IMAGE DETECTION METHOD, AND ELECTRONIC DEVICE AND STORAGE MEDIUM

[0001] The present disclosure claims the priority to Chinese Patent Application No.202311813188.7, titled "IMAGE DETECTION METHOD, AND ELECTRONIC DEVICE AND STORAGE MEDIUM", filed on December 27, 2023 with the China National Intellectual Property Administration, which is incorporated herein by reference in its entirety.

FIELD

[0002] The present disclosure relates to the technical field of image detection, and in particular, to an image detection method, an electronic device, and a storage medium.

BACKGROUND

[0003] Industrial defect detection is a significant application direction in the technical field of image processing. Industrial defect detection refers to a process of detecting products that may have defects in industrial application scenarios and performing anomaly determination on the products, thereby classifying the products as either acceptable or defective.

[0004] In the related art, industrial defect detection is generally implemented by using an unsupervised anomaly detection method. In this method, features of a set of predefined normal samples during training form a feature library, and the feature library is compressed based on a clustering algorithm to improve defect detection efficiency. However, since the accuracy of defect detection is limited by the compression rate of the feature library, compressing the feature library leads to a reduction in the accuracy of defect detection.

SUMMARY

[0005] In view of above issues, the present disclosure is provided. An image detection method, an electronic device, and a storage medium are provided according to the present disclosure.

[0006] According to an aspect of the present disclosure, an image detection method is

provided, comprising: obtaining a to-be-detected image;

inputting the to-be-detected image into a feature extraction model for feature extraction to obtain a to-be-detected feature vector set corresponding to a to-be-detected region in the to-be-detected image, where the to-be-detected region corresponds to a feature vector library, the feature vector library comprises at least one reference feature vector set, the at least one reference feature vector set corresponds to at least one reference region in a one-to-one manner, and the to-be-detected region corresponding to the feature vector library and the at least one reference region all correspond to a same image position;

comparing the to-be-detected feature vector set corresponding to the to-be-detected region with the reference feature vector set stored in the feature vector library corresponding to the to-be-detected region to determine whether the to-be-detected region meets a first preset condition, where the first preset condition is that a similarity between the to-be-detected feature vector set corresponding to the to-be-detected region and at least a portion of the reference feature vector set in the feature vector library is greater than a first similarity threshold; and

determining the to-be-detected region as an abnormal region in a case that the to-be-detected region does not meet the first preset condition.

[0007] In the foregoing technical solution, the to-be-detected feature vector set corresponding to the to-be-detected region in the to-be-detected image is extracted, and the to-be-detected feature vector set is compared with the reference feature vector set stored in the corresponding feature vector library, thereby accurately determining whether the to-be-detected region is an abnormal region. Moreover, since the to-be-detected region corresponding to the feature vector library including the to-be-detected feature vector set and the reference region both correspond to the same image position, the number of reference feature vector sets to be compared with the to-be-detected feature vector set is reduced. Therefore, the computational load is reduced and the image detection efficiency is improved.

[0008] In a possible implementation, the comparing the to-be-detected feature vector set corresponding to the to-be-detected region with the reference feature vector set stored in the feature vector library corresponding to the to-be-detected region to determine whether the to-be-detected region meets a first preset condition comprises:

calculating a first feature difference value between the to-be-detected feature vector set

corresponding to the to-be-detected region and each reference feature vector set of the at least one reference feature vector set stored in the feature vector library corresponding to the to-be-detected region;

determining a smallest first feature difference value among the first feature difference value as a first anomaly score; and

determining whether the first anomaly score is greater than or equal to a first score threshold, and determining that the to-be-detected region does not meet the first preset condition in a case that the first anomaly score is greater than or equal to the first score threshold, where the first anomaly score is negatively correlated with the similarity, and the first score threshold is negatively correlated with the first similarity threshold.

[0009] In the foregoing technical solution, the first feature difference value determined between the to-be-detected feature vector set and the reference feature vector set in the feature vector library that is most similar to the to-be-detected feature vector set serves as the first anomaly score, and based on the comparison result between the first anomaly score and the first score threshold, it is determined whether the to-be-detected region meets the first preset condition, providing an accurate basis for subsequent steps, thereby further improving the image detection efficiency.

[0010] In a possible implementation, to-be-detected region is at least two in number, and the at least two to-be-detected regions correspond to at least two feature vector libraries in a one-to-one manner, where in a case that the to-be-detected regions do not meet the first preset condition, the image detection method further comprises:

for each to-be-detected region of the at least two to-be-detected regions,

comparing the to-be-detected feature vector set corresponding to the to-be-detected region with a second reference feature vector set stored in a second feature vector library corresponding to another to-be-detected region to determine whether the to-be-detected region meets a second preset condition, where the second preset condition is that a similarity between the to-be-detected feature vector set corresponding to the to-be-detected region and at least a portion of the second reference feature vector set in the second feature vector library is greater than a second similarity threshold; and

determining the to-be-detected region as an offset-induced abnormal region in a case

that the to-be-detected region meets the second preset condition.

[0011] In the foregoing technical solution, in the case that the to-be-detected region does not meet the first preset condition, the to-be-detected feature vector set corresponding to the to-be-detected region is compared with the second reference feature vector set in the second feature vector library, enabling accurate determination of whether the to-be-detected region is an offset-induced abnormal region. Therefore, in addition to detecting abnormal regions caused by structural anomalies (such as dirt or damage), this solution also enables the detection of offset-induced abnormal regions. In conclusion, this solution expands the applicability of the image detection method and further enhances the user experience.

[0012] In a possible implementation, the comparing the to-be-detected feature vector set corresponding to the to-be-detected region with a second reference feature vector set stored in a second feature vector library corresponding to another to-be-detected region to determine whether the to-be-detected region meets a second preset condition, comprises:

calculating a second feature difference value between the to-be-detected feature vector set corresponding to the to-be-detected region and each second reference feature vector set stored in the second feature vector library;

determining a smallest second feature difference value among the second feature difference value as a second anomaly score;

determining whether the second anomaly score is less than a second score threshold; and

determining that the to-be-detected region corresponding to the to-be-detected feature vector set meets the second preset condition in a case that the second anomaly score is less than the second score threshold, where the second anomaly score is negatively correlated with the similarity.

[0013] In the foregoing technical solution, the second feature difference value determined between the to-be-detected feature vector set and the reference feature vector set in the second feature vector library that is most similar to the to-be-detected feature vector set serves as the second anomaly score, and based on the comparison result between the second anomaly score and the second score threshold, it is determined whether the to-be-detected region meets the second preset condition, enabling accurate determination of whether the to-be-detected region is an offset-induced abnormal region. The solution features simple computation, providing an

accurate basis for subsequent steps.

[0014] In a possible implementation, before inputting the to-be-detected image into a feature extraction model for feature extraction, the image detection method further comprises:

matching a to-be-detected product in the to-be-detected image with a template product in a template image to obtain a first matching result;

adjusting an image parameter of the to-be-detected image based on the first matching result, to make the to-be-detected image the same size as the template image and the to-be-detected product in the to-be-detected image be aligned with the template product in the template image after the adjustment, where the image parameter comprises one or more of: position, shape, and size,

where the at least one reference region is an image region within at least one reference image, each reference image of the at least one reference image is the same size as the template image, and a reference product in the reference image is aligned with the template product in the template image.

[0015] In the foregoing technical solution, before the to-be-detected image is input into the feature extraction model for feature extraction, the to-be-detected image is adjusted based on the template image, so that the adjusted to-be-detected image is the same size as the template image and the to-be-detected product in the to-be-detected image is aligned with the template product in the template image, thereby eliminating rotation and distortion in the to-be-detected image, unifying the scale between the to-be-detected image and the template image. Therefore, computational efficiency is improved, the robustness of the image detection method is enhanced, and the accuracy of image detection is improved.

[0016] In a possible implementation, the to-be-detected region is located at a target image position in the to-be-detected image; and the feature vector library is obtained through a feature vector library generation process, which comprises:

obtaining a sample image set, where the sample image set comprises at least one sample image, and each sample image of the at least one sample image in the sample image set does not comprise an abnormal region;

for each sample image of the at least one sample image in the sample image set, extracting an image feature vector corresponding to the sample image by using the feature extraction

model;

determining, from the image feature vector corresponding to the at least one sample image in the sample image set, a feature vector set corresponding to a reference region located at the target image position, and obtaining a feature vector library corresponding to the target image position based on the determined feature vector set,

where the feature vector library corresponding to the target image position is the feature vector library corresponding to the to-be-detected region located at the target image position, the at least one reference region is an image region within at least one reference image, and the at least one reference image comprises at least a portion of the sample image in the sample image set.

