

METHOD FOR DEFECT DETECTION AND RELATED APPARATUS, DEVICE, STORAGE MEDIUM AND COMPUTER PROGRAM PRODUCT

CROSS-REFERENCE TO RELATED APPLICATION

[0001] The present application is based on and claims priority to Chinese patent application No. 202010837709.2 and filed on August 19, 2020, the disclosure of which should be interpreted as being hereby incorporated by reference in its entirety.

TECHNICAL FIELD

[0002] The application relates, but is not limited, to the field of image processing, in particular to a method for defect detection and a related apparatus, a device, a storage medium and a computer program product.

BACKGROUND

[0003] At present, in order to improve the protection of devices in the environment, defects of the devices in the environment often need to be detected, so as to ensure that the device defects can be found in time and then maintenance is performed.

[0004] In the related art, defect detection of a device is based on a shot image of the device, and the image is manually checked to determine whether a defect exists. However, when the defect is detected manually, omission or false detection is very prone to occurring due to human negligence. Therefore, how to perform defect detection to reduce the defect omission rate and improve the accuracy in detection is a critical issue at present.

SUMMARY

[0005] Embodiments of the disclosure at least provide a method for defect detection and a related apparatus, a device, a storage medium, and a computer program product.

[0006] The embodiments of the disclosure provide a method for defect detection, including: performing first object detection on a to-be-detected image to obtain an object detection result, the object detection result including a first class of object and first location information of the first class of object in the to-be-detected image; acquiring, based on the first location information, an image area containing the first class of object; and performing defect detection on the image area to obtain a defect detection result related to the first class of object.

[0007] Therefore, the location of the first class of object in the to-be-detected image is firstly detected, and then defect detection is performed on the image area containing the first class of object. Since the defect detection is performed directly on the image area where the object is located, the defect detection area is reduced, and the detection omission rate and false detection rate can be reduced.

[0008] In some embodiments of the disclosure, performing the defect detection on the image area to obtain the defect detection result related to the first class of object includes at least one of: performing second object detection on the image area to obtain a first defect detection result, the first defect detection result including information of at least one first defect existing on the first class of object; or classifying the image area to obtain a second defect detection result, the second defect detection result including information of a second defect that the first class of object has.

[0009] Therefore, the image area containing the first class of object is detected in various ways, and different detection modes may be used for different first class of objects respectively, so that the defect detection is more targeted, the obtained defect detection result is more accurate. Alternatively, different detection modes may be employed for the same first class of object, and then the results obtained by the different defect detection modes may be combined to obtain final defect information of the to-be-

detected image, so that the defect detection result is more accurate.

[0010] In some embodiments of the disclosure, performing the defect detection on the image area to obtain the defect detection result related to the first class of object includes: in response to that the first class of object is a first subclass of object, executing the action of performing second object detection on the image area to obtain the first defect detection result; or in response to that the first class of object is a second subclass of object, executing the action of classifying the image area to obtain the second defect detection result.

[0011] Therefore, by employing different defect detection modes for different target objects, the defect detection is more flexible. In comparison with that the same detection mode is employed for all classes of objects, the scheme of employing different defect detection modes for different target objects is more targeted, enabling the defect detection result to be more accurate.

[0012] In some embodiments of the disclosure, the first defect detection result includes location information of each of the at least one first defect and a first probability of the first defect; and after performing the second object detection on the image area to obtain the first defect detection result, the method further includes at least one of following: for each of the at least one first defect, determining a defect area of the first defect in the to-be-detected image according to the location information of the first defect, and performing overlap suppression on an overlapping area between different defect areas; or for each of the at least one first defect, filtering away the first defect in response to that the first probability of the first defect meets a first filtering condition.

[0013] Therefore, by acquiring the probability of the first defect in the first subclass of object and the location information of the first defect in the image area containing the first subclass of object, finding the area corresponding to the first defect in the to-be-detected image based on the location information, and performing overlap suppression on an overlapping part, one first defect corresponds to one respective defect area, thus reducing the problem of subsequently processing one defect for multiple times and improving the accuracy of final output. By eliminating an area meeting the filtering condition, the number of areas to be processed subsequently can be reduced and the

processing efficiency can be improved.

[0014] In some embodiments of the disclosure, the first filtering condition is that the first probability is lower than a first probability threshold; and/or the operation of performing overlap suppression on the overlapping area between the different defect areas may include that: non-maximum suppression is performed on the overlapping area between the different defect areas.

[0015] Therefore, by eliminating areas with low defect probabilities, many unnecessary processing areas are reduced, and the processing efficiency is improved. Overlap suppression is performed by means of non-maximum suppression, so that the defect areas with higher defect probabilities can be retained, and the accuracy of defect detection can also be guaranteed while the processing efficiency is improved.

[0016] In some embodiments of the disclosure, the second defect detection result may include a second probability that the first class of object has the second defect. Classifying the image area to obtain the second defect detection result includes: classifying the image area to obtain a probability that the first class of object has each of preset defects; and adding probabilities of the preset defects to obtain the second probability that the first class of object has the second defect.

