking result of the data processing model subjected to second-stage training, is obtained by checking with verification data. The data augmentation strategy is updated based on the checking result to obtain an updated data augmentation strategy.

[00186] It can be seen with reference to FIG. 7 that, in the embodiment of the disclosure, multiple data augmentation strategies may be concurrently updated. In each updating process of the data augmentation strategy, merely the second-stage training on the data processing model is needed. After every preset number of updates, each updated data augmentation strategy is replaced with a present optimal data augmentation strategy. The amount of calculation for updating the strategy parameter is small. Therefore, the efficiency of updating the data augmentation strategy is effectively improved, the efficiency of the generation of the data augmentation strategy is improved, and the quality of the data augmentation strategy is ensured.

[00187] In an embodiment, a proportion of a training count of the first-stage training to a total training count or a proportion of a training count of the second-stage training to the total training count may be adjusted to improve the efficiency of the generation of the data augmentation strategy.

[00188] FIG. 8 is a flowchart of a data processing method according to an embodiment of the disclosure. As shown in FIG. 8, the method includes the following operations.

[00189] In S801, data to be processed is acquired.

[00190] Data to be processed input by a user may be acquired. Alternatively, the data to

be processed may be acquired in advance.

[00191] In S802, the data to be processed is processed by using a pre-trained data processing model. The data processing model is sequentially subjected to first-stage training and second-stage training. The data processing model is trained by a preset data augmentation strategy and preset training data in a second training stage.

[00192] The data processing model is pre-trained. First-stage training is performed on the data processing model at first in a training process of the data processing model. Then, second-stage training is performed on the data processing model according to the data augmentation strategy and the training data. Therefore, the characteristic that data augmentation has a greater influence on the later training of the data processing model is fully utilized, and the data processing effect and model training efficiency of the data processing model are improved.

[00193] In a possible implementation mode, the data to be processed is input to the data processing model, and the data processing model processes the data to be processed to obtain a corresponding processing result.

[00194] In a possible implementation mode, the data augmentation strategy used for second-stage training of the data processing model may be obtained by the method for updating a data augmentation strategy provided in any above-mentioned embodiment. Therefore, the quality and generation efficiency of the data augmentation strategy are improved, and the data processing effect and model training efficiency of the data processing model are further improved.

[00195] In a possible implementation mode, first-stage training may be performed at first on the data processing model according to the training data to obtain a data processing model subjected to first-stage training in the training process of the data processing model. Then, data augmentation is performed on the training data by using the data augmentation strategy. Second-stage training is performed on the data processing model subjected to first-stage training based on the training data subjected to data augmentation to obtain the trained data processing model. Therefore, the characteristic that data augmentation has a greater influence on the later training of the data processing

model is fully utilized, and the data processing effect and model training efficiency of the data processing model are improved.

[00196] In a possible implementation mode, a data augmentation strategy may be uniformly and randomly selected from the preset strategies as the data augmentation strategy for first-stage training in a process of performing first-stage training on the data processing model. Data augmentation is performed on the training data by using the selected data augmentation. First-stage training is performed on the data processing model according to the training data subjected to data augmentation. The data augmentation strategy is uniformly and randomly selected to improve the training effect of first-stage training over the data processing model and further improve the overall training effect of the data processing model without prolonging time spent on model training.

[00197] In a possible implementation mode, the data to be processed and the training data may be image data or text data. The data to be processed and training data are image data in the case that the data processing model is an image processing model. The data to be processed and the training data are text data in the case that the data processing model is a natural language processing model. Therefore, an image processing effect or a natural language processing effect is improved.

[00198] In the embodiment of the disclosure, the data to be processed is processed by the pre-trained data processing model. The training process of the data processing model is divided into first-stage training and second-stage training. A preset data augmentation strategy is used in a second-stage training process. Therefore, the data processing effect and model training efficiency of the data processing model are improved, and the data processing effect is further improved.

[00199] An exemplary application of the embodiment of the disclosure in a practical application scene will now be described.