[0017] In the foregoing technical solution of the present disclosure, each target image position corresponds to a respective feature vector library, that is, each feature vector library only stores the feature vector sets corresponding to the target image position. Compared with the related art, this solution allows more sample images to be used for generating the feature vector library. Thereby, the feature vector library generated by this solution further improves the accuracy of image detection.

[0018] In a possible implementation, obtaining a feature vector library corresponding to the target image position based on the determined feature vector set comprises:

storing the determined feature vector set into the feature vector library corresponding to the target image position;

where the at least one reference region is the image region within the at least one reference image, and the at least one reference image comprises all sample images in the sample image set.

[0019] In this embodiment, all sample images used for generating the feature vector library are taken as reference images. Each feature vector library corresponds to a target image position and includes feature vector sets corresponding to the target image position in all the sample images. Thereby, when image detection is performed on the to-be-detected image based on the feature vector library, the accuracy of image detection is improved.

[0020] In a possible implementation, the sample image is more than one in number, and the obtaining a feature vector library corresponding to the target image position based on the

determined feature vector set comprises:

performing clustering on the determined feature vector sets to obtain at least two clusters;

storing a representative feature vector set comprised in each cluster of the at least two clusters into the feature vector library corresponding to the target image position,

where the representative feature vector set is any feature vector set in the cluster, each feature vector library corresponds to a reference region, and a reference image to which the reference region belongs is a sample image corresponding to a representative feature vector set stored in the feature vector library.

[0021] In the technical solution of this embodiment, the feature vector library only includes the representative feature vector sets of respective clusters, thereby reducing the memory required by the feature vector library. In addition, since this solution reduces the number of feature vector sets in the feature vector library through clustering, accordingly, it reduces the number of reference feature vector sets used for comparison with the to-be-detected feature vector set during the image detection process. Therefore, the efficiency of image detection is improved.

[0022] In a possible implementation, the feature vector library generation process further comprises:

obtaining a new sample image;

comparing image information of the new sample image with image information of the at least a portion of the sample image in the sample image set, to determine whether a change degree of the image information of the new sample image relative to the image information of the at least a portion of the sample image in the sample image set exceeds a preset degree threshold;

if the change degree of the image information of the new sample image relative to the image information of the at least a portion of the sample image in the sample image set exceeds the preset degree threshold, adding the new sample image to the sample image set, and determining the feature vector library based on the sample image set after the addition, where the image information comprises at least one of: texture information, color information, pixel information.

[0023] In the foregoing technical solution, when the change degree of image information of

the new sample image relative to the at least a portion of the sample images in the sample image set exceeds the preset degree threshold, the new sample image is added to the sample image set to determine the feature vector library based on the sample image set after addition, thereby ensuring the accuracy of image detection results based on the feature vector library. In addition, this incremental training method further expands the application scope of the image detection method.

[0024] According to another aspect of the present disclosure, an electronic device is provided, comprising a processor and a memory. The computer program instructions are stored in the memory, and the computer program instructions, when executed by the processor, performs the aforementioned image detection method.

[0025] In the foregoing technical solution, the to-be-detected feature vector set corresponding to the to-be-detected region in the to-be-detected image is extracted, and the to-be-detected feature vector set is compared with the reference feature vector set stored in the corresponding feature vector library, thereby accurately determining whether the to-be-detected region is an abnormal region. Moreover, since the to-be-detected region corresponding to the feature vector library including the to-be-detected feature vector set and the reference region both correspond to the same image position, the number of reference feature vector sets to be compared with the to-be-detected feature vector set is reduced. Therefore, the computational load is reduced and the image detection efficiency is improved.

[0026] According to another aspect of the present disclosure, a storage medium storing program instructions is provided. The program instructions, when executed, performs the aforementioned image detection method.

[0027] In the foregoing technical solution, the to-be-detected feature vector set corresponding to the to-be-detected region in the to-be-detected image is extracted, and the to-be-detected feature vector set is compared with the reference feature vector set stored in the corresponding feature vector library, thereby accurately determining whether the to-be-detected region is an abnormal region. Moreover, since the to-be-detected region corresponding to the feature vector library including the to-be-detected feature vector set and the reference region both correspond to the same image position, the number of reference feature vector sets to be compared with the to-be-detected feature vector set is reduced. Therefore, the computational load is reduced and the image detection efficiency is improved.

BRIEF DESCRIPTION OF THE DRAWINGS

[0028] The above and other objectives, features, and advantages of the present disclosure will become more apparent through a more detailed description of the embodiments of the present disclosure with reference to the accompanying drawings. The drawings are provided to further understand the embodiments of the present disclosure and constitute a part of the specification, and are used together with the embodiments of the present disclosure to explain the present disclosure, and do not constitute a limitation on the present disclosure. In the drawings, the same reference numerals generally represent the same components or steps.

[0029] FIG. 1 shows a schematic flowchart of an image detection method according to an embodiment of the present disclosure;

[0030] FIG. 2 shows a schematic diagram of a to-be-detected image according to an embodiment of the present disclosure;

[0031] FIG. 3 shows a schematic diagram of an adjusted to-be-detected image according to an embodiment of the present disclosure;

[0032] FIG. 4 shows a schematic diagram of a generation process of a feature vector library according to an embodiment of the present disclosure; and

[0033] FIG. 5 shows a schematic block diagram of an electronic device according to an embodiment of the present disclosure.

DETAILED DESCRIPTION

[0034] To make the objectives, technical solutions, and advantages of the present disclosure clearer, embodiments according to the present disclosure will be described in detail with reference to the accompanying drawings. Apparently, the described embodiments are only a part of the embodiments of the present disclosure, rather than all of them. It should be understood that the present disclosure is not limited to the embodiments described herein. Based on the embodiments of the present disclosure described herein, all other embodiments obtained by those skilled in the art without creative effort shall fall within the protection scope of the present disclosure.

[0035] Industrial defect detection is a significant application direction in the technical field of image processing. Currently, anomaly detection algorithms in the industrial field are generally implemented by an unsupervised algorithm. The algorithm performs defect detection without collecting defect datasets, avoiding problems such as difficulty in collecting low-frequency defect datasets and difficulty in intercepting defects of unknown categories. Moreover, in the case of insufficient defect datasets, the algorithm effectively addresses the detection issue, significantly reducing a project period, and bringing faster efficiency gains to the production process for enterprises. Anomaly detection methods may be implemented using any one of the following algorithms: Patchcore algorithm, CSlow algorithm, DFM algorithm, DRAEM algorithm, FastFlow algorithm, GANomaly algorithm, PADIM algorithm, etc. Taking the Patchcore algorithm as an example, the algorithm extracts mid-level features in images through a pre-trained model based on ImageNet, and forms a feature library by using the extracted mid-level features. During defect detection, it is determined whether there is a defect region in the to-be-processed image by performing matching search between features of the to-be-processed image and features stored in the feature library. In the related art, to improve the efficiency of defect detection, the feature library is often compressed using, for example, a clustering algorithm. However, compressing the feature library leads to a reduction in the number of sample features in the feature library. The reduction in the number of sample features in the feature library then leads to a decrease in the accuracy of defect detection. In other words, the accuracy of defect detection is limited by the compression rate of the feature library, and as the compression rate of the feature library increases, the accuracy of defect detection decreases. Therefore, the defect detection methods in the related art suffer from poor accuracy. In view of this, the present disclosure provides an image detection method, an electronic device, and a storage medium. The image detection method improves the accuracy of image detection. The image detection method, electronic device, and storage medium are described in detail below.

[0036] According to an aspect of the present disclosure, an image detection method is provided. FIG. 1 shows a schematic flowchart of an image detection method according to an embodiment of the present disclosure. As shown in FIG. 1, the image detection method 100 may include step S110, step S120, step S130, and step S140.

[0037] In step S110, a to-be-detected image is obtained.

[0038] The to-be-detected image according to an embodiment of the present disclosure may

be an image of any object to be subjected to defect detection. In other words, the to-be-detected image includes a target object to be subjected to defect detection. The target object may be any suitable object, including but not limited to metals, glass, paper, electronic components, etc., which have strict appearance requirements and clear criteria. The present disclosure does not impose limitations thereon.

[0039] In a possible implementation, the to-be-detected image may be a black-and-white image or a color image. In a possible implementation, the to-be-detected image may be an image of any size or resolution. Alternatively, the to-be-detected image may also be an image meeting preset resolution requirements. In an example, the to-be-detected image may be a black-and-white image with a size of 512*512 pixels. Requirements for the to-be-detected image may be set based on practical detection needs, hardware conditions of the image acquisition apparatus, and requirements of the model (e.g., the feature extraction model mentioned below) for input images. The present disclosure does not impose limitations thereon.