[0017] Therefore, the probabilities of various preset defects are combined, to determine the defect probability of the first class of object. The possibilities that the first class of object has various preset defects are widely considered, so that the defect probability of the first class of object is more accurate.

[0018] In some embodiments of the disclosure, performing the first object detection on the to-be-detected image to obtain the object detection result includes: performing the first object detection on the to-be-detected image using a first area detection network of a defect detection model, to obtain the object detection result. Performing the second object detection on the image area to obtain the first defect detection result includes: performing the second object detection on the image area using a second area detection network of the defect detection model, to obtain the first defect detection result, wherein the second area detection network has a smaller depth than the first area detection

network. Classifying the image area to obtain the second defect detection result includes: classifying the image area using a classification network of the defect detection model, to obtain the second defect detection result.

[0019] Therefore, by considering properties of different defects of different components, detection is performed using different network models by segments, so that the problem of high omission rate in a single network detection or classification device can be alleviated to a certain extent. At the same time, the network model applied to the second object detection has a smaller depth than the network model applied to the first object detection, so that the amount of calculation during the second object detection can be reduced and the detection efficiency can be improved.

[0020] In some embodiments of the disclosure, the object detection result further includes a second class of object and second location information of the second class of object in the to-be-detected image. After performing the first object detection on the to-be-detected image to obtain the object detection result, the method further includes: determining the second class of object as a third defect in the to-be-detected image, and taking the second class of object and the second location information as a defect detection result related to the second class of object.

[0021] Therefore, by setting the second class of object that should not exist as the third defect, one defect detection model can detect multiple kinds of objects, which enhances the adaptability of the defect detection model.

[0022] In some embodiments of the disclosure, after obtaining the defect detection result, the method further includes: for each defect in the defect detection result, filtering away the defect in response to that the defect meets a second filtering condition, to obtain a final defect detection result of the to-be-detected image.

[0023] Therefore, by filtering away defects that do not meet the condition, the obviously unreasonable defects are reduced, and the final defect detection result is more accurate.

[0024] In some embodiments of the disclosure, for each defect in the defect detection result, filtering away the defect in response to that the defect meets the second filtering

condition includes at least one of: for each defect in the defect detection result, obtaining a size of the defect using the location information of the defect in the defect detection result, and filtering away the defect in the defect detection result in response to that the size of the defect does not meet a preset size condition; or for each defect in the defect detection result, filtering away the defect in response to that a probability of the defect is lower than a second probability threshold.

[0025] Therefore, by filtering away defects of which a size or probability does not meet the conditions, the obviously unreasonable defects are reduced, and the final defect detection result is more accurate.

[0026] In some embodiments of the disclosure, after obtaining the final defect detection result of the to-be-detected image, the method further includes: outputting information of each defect in the final defect detection result in a sequence of probabilities of all defects in the final defect detection result from high to low.

[0027] Therefore, by outputting the information of defects in the final defect detection result according to the sequence of the probabilities of the defects from high to low, an output result is neater and further observation is facilitated.

[0028] In some embodiments of the disclosure, acquiring the image area containing the first class of object based on the first location information includes: determining, based on the first location information, a target area corresponding to the first class of object in the to-be-detected image; expanding the target area outwards in the to-be-detected image by a preset multiple; and extracting the expanded target area from the to-be-detected image to obtain the image area containing the first class of object.

[0029] Therefore, by expanding the target area outwards in the to-be-detected image by the preset multiple, some background information can be retained after the corresponding target area containing the first class of object is acquired, and the accuracy of defect detection on the first class of object can be improved.

[0030] In some embodiments of the disclosure, before performing the first object detection on the to-be-detected image to obtain the object detection result, the method further includes: scaling an initial image to a preset size; compressing a pixel value of the

scaled initial image to a preset pixel value range; and normalizing the compressed initial image to obtain the to-be-detected image.

[0031] Therefore, by processing the size or pixels of the acquired initial image to obtain images of a uniform style, the robustness of the input image is improved to a certain extent.

[0032] In some embodiments of the disclosure, the first object detection and the defect detection are executed by a defect detection model; and the defect detection model is trained at least by: acquiring a first sample set, wherein the first sample set includes at least one sample image labeled with real defect information related to an object, and the object includes a first class of object; detecting, by the defect detection model, the at least one sample image in the first sample set, to obtain the defect detection result related to the object; and adjusting a parameter of the defect detection model using the real defect information and the defect detection result related to the object.

[0033] Therefore, the accuracy of the defect detection model trained in the above mode is higher, so that the detection result obtained through the first object detection and defect detection are executed to make is more accurate.