[00200] Automatic machine learning is a hot spot in the field of machine learning at present, and a related technology thereof may improve model performance and reduce manpower needed by tuning in many fields. An image data augmentation technology has

also been applied extensively to the field of image processing. An automatic image data augmentation process based on the automatic machine learning technology may improve the pertinence of data augmentation also reduces unnecessary manual adjustment. However, it is relatively complex to find a proper augmentation strategy based on a dataset of a specific task. This is because a dataset usually includes a relatively large amount of data, the overhead of direct seeking is unacceptable, and if a universal strategy is sought and used for all tasks, the model may not be improved greatly. Some of existing automatic data augmentation search technologies still cost much, while some have non-ideal improvement effects. The most common task in a partial or whole automatic machine learning process is parameter adjustment for automatic machine learning, e.g., automatically seeking a proper model structure, a proper data augmentation strategy, a proper loss function and a proper optimizer.

[00201] According to the method for updating a data augmentation strategy provided in the embodiments of the disclosure, the time consumption and the evaluation accuracy may be balanced well, namely a dataset of a normal scale may be directly searched and a stable improvement may be achieved. In addition, the method is applicable to multiple image classification datasets and transferrable to some extents. Moreover, the method may be easily embedded into each image classification task.

[00202] In a possible implementation mode, the method for updating a data augmentation strategy includes the search of an image data augmentation strategy. A search process may be divided into the following three operations. First, earlier-stage training is performed on the model in a uniform random strategy. Then, a One-shot (search strategy) search stage is entered, namely an ending state of earlier-stage training is repeatedly loaded, later-stage training is performed, and meanwhile, searching is performed. A search target is optimal performance achieved by later-stage training. Finally, a found strategy is applied to an original task for overall re-training to obtain final model performance. As a search strategy, One-Shot originally means using one "path" each time in a whole search space, and may also be broadly understood as multiple repeated single-sample updates. Proportions of earlier- and later-stage training in this method may be reasonably adjusted to greatly improve the time efficiency of search. Moreover, it is noticed by the inventor that later-stage training is more sensitive

to the influence of data augmentation, so the stability of an evaluation index is found undamaged in an experiment. By the method, the performance of each image classification model under a given dataset may be improved to help the model to achieve higher performance in multiple task scenes.

[00203] The method for updating a data augmentation strategy provided in the embodiments of the disclosure mainly includes the following steps.

[00204] In A, earlier-stage training is performed by uniform and random data augmentation.

[00205] In a possible implementation mode, A includes the following operations. An untrained initial model is acquired. Training is performed under uniform and random data augmentation. A model subjected to earlier-stage training is obtained. An input of A is a specified image classification dataset and the completely untrained model, while an output is the model subjected to earlier-stage training.

[00206] In a possible implementation mode, A includes the following operations.

[00207] The untrained initial model is used as a starting point. In an experiment, multiple models may be selected for independent experiments respectively.

[00208] Various types of data augmentation may be performed on an image at an equal property in an earlier-stage training process. Practical experimental observations show that, compared with not performing data augmentation, performing uniform data augmentation may achieve a better effect. A selected data augmentation operation may be various automatic data augmentation operations, to ensure the fairness. An operation list is shown in Table 1. The second column shows different amplitude values of each operation. Considering differences between amplitude values, there are totally 36 possible data augmentation operations. Two operations may be uniformly and randomly used for each image during training. An image obtained by the data augmentation operations is used as a practical input of the model.

[00209] The model subjected to earlier-stage training is saved for use in later-stage training.

[00210] In B, One-Shot search is performed, namely later-stage training is repeatedly performed, and a data augmentation strategy is constantly updated.

[00211] In a possible implementation mode, details about repeated training may refer to FIG. 7. As shown in FIG. 7, each cuboid represents a data augmentation strategy. Each cube represents a data processing model. ACC represents a checking result of a data processing model obtained by second-stage training. Each row represents an updating process of a data augmentation strategy. Each column represents an update of each data augmentation strategy.