[0040] In a possible implementation, the to-be-detected image may be an original image captured by an image acquisition apparatus. According to an embodiment of the present disclosure, any existing or future image acquisition method may be adopted to obtain the to-be-detected image. In a possible implementation, an image acquisition apparatus in a machine vision inspection system may be used to obtain the to-be-detected image, such as using a lighting device, lens, high-speed camera, and image acquisition card that match the inspection environment and the to-be-detected object to obtain the to-be-detected image.

[0041] In another example, the to-be-detected image may be an image obtained after performing a preprocessing operation on an original image.

[0042] In a possible implementation, the preprocessing operation may be any preprocessing operation that meets the requirements of subsequent image detection steps, and the preprocessing operation may include all operations that facilitate image detection on the to-be-detected image, such as improving the visual effect of the image, enhancing image clarity, or highlighting certain features in the image. In a possible implementation, the preprocessing operation may include denoising operations such as filtering, and may further include adjustments of image parameters, such as image enhancement, and adjustments of grayscale, contrast, and brightness. Alternatively, the preprocessing operation may include normalization processing performed on pixels of the to-be-detected image. For example, each pixel of the to-

be-detected image may be divided by 255 so that pixels of the preprocessed to-be-detected image are in the range of 0 to 1, which improves the efficiency of subsequent image detection.

[0043] In a possible implementation, the preprocessing operation may further include operations such as cropping an image or deleting an image. For example, an original image may be cropped to a size required by the model, or an original image that does not meet image quality requirements may be deleted to obtain to-be-detected images that meet image quality requirements, etc.

[0044] In a possible implementation, the number of to-be-detected images may be one or multiple. In a possible implementation, the number of to-be-detected image is one, for example, only one to-be-detected image is obtained each time. Alternatively, the number of to-be-detected images may be multiple, for example, 10 or 500, and multiple to-be-detected images may be obtained at one time, and then the multiple to-be-detected images are input as a batch into the subsequently described model for defect detection.

[0045] In step S120, the to-be-detected image is input into a feature extraction model for feature extraction to obtain a to-be-detected feature vector set corresponding to a to-be-detected region in the to-be-detected image. The to-be-detected region corresponds to a feature vector library, and the feature vector library includes at least one reference feature vector set. The at least one reference feature vector set corresponds to at least one reference region in a one-to-one manner, and the to-be-detected region corresponding to the feature vector library and the at least one reference region all correspond to a same image position.

[0046] In a possible implementation, the number of the to-be-detected region in the to-be-detected image may be one. In some embodiments, the entire to-be-detected image may be taken as a to-be-detected region, and then the to-be-detected feature vector set corresponding to the to-be-detected region is obtained. In other embodiments, the to-be-detected region may be a region on the target object to be subjected to defect detection, where the region is determined to be prone to defects based on experience.

[0047] In a possible implementation, the number of to-be-detected regions in the to-be-detected image may be at least two. The at least two to-be-detected regions correspond to at least two feature vector libraries in a one-to-one manner. Each to-be-detected region of the at least two to-be-detected regions may be referred to as a position in the to-be-detected image. Similarly, each reference region may be referred to as a position in the reference image. It can

be understood that a to-be-detected region and a reference region which correspond to a same feature vector library correspond to a same image position, which indicates that the position of the to-be-detected region in the corresponding to-be-detected image is the same as the position of the reference region in the corresponding reference image.

[0048] In the present disclosure, the reference image and the to-be-detected image include the same type of target object. Moreover, the target object in the reference image is a defect-free target object. Taking the target object being a wafer as an example, the reference image may be an image of a normal wafer (i.e., a defect-free wafer), and the to-be-detected image may be an image of a to-be-detected wafer.

[0049] In a possible implementation, the reference region may be any image region in the reference image, and the size of the reference region is smaller than the size of the reference image. In the related art, features of a predefined set of normal samples (equivalent to reference images) during training are usually stored in a feature library. In the solution of the present disclosure, for each reference region in the reference image, the reference feature vector set of the reference region forms a respective feature vector library. Therefore, when performing the subsequent step S130, the computational focus is narrowed to a critical range, and thus improving both accuracy and efficiency.

[0050] In a possible implementation, the number of to-be-detected regions is at least two, and the size of the to-be-detected region may be determined as needed. For example, each pixel point in the to-be-detected image may be taken as a to-be-detected region. For another example, a region formed by a preset number of pixel points in the to-be-detected image may be taken as a to-be-detected region.

[0051] In a possible implementation, the feature extraction model may be any existing or future-developed neural network model for feature extraction. For example, the neural network may be any one of the following neural networks or a combination of several neural networks: Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), etc. In an embodiment, the feature extraction model may be a pre-trained model based on ImageNet. In another embodiment, the feature extraction model may be a pre-trained convolutional neural network model. In the embodiment, the to-be-detected image may be input into the pre-trained convolutional neural network model to obtain to-be-detected feature vector sets in one-to-one correspondence with at least two to-be-detected regions. The to-be-detected feature vector set

may be a feature vector corresponding to the to-be-detected region at an intermediate layer of the pre-trained convolutional neural network model.

[0052] In a possible implementation, the to-be-detected region may be a partial region of the to-be-detected image, and each to-be-detected region corresponds to a feature vector library. The to-be-detected image includes at least two to-be-detected regions, and the at least two to-be-detected regions may together cover the entire region of the to-be-detected image or only a partial region of the to-be-detected image. In a possible implementation, the image region covered by the at least two to-be-detected regions is the same as the image region of the to-be-detected image. In other words, the at least two to-be-detected regions are obtained by dividing the to-be-detected image according to a preset division rule. For example, when the number of at least two to-be-detected regions is four, the to-be-detected image may be divided into four regions of the same size, each region being a to-be-detected region. FIG. 2 shows a schematic diagram of a to-be-detected image according to an embodiment of the present disclosure. As shown in FIG. 2, the to-be-detected image is divided into four regions: region A, region B, region C, and region D. In this embodiment, all of region A, region B, region C, and region D may be taken as to-be-detected regions.

[0053] In a possible implementation, the method 100 may further include: dividing the to-be-detected image into at least two to-be-detected regions, where a size of each to-be-detected region of the at least two to-be-detected regions is the same as a size of a corresponding reference region, and the at least two to-be-detected regions do not overlap each other.

[0054] In this embodiment, the to-be-detected image is divided into at least two to-be-detected regions in a non-overlapping manner, ensuring that defect detection is performed on each region in the to-be-detected image, thereby avoiding missed detection.

[0055] In a possible implementation, the at least two to-be-detected regions may be regions in the to-be-detected image that require focused observation. For example, based on experience, regions on the target object to be defect-detected that are prone to defects may be determined, and then image regions in the to-be-detected image corresponding to those defect-prone regions are determined as to-be-detected regions. FIG. 2 is taken as an example to explain the to-be-detected regions in the to-be-detected image. As shown in FIG. 2, the to-be-detected image is divided into region A, region B, region C, and region D. The user may determine region B and region C as to-be-detected regions based on experience. In the solution of this embodiment,

only a portion of the regions in the to-be-detected image are taken as to-be-detected regions, thereby reducing computational load and improving image detection efficiency.

[0056] In step S130, the to-be-detected feature vector set corresponding to the to-be-detected region is compared with the reference feature vector set stored in the feature vector library corresponding to the to-be-detected region to determine whether the to-be-detected region meets a first preset condition. The first preset condition is that a similarity between the to-be-detected feature vector set corresponding to the to-be-detected region and at least a portion of the reference feature vector set in the feature vector library is greater than a first similarity threshold.

[0057] In a possible implementation, in a case that the number of to-be-detected regions is at least two, for each to-be-detected region of the at least two to-be-detected regions, the feature vector library corresponding to the to-be-detected region may be referred to as a first feature vector library, and another feature vector library of the at least two feature vector libraries other than the first feature vector library may be referred to as a second feature vector library. The reference feature vector set stored in the first feature vector library may be referred to as a first reference feature vector set. Step S130, comparing the to-be-detected feature vector set corresponding to the to-be-detected region with the reference feature vector set stored in the feature vector library corresponding to the to-be-detected region to determine whether the to-be-detected region meets a first preset condition, may include: for each to-be-detected region of the at least two to-be-detected regions, comparing the to-be-detected feature vector set corresponding to the to-be-detected region with the first reference feature vector set stored in the first feature vector library corresponding to the to-be-detected region to determine whether the to-be-detected region meets the first preset condition. As mentioned above, in the related art, features of a predefined set of normal samples (equivalent to reference images) during training are usually stored in a feature library. When detecting the to-be-detected image by using algorithms such as Patchcore, features of each region in the to-be-detected image are compared with all features in the feature library, resulting in a large amount of computation and low efficiency. In this embodiment, it is only required to compare the to-be-detected feature vector set of each to-be-detected region with the first reference feature vector set stored in the first feature vector library corresponding to the to-be-detected region. Therefore, computational focus is narrowed to a critical range, computational load is reduced, and image detection efficiency is improved.