[00212] A single updating process may include the loading of the model subjected to earlier-stage training. That is, a model parameter may be reset to a parameter obtained by earlier-stage training after each later-stage training. A present strategy is used to control data augmentation for later-stage training. The present strategy is a parameterized model, and a probability of each data augmentation operation may be derived from a parameter thereof. Two data augmentation operations may be performed on each image, so there are totally $36 \times 36 = 1,296$ augmentation methods in consideration of a sequential relationship. It is to be noted that the strategy may be not reset but constantly updated after each later-stage training till the end of the whole search period. A model subjected to later-stage training is obtained. In such case, the model is evaluated. Image classification is selected as a practical task, so the evaluation index is classification accuracy. In order to improve the stability and relativity of the evaluation index, a historical exponential moving average value may be subtracted for each evaluation. The strategy is updated using the present evaluation index of the model. Here, reinforcement learning is used for updating. A purpose of updating is to improve the evaluation index of the model.

[00213] A final strategy may be obtained by a plurality of repeated trainings and updates. A short script may be derived from the final strategy to be conveniently added to an expected training process.

[00214] In C, retraining is performed using a final strategy to obtain a final model and final performance. Data augmentation may be performed on each image under the control of the final strategy (under a corresponding probability value) in C. The final model and performance are obtained after C is completed.

[00215] According to the method for updating a data augmentation strategy provided in the embodiments of the disclosure, the One-Shot idea is used, so that the search efficiency and the evaluation accuracy are balanced well, and a better experimental effect is achieved under the same condition. In addition, a search result of the algorithm may be easily derived and flexibly used for other tasks.

[00216] According to the method for updating a data augmentation strategy provided in the embodiments of the disclosure, data augmentation may be directly performed in a training process of an image classification task or another image processing task to expect to achieve better performance and higher generalization performance. Data augmentation strategy search may be implemented under a specified dataset and a specified model to obtain a highly customized data augmentation strategy. Data augmentation strategy search for more tasks may be implemented in combination with a self-defined search space, e.g., in the field of natural language processing.

[00217] FIG. 9 is a structure diagram of an apparatus for updating a data augmentation strategy according to an embodiment of the disclosure. As shown in FIG. 9, the apparatus includes an acquisition part 901, a training part 902 and an updating part 903.

[00218] The acquisition part 901 is configured to acquire an initial data augmentation strategy.

[00219] The training part 902 is configured to perform second-stage training on a preset data processing model subjected to first-stage training according to the data augmentation strategy and preset training data.

[00220] The updating part 903 is configured to update the data augmentation strategy according to the data processing model subjected to second-stage training to obtain an updated data augmentation strategy.

[00221] In a possible implementation mode, the updating part 903 is further configured to:

[00222] update the initial data augmentation strategy to the updated data augmentation strategy to update the data augmentation strategy for multiple times.

- [00223] In a possible implementation mode, there are multiple initial data augmentation strategies, and the data augmentation strategies are concurrently updated. The updating part 903 is further configured to:
- [00224] after every preset number of updates, select an optimal data augmentation strategy from the updated data augmentation strategies according to the data processing model subjected to second-stage training; and
- [00225] replace each of the updated data augmentation strategies other than the optimal data augmentation strategy with the optimal data augmentation strategy.
- [00226] In a possible implementation mode, the data augmentation strategy includes multiple preset data augmentation operations. The training part 902 is further configured to:
- [00227] sequentially perform data augmentation on the training data according to the data augmentation operations; and
- [00228] perform second-stage training on the data processing model subjected to first-stage training according to the training data subjected to data augmentation.
- [00229] In a possible implementation mode, the training data is image data or text data.
- [00230] In a possible implementation mode, the updating part 903 is further configured to:
- [00231] update a preset strategy model according to the data processing model subjected to second-stage training;
- [00232] determine a selection probability of each preset strategy by using the updated strategy model; and
- [00233] select an updated data augmentation strategy from the preset strategies according to the selection probabilities of preset strategies.
- [00234] In a possible implementation mode, in the case that the data augmentation strategy has been updated for multiple times, the updating part 903 is further configured

to:

- [00235] check the data processing model subjected to second-stage training according to preset verification data to obtain a checking result;
- [00236] acquire historical checking results of the data processing model subjected to second-stage training in first N-1 updates of the data augmentation strategy, N being a present total number of times for which the data augmentation strategy has been updated; and
- [00237] update the strategy model according to the historical checking results and the checking result.
- [00238] In a possible implementation mode, the updating part 903 is further configured to:
- [00239] determine an average value of the historical checking results;
- [00240] determine a difference between the checking result and the average value; and
- [00241] update a strategy parameter in the strategy model according to the difference.
- [00242] In a possible implementation mode, the training part 902 is further configured to:
- [00243] uniformly and randomly select a data augmentation strategy in first-stage training selected from preset strategies; and
- [00244] perform first-stage training on the data processing model according to the data augmentation strategy in first-stage training and the training data.
- [00245] The apparatus for updating a data augmentation strategy provided in FIG. 9 may execute the corresponding method embodiment, and an implementation principle and technical effects thereof are similar and will not be elaborated herein.
- [00246] FIG. 10 is a structure diagram of a data processing apparatus according to an embodiment of the disclosure. As shown in FIG. 10, the apparatus includes an

acquisition part 1001 and a processing part 1002.

[00247] The acquisition part 1001 is configured to acquire data to be processed.

[00248] The processing part 1002 is configured to process the data to be processed by using a pre-trained data processing model. The data processing model is sequentially subjected to first-stage training and second-stage training. The data processing model is trained by using a preset data augmentation strategy and preset training data in a second training stage.

[00249] In a possible implementation mode, the data augmentation strategy is generated by the method for updating a data augmentation strategy as described in any above-mentioned embodiment.

[00250] In a possible implementation mode, the apparatus further includes a training part, further configured to:

[00251] perform first-stage training on the data processing model according to the training data;

[00252] perform data augmentation on the training data by using the data augmentation strategy; and

[00253] perform second-stage training on the data processing model subjected to first-stage training according to the training data subjected to data augmentation.

[00254] In a possible implementation mode, the training part is further configured to:

[00255] uniformly and randomly select a data augmentation strategy in the first-stage training from preset strategies; and

[00256] perform first-stage training on the data processing model according to the data augmentation strategy and the training data in the first-stage training.

[00257] In a possible implementation mode, the data to be processed and the training data are image data or text data.

[00258] The data processing apparatus provided in FIG. 10 may execute the corresponding method embodiment, and an implementation principle and technical effects thereof are similar and will not be elaborated herein.

[00259] FIG. 11 is a structure diagram of an electronic device according to an embodiment of the disclosure. As shown in FIG. 11, the terminal device may include a processor 1101 and a memory 1102. The memory 1102 is configured to store computer-executable instructions. The processor 1101, when executing a computer program, implements the method in any above-mentioned embodiment.

[00260] The processor 1101 may be a universal processor, including a Central Processing Unit (CPU), a Network Processor (NP), etc., or may be a Digital Signal Processor (DSP), an Application Specific Integrated Circuit (ASIC), a Field-Programmable Gate Array (FPGA) or another programmable logic device, a discrete gate or transistor logic device, and a discrete hardware component. The memory 1102 may include a high-speed Random Access Memory (RAM), or may include a non-volatile memory, for example, at least one disk memory.

[00261] An embodiment of the disclosure also provides a computer-readable storage medium having stored therein an instruction which, when running in a computer, enables the computer to execute the method in any above-mentioned embodiment.

[00262] An embodiment of the disclosure also provides a program product, which includes a computer program stored in a storage medium. At least one processor may read the computer program from the storage medium. The at least one processor, when executing the computer program, may implement the method in any above-mentioned embodiment.

[00263] FIG. 12 is a block diagram of an apparatus for updating a data augmentation strategy 1200 according to the present embodiment. For example, the apparatus 1200 may be provided as a server or a computer. Referring to Fig. 12, the apparatus 1200 includes a processing component 1201, further including one or more processors, and a memory resource represented by a memory 1202, configured to store instructions executable by the processing component 1201, for example, an application program. The

application program stored in the memory 1202 may include one or more than one part, each corresponding to a set of instructions. In addition, the processing component 1201 is configured to execute the instructions to execute the method in any embodiment shown in FIGS. 3 to 6.

[00264] The apparatus 1200 may further include a power component 1203 configured to execute power management of the apparatus 1200, a wired or wireless network interface 1204 configured to connect the apparatus 1200 to a network, and an Input/Output (I/O) interface 1205. The apparatus 1200 may be operated on an operating system stored in the memory 1202, for example, Windows Server™, Mac OS X™, Unix™, Linux™, FreeBSD™, or the like.

[00265] In the embodiments of the disclosure, "at least one" refers to one or more, and "multiple" refers to two or more than two. "And/or" describes an association relationship between associated objects and represents that three relationships may exist. For example, A and/or B may represent three conditions: i.e., independent existence of A, existence of both A and B, and independent existence of B. A and B may be singular or plural. Character "/" usually represents that previous and next associated objects are in an "or" relationship. Character "/" in the formula represents that previous and next associated objects are in a "divided" relationship. "At least one (item) of the followings" or a similar expression refers to any combination in these options, including any combination of the options in a singular or plural form. For example, at least one (item) in a, b, or c may represent a, b, c, a-b, a-c, b-c, or a-b-c, a, b, and c being singular or plural.

[00266] It can be understood that various numerals involved in the embodiments of the disclosure are only for distinguishing for the ease of description and not intended to limit the scope of the embodiments of the disclosure.

[00267] It can be understood that, in the embodiments of the disclosure, the magnitude of the sequence number of each process does not mean an execution sequence and the execution sequence of each process should be determined by its function and an internal logic and should not form any limit to an implementation process of the embodiments of the disclosure.

[00268] Other implementation solutions of the disclosure are apparent to those skilled in the art upon considering the specification and practicing the disclosure disclosed here. The embodiments of the disclosure are intended to cover any transformations, uses or adaptive variations of the disclosure, and these transformations, uses or adaptive variations follow the general principle of the disclosure and include common general knowledge or conventional technical means undisclosed in the disclosure in the technical field. The specification and the embodiments are only considered as examples, and the practical scope and spirit of the disclosure are pointed out by the appended claims.

[00269] It should be understood that the disclosure is not limited to the precise structures described above and shown in the drawings, and various modifications and variations may be made without departing from the scope thereof. The scope of the disclosure is only defined by the appended claims.

INDUSTRIAL APPLICABILITY

[00270] According to the embodiments of the disclosure, second-stage training is performed on the data processing model subjected to first-stage training according to the initial data augmentation strategy and the training data, and the data augmentation strategy is updated according to the data processing model subjected to second-stage training. Therefore, the characteristic that the data augmentation strategy has a greater influence on the later training of the data processing model may be fully utilized, the quality of the data augmentation strategy is ensured, and meanwhile, the efficiency of the generation of the data augmentation strategy is improved.

CLAIMS

1. A method for updating a data augmentation strategy, comprising:

acquiring an initial data augmentation strategy;

performing second-stage training on a preset data processing model subjected to first-stage training according to the data augmentation strategy and preset training data; and

updating, according to the data processing model subjected to second-stage training, the data augmentation strategy to obtain an updated data augmentation strategy.

2. The method of claim 1, further comprising:

acquiring a data augmentation strategy subjected to M updates, M being more than or equal to 1;

performing second-stage training on the data processing model subjected to first-stage training according to the data augmentation strategy subjected to M updates and the training data; and

performing an $(M+1)^{\text{th}}$ update on the data augmentation strategy according to the data processing model subjected to second-stage training.

3. The method of claim 2, wherein the initial data augmentation strategy comprises multiple initial data augmentation strategies, which are concurrently updated, and the method further comprises:

after every preset number of updates, selecting an optimal data augmentation strategy from the updated multiple data augmentation strategies according to the data processing model subjected to second-stage training; and

replacing each of the updated data augmentation strategies other than the optimal data augmentation strategy with the optimal data augmentation strategy.