[0058] In a possible implementation, the number of reference feature vector sets in the at least a portion of reference feature vector set may be set as needed. For example, the number of the reference feature vector set in the at least a portion of reference feature vector set may be one. In this embodiment, if the similarity between at least one reference feature vector set in the feature vector library corresponding to the to-be-detected region and the to-be-detected feature vector set is greater than the similarity threshold, it is determined that the to-be-detected region meets the first preset condition. For another example, the number of reference feature vector sets in the at least a portion of reference feature vector set may be two. In this embodiment, if the similarity between each of the at least two reference feature vector sets in the feature vector library corresponding to the to-be-detected region and the to-be-detected feature vector set is greater than the similarity threshold, it is determined that the to-be-detected region meets the first preset condition.

[0059] In a possible implementation, any existing or future method for calculating the similarity between two feature vector sets may be used to compare the to-be-detected feature vector set with the corresponding reference feature vector set. For example, the Euclidean distance or Mahalanobis distance between the to-be-detected feature vector set and the corresponding reference feature vector set may be calculated, and the similarity between the to-be-detected feature vector set and the corresponding reference feature vector set may be determined based on the calculation result.

[0060] In a possible implementation, the first similarity threshold may be set as needed. A greater first similarity threshold indicates a stricter standard for determining whether the to-be-detected region is an abnormal region and leads to a higher accuracy of image detection. Conversely, a smaller first similarity threshold indicates a looser standard and a lower accuracy. Therefore, in some embodiments, the first similarity threshold may be set according to the image detection accuracy required by the user.

[0061] In step S140, in a case that the to-be-detected region does not meet the first preset condition, the to-be-detected region is determined as an abnormal region.

[0062] In the case that the to-be-detected region does not meet the first preset condition, it is indicated that the similarity between the to-be-detected region and the reference feature vector set in the corresponding feature vector library is low. In other words, the to-be-detected region is highly likely to be an abnormal region deviating from the normal region (i.e., the reference

region corresponding to the reference feature vector set). Therefore, when the to-be-detected region does not meet the first preset condition, it is determined that the to-be-detected region is an abnormal region.

[0063] In the foregoing technical solution, the to-be-detected feature vector set corresponding to the to-be-detected region in the to-be-detected image is extracted, and the to-be-detected feature vector set is compared with the reference feature vector set stored in the corresponding feature vector library, thereby accurately determining whether the to-be-detected region is an abnormal region. Moreover, since the to-be-detected region corresponding to the feature vector library including the to-be-detected feature vector set and the reference region both correspond to the same image position, the number of reference feature vector sets to be compared with the to-be-detected feature vector set is reduced. Therefore, the computational load is reduced and the image detection efficiency is improved.

[0064] In a possible implementation, in step S130, comparing the to-be-detected feature vector set corresponding to the to-be-detected region with the reference feature vector set stored in the feature vector library corresponding to the to-be-detected region to determine whether the to-be-detected region meets a first preset condition may include: calculating a first feature difference value between the to-be-detected feature vector set corresponding to the to-be-detected region and each reference feature vector set of the at least one reference feature vector set stored in the feature vector library corresponding to the to-be-detected region; determining a smallest first feature difference value among the first feature difference value as a first anomaly score; determining whether the first anomaly score is greater than or equal to a first score threshold; and determining that the to-be-detected region does not meet the first preset condition in a case that the first anomaly score is greater than or equal to the first score threshold, where the first anomaly score is negatively correlated with the similarity, and the first score threshold is negatively correlated with a similarity threshold.

[0065] In a possible implementation, the first feature difference value may be determined using any existing or future method for calculating the similarity between two feature vector sets. For example, any one of the Euclidean (L2) norm, Euclidean distance, and Mahalanobis distance between the to-be-detected feature vector set and the reference feature vector set may be calculated to determine the first feature difference value. In an embodiment, the L2 norm between the to-be-detected feature vector set corresponding to the to-be-detected region and

each reference feature vector set stored in the feature vector library corresponding to the to-be-detected region may be calculated, and the L2 norm serves as the first feature difference value between the to-be-detected feature vector set and the reference feature vector set.

[0066] After first feature difference values between the to-be-detected feature vector set and respective reference feature vector sets are obtained, the smallest first feature difference value may be taken as the first anomaly score. Taking the embodiment where the L2 norm serves as the first feature difference value as an example, after L2 norms corresponding to respective reference feature vector sets are obtained, the L2 norms are compared with each other, and the smallest L2 norm is determined as the first anomaly score. The L2 norm serving as the first anomaly score is calculated between the to-be-detected feature vector set and a reference feature vector set in the feature vector library that is most similar to the to-be-detected feature vector set.

[0067] In a possible implementation, the first score threshold may be set as needed. As mentioned above, the first anomaly score is negatively correlated with the similarity, and the first score threshold is negatively correlated with the similarity threshold. Therefore, in a case that the first anomaly score is greater than or equal to the first score threshold, it is indicated that the similarity between the to-be-detected feature vector set and the first reference feature vector set in the feature vector library is less than or equal to the similarity threshold. In this case, it is determined that the to-be-detected region does not meet the first preset condition. That is, the to-be-detected region is an abnormal region.

[0068] In the foregoing technical solution, the first feature difference value determined between the to-be-detected feature vector set and the reference feature vector set in the feature vector library that is most similar to the to-be-detected feature vector set serves as the first anomaly score, and based on the comparison result between the first anomaly score and the first score threshold, it is determined whether the to-be-detected region meets the first preset condition, providing an accurate basis for subsequent steps (e.g., step S140), thereby further improving the image detection efficiency.

[0069] In a possible implementation, the number of to-be-detected regions is at least two, and the at least two to-be-detected regions correspond to at least two feature vector libraries in a one-to-one manner. When the to-be-detected regions do not meet the first preset condition, the image detection method may further include the following step S150 and step S160.

[0070] In step S150, for each to-be-detected region of the at least two to-be-detected regions, the to-be-detected feature vector set corresponding to the to-be-detected region is compared with a second reference feature vector set stored in a second feature vector library corresponding to another to-be-detected region to determine whether the to-be-detected region meets a second preset condition, where the second preset condition is that a similarity between the to-be-detected feature vector set corresponding to the to-be-detected region and at least a portion of the second reference feature vector set in the second feature vector library is greater than a second similarity threshold.

[0071] In step S160, for each to-be-detected region of the at least two to-be-detected regions, in a case that the to-be-detected region meets the second preset condition, the to-be-detected region is determined as an offset-induced abnormal region.

[0072] The manner of comparing the to-be-detected feature vector set corresponding to the to-be-detected region with the second reference feature vector set stored in the second feature vector library corresponding to another to-be-detected region is similar to that of comparing the to-be-detected feature vector set with the corresponding reference feature vector set, and will not be repeated here.

[0073] In a possible implementation, comparing the to-be-detected feature vector set corresponding to the to-be-detected region with a second reference feature vector set stored in a second feature vector library corresponding to another to-be-detected region may include: comparing the to-be-detected feature vector set corresponding to the to-be-detected region with a second reference feature vector set stored in each of second feature vector libraries corresponding to to-be-detected regions other than the to-be-detected region. In this embodiment, the to-be-detected feature vector set is compared with the second reference feature vector sets stored in the second feature vector libraries corresponding to other to-be-detected regions, and the similarity between the to-be-detected feature vector set and the second reference feature vector set in each of the second feature vector libraries is determined.

[0074] In a possible implementation, the number of second reference feature vector sets in the at least a portion of the second reference feature vector sets in the second feature vector library may be set as needed. For example, the number of the second reference feature vector set in the at least a portion of the second reference feature vector sets may be one. In this embodiment, when a similarity between at least one second reference feature vector set in the

second feature vector library and the to-be-detected feature vector set is greater than the similarity threshold, it may be determined that the to-be-detected region meets the second preset condition. For another example, the number of second reference feature vector sets in the at least a portion of the second reference feature vector sets may be two. In this embodiment, when a similarity between each of at least two second reference feature vector sets in the second feature vector library and the to-be-detected feature vector set is greater than the similarity threshold, it may be determined that the to-be-detected region meets the second preset condition.

[0075] In a possible implementation, the second similarity threshold may be set as needed. For example, the second similarity threshold may be set according to the image detection accuracy required by the user. In a case that the to-be-detected region meets the second preset condition, it is indicated that the similarity between the to-be-detected region and the second reference feature vector set in the second feature vector library is high. That is, the similarity between the to-be-detected region and the reference region corresponding to the second reference feature vector set is high. In this case, it is determined that the to-be-detected region is an offset-induced abnormal region. In other words, the cause of the anomaly in the to-be-detected region may be that the image acquisition apparatus (e.g., a camera) is not aligned with the target object when capturing the to-be-detected image.

[0076] In the foregoing technical solution, in the case that the to-be-detected region does not meet the first preset condition, the to-be-detected feature vector set corresponding to the to-be-detected region is compared with the second reference feature vector set in the second feature vector library, enabling accurate determination of whether the to-be-detected region is an offset-induced abnormal region. Therefore, in addition to detecting abnormal regions caused by structural anomalies (such as dirt or damage), this solution also enables the detection of offset-induced abnormal regions. In conclusion, this solution expands the applicability of the image detection method and further enhances the user experience.

[0077] In a possible implementation, step S150, comparing the to-be-detected feature vector set corresponding to the to-be-detected region with a second reference feature vector set stored in a second feature vector library corresponding to another to-be-detected region to determine whether the to-be-detected region meets a second preset condition may include: calculating a second feature difference value between the to-be-detected feature vector set corresponding to the to-be-detected region and each second reference feature vector set stored in the second

feature vector library; determining a smallest second feature difference value among the second feature difference value as a second anomaly score; determining whether the second anomaly score is less than a second score threshold; and determining that the to-be-detected region corresponding to the to-be-detected feature vector set meets the second preset condition in a case that the second anomaly score is less than the second score threshold, where the second anomaly score is negatively correlated with the similarity.

[0078] Similar to the first feature difference value, the second feature difference value may further be determined by using any one of the methods such as the Euclidean (L2) norm, Euclidean distance, or Mahalanobis distance. In an embodiment, the L2 norm between the to-be-detected feature vector set corresponding to the to-be-detected region and each second reference feature vector set stored in the second feature vector library may be calculated, and the L2 norm serves as the second feature difference value between the to-be-detected feature vector set and the second reference feature vector set.

[0079] After second feature difference values between the to-be-detected feature vector set and respective reference feature vector sets are obtained, the smallest second feature difference value may be taken as the second anomaly score. Taking the embodiment where the L2 norm serves as the second feature difference value as an example, after L2 norms corresponding to respective first reference feature vector sets are obtained, the L2 norms are compared with each other, and the smallest L2 norm is determined as the second anomaly score. The L2 norm serving as the second anomaly score is calculated between the to-be-detected feature vector set and a reference feature vector set in the second feature vector library that is most similar to the to-be-detected feature vector set.

[0080] In a possible implementation, the second score threshold may be set as needed. As mentioned above, the second anomaly score is negatively correlated with the similarity. Therefore, in a case that the second anomaly score is less than the second score threshold, it is indicated that the similarity between the to-be-detected feature vector set and at least a portion of the second reference feature vector set in the second feature vector library is greater than the similarity threshold. In this case, it is determined that the to-be-detected region meets the second preset condition. That is, the to-be-detected region is an offset-induced abnormal region.

[0081] In the foregoing technical solution, the second feature difference value determined between the to-be-detected feature vector set and the reference feature vector set in the second

feature vector library that is most similar to the to-be-detected feature vector set serves as the second anomaly score, and based on the comparison result between the second anomaly score and the second score threshold, it is determined whether the to-be-detected region meets the second preset condition, enabling accurate determination of whether the to-be-detected region is an offset-induced abnormal region. The solution features simple computation, providing an accurate basis for subsequent steps (e.g., step S160).

[0082] In a possible implementation, before inputting the to-be-detected image into the feature extraction model for feature extraction, the image detection method may further include: matching a to-be-detected product in the to-be-detected image with a template product in a template image to obtain a first matching result; adjusting an image parameter of the to-be-detected image based on the first matching result, to make the to-be-detected image the same size as the template image and the to-be-detected product in the to-be-detected image be aligned with the template product in the template image after the adjustment, where the image parameter includes one or more of position, shape, and size; the at least one reference region is an image region within at least one reference image, each reference image of the at least one reference image is the same size as the template image, and a reference product in the reference image is aligned with the template product in the template image.

[0083] In the present disclosure, the template image and the to-be-detected image include the same type of target object. Taking the target object being a wafer as an example, the to-be-detected image may be an image of a to-be-detected wafer, and the template image may be an image of a normal wafer, and the position of the wafer in the template image meets the user requirements.

[0084] In a possible implementation, the first matching result may be determined using any existing or future method for image matching. For example, a method based on grayscale matching may be used to determine the first matching result. The method based on grayscale matching may be any one of the following algorithms: Normalized Cross-Correlation (NCC) matching method, Mean Absolute Difference (MAD) algorithm, Sequential Similarity Detection Algorithm (SSDA), etc. For another example, a method based on feature matching may be used to determine the first matching result. The method based on feature matching may be any one of the following algorithms: Scale-Invariant Feature Transform (SIFT) algorithm, Speeded-Up Robust Features (SURF) algorithm, etc.

[0085] After the first matching result is obtained, the image parameter of the to-be-detected image may be adjusted based on the first matching result. The adjustment operation may include, but is not limited to, at least one of the following operations: translation, rotation, scaling, etc. In an embodiment, the first matching result may be the grayscale correlation between each of pixel points corresponding to the to-be-detected product in the to-be-detected image and each of pixel points corresponding to the template product in the template image, where the grayscale correlation is determined using a method based on grayscale matching. After the first matching result is obtained, the to-be-detected image may be transformed through translation, rotation, scaling, etc., according to the grayscale correlation determined for the respective pixel points corresponding to the to-be-detected product, so that the to-be-detected image is the same size as the template image and the to-be-detected product in the to-be-detected image is aligned with the template product in the template image. FIG. 3 shows a schematic diagram of an adjusted to-be-detected image according to an embodiment of the present disclosure. In this embodiment, the target object is a wafer. After the to-be-detected image is obtained, the to-be-detected image may be matched to align with the template image, so that the to-be-detected image is the same size as the template image and the to-be-detected product in the to-be-detected image is aligned with the template product in the template image. The specific matching and alignment method has been described in detail above and will not be elaborated here.

[0086] In the foregoing technical solution, before the to-be-detected image is input into the feature extraction model for feature extraction, the to-be-detected image is adjusted based on the template image, so that the adjusted to-be-detected image is the same size as the template image and the to-be-detected product in the to-be-detected image is aligned with the template product in the template image, thereby eliminating rotation and distortion in the to-be-detected image, unifying the scale between the to-be-detected image and the template image. Therefore, computational efficiency is improved, the robustness of the image detection method is enhanced, and the accuracy of image detection is improved.

[0087] In a possible implementation, the to-be-detected region is located at a target image position in the to-be-detected image, and the feature vector library is obtained through the following feature vector library generation process: obtaining a sample image set, where the sample image set includes at least one sample image, and each sample image in the sample image set does not include an abnormal region; for each sample image of the at least one sample image in the sample image set, extracting an image feature vector corresponding to the sample

image by using the feature extraction model; determining, from the image feature vector corresponding to the at least one sample image in the sample image set, a feature vector set corresponding to a reference region located at the target image position, and obtaining a feature vector library corresponding to the target image position based on the determined feature vector set, where the feature vector library corresponding to the target image position is the feature vector library corresponding to the to-be-detected region located at the target image position, the at least one reference region is an image region within at least one reference image, and the at least one reference image includes at least a portion of the sample image in the sample image set.

[0088] In a possible implementation, the at least one sample image in the sample image set may be obtained using any existing or future image acquisition method. For example, images of the target object may be captured manually, and the captured images may be used as sample images. For another example, sample images of the target object may be obtained through web crawling.

[0089] In a possible implementation, the number of sample images in the sample image set may be set as needed. A large number of sample images indicates accurate detection results of image detection on the to-be-detected image by using the obtained feature vector library, however, it further indicates long training time. Therefore, the number of sample images may be determined by comprehensively considering the accuracy of image detection results and the training time.