4. The method of any one of claims 1-3, wherein the data augmentation strategy comprises multiple preset data augmentation operations; and the performing second-stage training on the preset data processing model subjected to first-stage training according to the data augmentation strategy and preset training data comprises:

sequentially performing data augmentation on the training data according to each of the data augmentation operations, and

performing second-stage training on the data processing model subjected to first-stage training according to the training data subjected to data augmentation.

5. The method of any one of claims 1-3, wherein the updating the data augmentation strategy according to the data processing model subjected to second-stage training comprises:

updating a preset strategy model according to the data processing model subjected to second-stage training;

determining a selection probability of each of preset strategies by using the updated strategy model; and

selecting the updated data augmentation strategy from the preset strategies according to the selection probabilities of the preset strategies.

6. The method of claim 5, wherein, in a case that the data augmentation strategy is updated for multiple times, the updating the preset strategy model according to the data processing model subjected to second-stage training comprises:

checking the data processing model subjected to second-stage training according to

preset verification data, to obtain a checking result;

acquiring historical checking results of the data processing model subjected to second-stage training in first N-1 updates of the data augmentation strategy, N being a present total number of times for which the data augmentation strategy has been updated; and

updating the strategy model according to the historical checking results and the checking result.

7. The method of claim 6, wherein the updating the strategy model according to the historical checking results and the checking result comprises:

determining an average value of the historical checking results;

determining a difference between the checking result and the average value; and

updating a strategy parameter in the strategy model according to the difference.

8. The method of any one of claims 1-3, wherein before the acquiring the initial data augmentation strategy, the method further comprises:

uniformly and randomly selecting a data augmentation strategy in the first-stage training from preset strategies; and

performing first-stage training on the data processing model according to the data augmentation strategy in the first-stage training and the training data.

9. A data processing method, comprising:

acquiring data to be processed; and

processing, by using a pre-trained data processing model, the data to be processed, the

data processing model being sequentially subjected to first-stage training and second-stage training, the data processing model being trained in the second-stage training by using a preset data augmentation strategy and preset training data, and the data augmentation strategy being generated by the method for updating the data augmentation strategy of any one of claims 1-8.

10. The method of claim 9, further comprising:

performing first-stage training on the data processing model according to the training data;

performing data augmentation on the training data by using the data augmentation strategy; and

performing second-stage training on the data processing model subjected to first-stage training according to the training data subjected to data augmentation.

11. The method of claim 10, wherein the performing first-stage training on the data processing model according to the training data comprises:

uniformly and randomly selecting a data augmentation strategy in first-stage training from preset strategies; and

performing first-stage training on the data processing model according to the data augmentation strategy in first-stage training and the training data.

12. The method of any one of claims 9-11, wherein the data to be processed and the training data are image data or text data.

13. An apparatus for updating a data augmentation strategy, comprising:

an acquisition part, configured to acquire an initial data augmentation strategy;

a training part, configured to perform second-stage training on a preset data processing model subjected to first-stage training according to the data augmentation strategy and preset training data; and

an updating part, configured to update the data augmentation strategy according to the data processing model subjected to second-stage training, to obtain an updated data augmentation strategy.

14. A data processing apparatus, comprising:

an acquisition part, configured to acquire data to be processed; and

a processing part, configured to process the data to be processed by using a pre-trained data processing model, the data processing model being sequentially subjected to first-stage training and second-stage training, the data processing model being trained by using a preset data augmentation strategy and preset training data in a second training stage, and the data augmentation strategy being generated by the method for updating the data augmentation strategy of any one of claims 1-8.

15. An electronic device, comprising a memory and a processor, wherein

the memory is configured to store program instructions; and

the processor is configured to call the program instructions in the memory to execute the method of any one of claims 1-8 or the method of any one of claims 9-12.

16. A computer-readable storage medium having stored therein a computer program which, when executed, implements the method of any one of claims 1-8 or the method of any one of claims 9-12.

17. A computer program, comprising a computer-readable code, which, when running in an electronic device, causes a processor in the electronic device to execute the method of any one of claims 1-8 or the method of any one of 9-12.