[0090] In a possible implementation, in a case that the number of to-be-detected regions is at least two, the positions of the at least two to-be-detected regions in the to-be-detected image are respectively at least two target image positions. In a possible implementation, the determining, from the image feature vector corresponding to the at least one sample image in the sample image set, a feature vector set corresponding to a reference region located at the target image position, and obtaining a feature vector library corresponding to the target image position based on the determined feature vector set may include: for each target image position of the at least two target image positions, determining, from the image feature vector corresponding to the at least one sample image in the sample image set, a feature vector set corresponding to a reference region located at the target image position, and obtaining a feature vector library corresponding to the target image position based on the determined feature vector

set. In this embodiment, after the image feature vectors corresponding to the respective sample images in the sample image set are obtained, for each target image position of the target image positions, the feature vector sets corresponding to the target image position of all the sample images may form a feature vector library corresponding to the target image position. Through this process, feature vector libraries corresponding to the respective target image positions in the sample images may be obtained.

[0091] In a possible implementation, all the target image positions may correspond to all pixels in the sample image in a one-to-one manner. Alternatively, the target image positions may respectively correspond to regions in the sample image, where each region of the regions is formed by a preset number of pixels. The preset number may be determined as needed, for example, the preset number may be 9.

[0092] In a possible implementation, the at least one reference image may be all sample images in the sample image set. In other words, all sample images used for generating the feature vector library may be taken as reference images. Each feature vector library corresponds to a target image position and includes feature vector sets corresponding to the target image position of all the sample images. Alternatively, the at least one reference image may be a portion of the sample images in the sample image set. In this embodiment, a portion of the sample images in the sample image set may be selected as reference images. Each feature vector library corresponds to a target image position and includes feature vector sets corresponding to the target image position in the portion of the sample images.

[0093] FIG. 4 shows a schematic diagram of a generation process of a feature vector library according to an embodiment of the present disclosure. As shown in FIG. 4, first, for each of sample images, based on the template product in the template image, the sample image is aligned and flattened, so that after the alignment and flattening, the sample image is the same size as the template image and the sample product in the sample image is aligned with the template product in the template image. After being aligned and flattened, the sample images are input into a pre-trained convolutional neural network (i.e., the feature extraction model) to obtain the image feature vectors of the sample images at an intermediate layer (in this embodiment, the intermediate layer is the second or third layer). Ultimately, after the image feature vectors of the respective sample images are obtained, for each of target image positions, feature vector sets corresponding to a reference region located at the target image position in

the respective sample images form a feature vector library, thereby obtaining feature vector libraries corresponding one-to-one with the respective target image positions in the sample image.

[0094] In the related art, all image feature vectors corresponding to the sample images are usually stored in the same feature vector library. When the storage space of the feature vector library is limited and a large number of sample images are used during training, it will lead to a longer compression time required for subsequent compression of image feature vectors. Therefore, in the related art, the number of sample images used for generating the feature vector library is greatly affected by the storage space of the feature vector library. In the aforementioned technical solution of the present disclosure, each target image position corresponds to a respective feature vector library, that is, each feature vector library only stores the feature vector sets corresponding to the target image position. Compared with the related art, this solution allows more sample images to be used for generating the feature vector library. Thereby, the feature vector library generated by this solution further improves the accuracy of image detection.

[0095] In a possible implementation, obtaining the feature vector library corresponding to the target image position based on the determined feature vector set may include: storing the determined feature vector set into the feature vector library corresponding to the target image position, where the at least one reference region is an image region within at least one reference image, and the at least one reference image includes all sample images in the sample image set.

[0096] In this embodiment, all sample images used for generating the feature vector library are taken as reference images. Each feature vector library corresponds to a target image position and includes feature vector sets corresponding to the target image position in all the sample images. Thereby, when image detection is performed on the to-be-detected image based on the feature vector library, the accuracy of image detection is improved.

[0097] In a possible implementation, the number of the sample images is multiple, and obtaining a feature vector library corresponding to the target image position based on the determined feature vector set may include: performing clustering on the determined feature vector sets to obtain at least two clusters; storing a representative feature vector set included in each cluster of the at least two clusters into the feature vector library corresponding to the target image position, where the representative feature vector set is any feature vector set in the cluster,

each feature vector library corresponds to a reference region, and a reference image to which the reference region belongs is a sample image corresponding to a representative feature vector set stored in the feature vector library.

[0098] In a possible implementation, the representative feature vector set may correspond to a cluster center of each cluster. In some embodiments, the feature vector set corresponding to the cluster center included in each cluster may be stored in the feature vector library corresponding to the target image position. In this embodiment, the feature vector library only includes the representative feature vector sets of respective clusters, thereby reducing the memory required by the feature vector library. In addition, since this solution reduces the number of feature vector sets in the feature vector library through clustering, accordingly, it reduces the number of reference feature vector sets used for comparison with the to-be-detected feature vector set during the image detection process. Therefore, the efficiency of image detection is improved.

[0099] In a possible implementation, the feature vector library generation process may further include: obtaining a new sample image; comparing image information of the new sample image with image information of the at least a portion of the sample image in the sample image set, to determine whether a change degree of the image information of the new sample image relative to the image information of the at least a portion of the sample image in the sample image set exceeds a preset degree threshold; if the change degree of the image information of the new sample image relative to the image information of the at least a portion of the sample image in the sample image set exceeds the preset degree threshold, adding the new sample image to the sample image set, and determining the feature vector library based on the sample image set after the addition, where the image information includes at least one of: texture information, color information, pixel information.

[0100] The preset degree threshold may be set as needed. It can be understood that when the change degree of the image information of the new sample image relative to the image information of the at least a portion of the sample image in the sample image set exceeds the preset degree threshold, it may indicate that there is a difference between the target object included in the new sample image and the target object corresponding to the current sample image set. Taking the target object being a wafer as an example, due to differences in production processes, etc., there may be differences between the new sample image and the sample images

in the sample image set. In this case, the feature vector library trained and generated based on the current sample image set cannot be used for accurately determining whether the target object corresponding to the new sample image has defects. In the embodiment of the present disclosure, the difference between the image information of the new sample image and the image information of the at least a portion of the sample image in the sample image set is obtained by comparing, and when the change degree of image information exceeds the preset degree threshold, the new sample image is added to the sample image set to generate a new feature vector library. The accuracy of image detection is ensured.

[0101] In a possible implementation, the image information of the new sample image may be compared with the image information of the at least a portion of the sample image in the sample image set manually. For example, it may be manually determined whether there is a difference between the new sample image and the sample images in the sample image set. Alternatively, an algorithm model may be used to compare the image information of the new sample image with the image information of the at least a portion of the sample image in the sample image set. For example, the image information of the new sample image and the image information of the at least a portion of the sample image in the sample image set are collected by using the algorithm model, and then the change degree of image information between the image information of the new sample image and the image information of the at least a portion of the sample image in the sample image set is obtained by comparing. In an embodiment, a Gaussian similarity algorithm may be used to calculate the similarity between the image information of the new sample image and the image information of the at least a portion of the sample image in the sample image set. The similarity is negatively correlated with the change degree of image information.

[0102] In a possible implementation, the image information includes at least one of the following: texture information, color information, pixel information. In some embodiments, the image information may include texture information. In this embodiment, texture information extraction algorithms such as the Histogram of Oriented Gradients (e.g., FHOG) feature extraction algorithm may be used to extract the texture information of the new sample image and the texture information of the at least a portion of the sample images in the sample image set. In other embodiments, the image information may include color information. In this embodiment, color information extraction algorithms such as the Color Name (CN) feature extraction algorithm may be used to extract the color information of the new sample image and

the color information of the at least a portion of the sample images in the sample image set. In some embodiments, the image information may include pixel information. In this embodiment, the pixel information may be the pixel value of each pixel on the corresponding image. In this embodiment, the change degree of image information may be determined by calculating any one of the mean, weighted mean, variance, etc., of corresponding pixels of the new sample image and the at least a portion of the sample images in the sample image set. It can be understood that the above methods for extracting image information are merely examples, and the present disclosure does not impose limitations on specific methods for extracting image information. For example, in a case that the image information is color information, respective color histograms of the new sample image and the at least a portion of the sample images in the sample image set may be respectively obtained, and the change degree of image information of the new sample image may be determined by comparing the differences between the color histogram of the new sample image and the color histograms of the at least a portion of the sample images in the sample image set.

[0103] In the foregoing technical solution, when the change degree of image information of the new sample image relative to the at least a portion of the sample images in the sample image set exceeds the preset degree threshold, the new sample image is added to the sample image set to determine the feature vector library based on the sample image set after addition, thereby ensuring the accuracy of image detection results based on the feature vector library. In addition, this incremental training method further expands the application scope of the image detection method.

[0104] According to another aspect of the present disclosure, an electronic device is provided. FIG. 5 shows a schematic block diagram of an electronic device according to an embodiment of the present disclosure. As shown in FIG. 5, the control device 500 includes a processor 510 and a memory 520. A computer program is stored in the memory 520. The processor 510 is configured to execute the computer program to implement the image detection method 100.

[0105] In a possible implementation, the processor may include any suitable processing device with data processing capability and/or instruction execution capability. For example, the processor may be implemented using one or a combination of several of the following: Programmable Logic Controller (PLC), Digital Signal Processor (DSP), Field Programmable Gate Array (FPGA), Programmable Logic Array (PLA), Central Processing Unit (CPU),

Application-Specific Integrated Circuit (ASIC), Microcontroller Unit (MCU), and other forms of processing units.

[0106] According to another aspect of the embodiment of the present disclosure, a storage medium is further provided. A computer program/instruction is stored in the storage medium, and when the computer program/instruction is executed by a processor, the aforementioned image detection method 100 is implemented. The storage medium may include, for example, Read-Only Memory (ROM), Erasable Programmable Read-Only Memory (EPROM), portable Compact Disc Read-Only Memory (CD-ROM), USB memory, or any combination of the aforementioned storage media. The computer-readable storage medium may be any combination of one or more computer-readable storage media.

[0107] Those skilled in the art, by reading the relevant description of the image detection method 100 above, can understand the specific implementation solutions of the aforementioned electronic device and storage medium. For the sake of brevity, they will not be elaborated here.

[0108] Embodiment 1. An image detection method is provided, and the image detection method comprises:

[0109] obtaining a to-be-detected image;

[0110] inputting the to-be-detected image into a feature extraction model for feature extraction to obtain a to-be-detected feature vector set corresponding to a to-be-detected region in the to-be-detected image, where the to-be-detected region corresponds to a feature vector library, the feature vector library comprises at least one reference feature vector set, the at least one reference feature vector set corresponds to at least one reference region in a one-to-one manner, and the to-be-detected region corresponding to the feature vector library and the at least one reference region all correspond to a same image position;

[0111] comparing the to-be-detected feature vector set corresponding to the to-be-detected region with the reference feature vector set stored in the feature vector library corresponding to the to-be-detected region to determine whether the to-be-detected region meets a first preset condition, where the first preset condition is that a similarity between the to-be-detected feature vector set corresponding to the to-be-detected region and at least a portion of the reference feature vector set in the feature vector library is greater than a first similarity threshold; and

[0112] determining the to-be-detected region as an abnormal region in a case that the to-be-

detected region does not meet the first preset condition.

[0113] Embodiment 2. The image detection method described in embodiment 1 is provided, where the comparing the to-be-detected feature vector set corresponding to the to-be-detected region with the reference feature vector set stored in the feature vector library corresponding to the to-be-detected region to determine whether the to-be-detected region meets a first preset condition comprises:

[0114] calculating a first feature difference value between the to-be-detected feature vector set corresponding to the to-be-detected region and each reference feature vector set of the at least one reference feature vector set stored in the feature vector library corresponding to the to-be-detected region;

[0115] determining a smallest first feature difference value among the first feature difference value as a first anomaly score; and

[0116] determining whether the first anomaly score is greater than or equal to a first score threshold; and

[0117] determining that the to-be-detected region does not meet the first preset condition in a case that the first anomaly score is greater than or equal to the first score threshold,

[0118] where the first anomaly score is negatively correlated with the similarity, and the first score threshold is negatively correlated with the first similarity threshold.

[0119] Embodiment 3. The image detection method described in any one of embodiments 1 to 2 is provided, where the to-be-detected region is at least two in number, and the at least two to-be-detected regions correspond to at least two feature vector libraries in a one-to-one manner, where in a case that the to-be-detected regions do not meet the first preset condition, the image detection method further comprises:

[0120] for each to-be-detected region of the at least two to-be-detected regions,

[0121] comparing the to-be-detected feature vector set corresponding to the to-be-detected region with a second reference feature vector set stored in a second feature vector library corresponding to another to-be-detected region to determine whether the to-be-detected region meets a second preset condition, where the second preset condition is that a similarity between the to-be-detected feature vector set corresponding to the to-be-detected region and at least a portion of the second reference feature vector set in the second feature vector library is greater

than a second similarity threshold; and

[0122] determining the to-be-detected region as an offset-induced abnormal region in a case that the to-be-detected region meets the second preset condition.

[0123] Embodiment 4. The image detection method described in any one of embodiments 1 to 3 is provided, where the comparing the to-be-detected feature vector set corresponding to the to-be-detected region with a second reference feature vector set stored in a second feature vector library corresponding to another to-be-detected region to determine whether the to-be-detected region meets a second preset condition, comprises:

[0124] calculating a second feature difference value between the to-be-detected feature vector set corresponding to the to-be-detected region and each second reference feature vector set stored in the second feature vector library;

[0125] determining a smallest second feature difference value among the second feature difference value as a second anomaly score;

[0126] determining whether the second anomaly score is less than a second score threshold; and

[0127] determining that the to-be-detected region corresponding to the to-be-detected feature vector set meets the second preset condition in a case that the second anomaly score is less than the second score threshold,

[0128] where the second anomaly score is negatively correlated with the similarity.

[0129] Embodiment 5. The image detection method described in any one of embodiments 1 to 4 is provided, where before inputting the to-be-detected image into a feature extraction model for feature extraction, the image detection method further comprises:

[0130] matching a to-be-detected product in the to-be-detected image with a template product in a template image to obtain a first matching result;

[0131] adjusting an image parameter of the to-be-detected image based on the first matching result, to make the to-be-detected image the same size as the template image and the to-be-detected product in the to-be-detected image be aligned with the template product in the template image after the adjustment, where the image parameter comprises one or more of: position, shape, and size,

[0132] where the at least one reference region is an image region within at least one reference image, each reference image of the at least one reference image is the same size as the template image, and a reference product in the reference image is aligned with the template product in the template image.

[0133] Embodiment 6. The image detection method described in any one of embodiments 1 to 5 is provided, where the to-be-detected region is located at a target image position in the to-be-detected image; and the feature vector library is obtained through a feature vector library generation process, which comprises:

[0134] obtaining a sample image set, where the sample image set comprises at least one sample image, and each sample image of the at least one sample image in the sample image set does not comprise an abnormal region;

[0135] for each sample image of the at least one sample image in the sample image set, extracting an image feature vector corresponding to the sample image by using the feature extraction model;

[0136] determining, from the image feature vector corresponding to the at least one sample image in the sample image set, a feature vector set corresponding to a reference region located at the target image position, and obtaining a feature vector library corresponding to the target image position based on the determined feature vector set,

[0137] where the feature vector library corresponding to the target image position is the feature vector library corresponding to the to-be-detected region located at the target image position, the at least one reference region is an image region within at least one reference image, and the at least one reference image comprises at least a portion of the sample image in the sample image set.

[0138] Embodiment 7. The image detection method described in any one of embodiments 1 to 6 is provided, where the obtaining a feature vector library corresponding to the target image position based on the determined feature vector set comprises:

[0139] storing the determined feature vector set into the feature vector library corresponding to the target image position;

[0140] where the at least one reference region is the image region within the at least one reference image, and the at least one reference image comprises all sample images in the sample

image set.

[0141] Embodiment 8. The image detection method described in any one of embodiments 1 to 7 is provided, where the sample image is more than one in number, and the obtaining a feature vector library corresponding to the target image position based on the determined feature vector set comprises:

[0142] performing clustering on the determined feature vector sets to obtain at least two clusters;

[0143] storing a representative feature vector set comprised in each cluster of the at least two clusters into the feature vector library corresponding to the target image position,

[0144] where the representative feature vector set is any feature vector set in the cluster, each feature vector library corresponds to a reference region, and a reference image to which the reference region belongs is a sample image corresponding to a representative feature vector set stored in the feature vector library.

[0145] Embodiment 9. The image detection method described in any one of embodiments 1 to 8 is provided, where the feature vector library generation process further comprises:

[0146] obtaining a new sample image;

[0147] comparing image information of the new sample image with image information of the at least a portion of the sample image in the sample image set, to determine whether a change degree of the image information of the new sample image relative to the image information of the at least a portion of the sample image in the sample image set exceeds a preset degree threshold;

[0148] if the change degree of the image information of the new sample image relative to the image information of the at least a portion of the sample image in the sample image set exceeds the preset degree threshold, adding the new sample image to the sample image set, and determining the feature vector library based on the sample image set after the addition,

[0149] where the image information comprises at least one of: texture information, color information, pixel information.

[0150] Embodiment 10. An electronic device, comprising a processor and a memory, is provided. Computer program instructions are stored in the memory, and the computer program

instructions, when executed by the processor, performs the image detection method described in any one of embodiments 1 to 9.

[0151] Embodiment 11. A storage medium, storing program instructions, is provided. The program instructions, when executed, performs the image detection method described in any one of embodiments 1 to 9.

[0152] Although the embodiments have been described herein with reference to the drawings, it should be understood that the above embodiments are merely exemplary and are not intended to limit the scope of the present disclosure. Many changes and modifications may be made by those skilled in the art without departing from the scope and spirit of the present disclosure. All such changes and modifications are deemed to fall into the scope claimed by the claims of the present disclosure.

[0153] Those skilled in the art may realize that the units and algorithm steps of each example described in conjunction with the embodiments herein may be implemented by electronic hardware, or a combination of computer software and electronic hardware. Whether these functions are performed in hardware or software depends on the specific application and design constraints of the technical solution. Those skilled in the art may use different methods to implement the described functions for each specific application, but such implementation should not be regarded as exceeding the scope of the present disclosure.

[0154] It should be noted that the above embodiments illustrate rather than limit the present disclosure, and alternative embodiments may be designed by those skilled in the art without departing from the scope of the claims. In the claims, any reference symbol between parentheses shall not be construed as a limitation of the claims. The word "comprising" /"including" does not exclude the presence of elements or steps not listed in the claims. The word "a" or "an" preceding an element does not exclude the presence of multiple of such elements. The present disclosure may be implemented by means of hardware including several different elements and by means of a suitably programmed computer. In a unit claim enumerating several devices, several of these devices may be embodied by a same hardware item. The use of the words "first", "second", "third" and the like does not indicate any order. These words may be interpreted as names.

[0155] Described above are only embodiments of the present disclosure or explanation for the embodiments of the present disclosure, and the protection scope of the present disclosure is

not limited thereto. Any changes or alternations that can be easily envisaged by those skilled in the art within the technical scope disclosed in the present disclosure shall fall within the protection scope of the present disclosure. Therefore, the protection scope of the present disclosure shall be in accordance with the protection scope of the claims.

CLAIMS

1. An image detection method, comprising:

obtaining a to-be-detected image;

inputting the to-be-detected image into a feature extraction model for feature extraction to obtain a to-be-detected feature vector set corresponding to a to-be-detected region in the to-be-detected image, wherein the to-be-detected region corresponds to a feature vector library, the feature vector library comprises at least one reference feature vector set, the at least one reference feature vector set corresponds to at least one reference region in a one-to-one manner, and the to-be-detected region corresponding to the feature vector library and the at least one reference region all correspond to a same image position;

comparing the to-be-detected feature vector set corresponding to the to-be-detected region with the reference feature vector set stored in the feature vector library corresponding to the to-be-detected region to determine whether the to-be-detected region meets a first preset condition, wherein the first preset condition is that a similarity between the to-be-detected feature vector set corresponding to the to-be-detected region and at least a portion of the reference feature vector set in the feature vector library is greater than a first similarity threshold; and

determining the to-be-detected region as an abnormal region in a case that the to-be-detected region does not meet the first preset condition.

2. The image detection method according to claim 1, wherein the comparing the to-be-detected feature vector set corresponding to the to-be-detected region with the reference feature vector set stored in the feature vector library corresponding to the to-be-detected region to determine whether the to-be-detected region meets a first preset condition comprises:

calculating a first feature difference value between the to-be-detected feature vector set corresponding to the to-be-detected region and each reference feature vector set of the at least one reference feature vector set stored in the feature vector library corresponding to the to-be-detected region;

determining a smallest first feature difference value among the first feature difference value as a first anomaly score; and

determining whether the first anomaly score is greater than or equal to a first score

threshold, and determining that the to-be-detected region does not meet the first preset condition in a case that the first anomaly score is greater than or equal to the first score threshold, wherein the first anomaly score is negatively correlated with the similarity, and the first score threshold is negatively correlated with the first similarity threshold.

3. The image detection method according to claim 1, wherein the to-be-detected region is at least two in number, and the at least two to-be-detected regions correspond to at least two feature vector libraries in a one-to-one manner, wherein in a case that the to-be-detected regions do not meet the first preset condition, the image detection method further comprises:

for each to-be-detected region of the at least two to-be-detected regions,

comparing the to-be-detected feature vector set corresponding to the to-be-detected region with a second reference feature vector set stored in a second feature vector library corresponding to another to-be-detected region to determine whether the to-be-detected region meets a second preset condition, wherein the second preset condition is that a similarity between the to-be-detected feature vector set corresponding to the to-be-detected region and at least a portion of the second reference feature vector set in the second feature vector library is greater than a second similarity threshold; and

determining the to-be-detected region as an offset-induced abnormal region in a case that the to-be-detected region meets the second preset condition.

4. The image detection method according to claim 3, wherein the comparing the to-be-detected feature vector set corresponding to the to-be-detected region with a second reference feature vector set stored in a second feature vector library corresponding to another to-be-detected region to determine whether the to-be-detected region meets a second preset condition, comprises:

calculating a second feature difference value between the to-be-detected feature vector set corresponding to the to-be-detected region and each second reference feature vector set stored in the second feature vector library;

determining a smallest second feature difference value among the second feature difference value as a second anomaly score;

determining whether the second anomaly score is less than a second score threshold; and

determining that the to-be-detected region corresponding to the to-be-detected feature

vector set meets the second preset condition in a case that the second anomaly score is less than the second score threshold, wherein the second anomaly score is negatively correlated with the similarity.

5. The image detection method according to any one of claims 1 to 4, wherein before inputting the to-be-detected image into a feature extraction model for feature extraction, the image detection method further comprises:

matching a to-be-detected product in the to-be-detected image with a template product in a template image to obtain a first matching result;

adjusting an image parameter of the to-be-detected image based on the first matching result, to make the to-be-detected image the same size as the template image and the to-be-detected product in the to-be-detected image be aligned with the template product in the template image after the adjustment, wherein the image parameter comprises one or more of: position, shape, and size,

wherein the at least one reference region is an image region within at least one reference image, each reference image of the at least one reference image is the same size as the template image, and a reference product in the reference image is aligned with the template product in the template image.

6. The image detection method according to any one of claims 1 to 4, wherein the to-be-detected region is located at a target image position in the to-be-detected image; and the feature vector library is obtained through a feature vector library generation process, which comprises:

obtaining a sample image set, wherein the sample image set comprises at least one sample image, and each sample image of the at least one sample image in the sample image set does not comprise an abnormal region;

for each sample image of the at least one sample image in the sample image set, extracting an image feature vector corresponding to the sample image by using the feature extraction model;

determining, from the image feature vector corresponding to the at least one sample image in the sample image set, a feature vector set corresponding to a reference region located at the target image position, and obtaining a feature vector library corresponding to the target image position based on the determined feature vector set,

wherein the feature vector library corresponding to the target image position is the feature vector library corresponding to the to-be-detected region located at the target image position, the at least one reference region is an image region within at least one reference image, and the at least one reference image comprises at least a portion of the sample image in the sample image set.

7. The image detection method according to claim 6, wherein the obtaining a feature vector library corresponding to the target image position based on the determined feature vector set comprises:

storing the determined feature vector set into the feature vector library corresponding to the target image position;

wherein the at least one reference region is the image region within the at least one reference image, and the at least one reference image comprises all sample images in the sample image set.

8. The image detection method according to claim 6, wherein the sample image is more than one in number, and the obtaining a feature vector library corresponding to the target image position based on the determined feature vector set comprises:

performing clustering on the determined feature vector sets to obtain at least two clusters;

storing a representative feature vector set comprised in each cluster of the at least two clusters into the feature vector library corresponding to the target image position,

wherein the representative feature vector set is any feature vector set in the cluster, each feature vector library corresponds to a reference region, and a reference image to which the reference region belongs is a sample image corresponding to a representative feature vector set stored in the feature vector library.

9. The image detection method according to claim 6, wherein the feature vector library generation process further comprises:

obtaining a new sample image;

comparing image information of the new sample image with image information of the at least a portion of the sample image in the sample image set, to determine whether a change degree of the image information of the new sample image relative to the image information of

the at least a portion of the sample image in the sample image set exceeds a preset degree threshold;

if the change degree of the image information of the new sample image relative to the image information of the at least a portion of the sample image in the sample image set exceeds the preset degree threshold, adding the new sample image to the sample image set, and determining the feature vector library based on the sample image set after the addition, wherein the image information comprises at least one of: texture information, color information, pixel information.

10. An electronic device, comprising a processor and a memory, wherein computer program instructions are stored in the memory, and the computer program instructions, when executed by the processor, performs the image detection method according to any one of claims 1 to 9.

11. A storage medium, storing program instructions, wherein the program instructions, when executed, performs the image detection method according to any one of claims 1 to 9.