[0034] In some embodiments of the disclosure, the defect detection model is trained by a plurality of graphics processors jointly, each graphics processor acquiring a respective different first sample set; detecting, by the defect detection model, the at least one sample image in the first sample set to obtain the defect detection result related to the object includes: synchronizing a batch standardization layer of the defect detection model in each graphics processor, and detecting the at least one sample image using the defect detection model with the synchronized batch standardization layer in each graphics processor, so as to obtain the defect detection result related to the object; and adjusting the parameter of the defect detection model using the real defect information and the defect detection result related to the object includes: determining an average loss value of the defect detection model based on all defect detection results and all real defect information; and adjusting the parameter of the defect detection model using the average loss value.

[0035] Therefore, by synchronizing batch standardization layers of the defect detection models on the multiple graphics processors, a global first sample set may be used for normalization, which is equivalent to increasing batch size, thus reducing the influence of using multiple graphics processors. The training speed is increased by training the defect detection models through the multiple graphics processors.

[0036] In some embodiments of the disclosure, before detecting, by the defect detection model, the at least one sample image in the first sample set to obtain the defect detection result related to the object, the method further includes at least one of following: pre-training the defect detection model using a second sample set; or performing preheating training on the defect detection model using the first sample set, wherein a learning rate used in the preheating training starts from a first learning rate lower than a preset learning rate and gradually increases to the preset learning rate; wherein each of the at least one sample image in the first sample set is obtained by at least following: preprocessing an original image to obtain the sample image, wherein the preprocessing includes one or more of: scaling, color transformation, horizontal flip, vertical flip, rotation and tailoring.

[0037] Therefore, by pre-training the defect detection model using the second sample set before formal training of the defect detection model so that the parameters of the defect detection model are initialized, or by initializing the learning rate before formal training, and by preprocessing the original images, the adaptability of the defect detection model is improved.

[0038] Embodiments of the disclosure provide an apparatus for defect detection, including: a first object detection module, configured to perform first object detection on a to-be-detected image to obtain an object detection result, the object detection result including a first class of object and first location information of the first class of object in the to-be-detected image; an image area acquisition module, configured to acquire, based on the first location information, an image area containing the first class of object; and a defect detection module, configured to perform defect detection on the image area to obtain a defect detection result related to the first class of object.

[0039] Embodiments of the disclosure provide an electronic device, including a

memory, and a processor, wherein the processor is configured to execute program instructions stored in the memory to implement some or all of the actions of the above method for defect detection.

[0040] Embodiments of the disclosure provide a computer readable storage medium having stored thereon program instructions that, when executed by a processor, implement some or all of the actions of the method for defect detection.

[0041] Embodiments of the disclosure provide a computer program product including a non-transitory computer readable storage medium having stored thereon computer program that, when read and executed by a computer, implement some or all of the actions of the method described in the embodiments of the disclosure. The computer program product may be a software package.

[0042] In the embodiments of the disclosure, the location of the first class of object in the to-be-detected image is firstly detected, and then defect detection is performed on the image area containing the first class of object. Since the defect detection is performed directly on the image area where the object is located, the defect detection area is reduced, and the detection omission rate and false detection rate can be reduced.

[0043] It is to be understood that the above general descriptions and the following detailed descriptions are only exemplary and explanatory, and do not limit the application.

BRIEF DESCRIPTION OF THE DRAWINGS

[0044] The drawings herein are incorporated into and constitute a part of the specification, which illustrate embodiments in accordance with the disclosure and together with the specification serve to explain the technical solutions of the embodiments of the disclosure.

[0045] FIG. 1a illustrates a schematic flowchart of a method for defect detection according to embodiments of the disclosure.

- [0046] FIG. 1b illustrates a schematic state diagram of a method for defect detection according to embodiments of the disclosure.
- [0047] FIG. 2 illustrates a schematic diagram of a to-be-detected image in a method for defect detection according to embodiments of the disclosure.
- [0048] FIG. 3 illustrates a schematic diagram of an image area in a method for defect detection according to embodiments of the disclosure.
- [0049] FIG. 4a illustrates a schematic flowchart of implementation of a method for defect detection according to embodiments of the disclosure.
- [0050] FIG. 4b illustrates a schematic flowchart of implementation of an image preprocessing module according to embodiments of the disclosure.
- [0051] FIG. 4c illustrates a schematic diagram of a network structure of a primary component detection module according to embodiments of the disclosure.
- [0052] FIG. 4d illustrates a schematic flowchart of implementation of vibration damper classification and recognition according to embodiments of the disclosure.
- [0053] FIG. 4e illustrates a schematic diagram of a network structure of a secondary insulator self-explosion detection module according to embodiments of the disclosure.
- [0054] FIG. 5 illustrates a schematic structural diagram of an apparatus for defect detection according to embodiments of the disclosure.
- [0055] FIG. 6 illustrates a schematic structural diagram of an electronic device according to embodiments of the disclosure.
- [0056] FIG. 7 illustrates a schematic structural diagram of a computer readable storage medium according to embodiments of the disclosure.

DETAILED DESCRIPTION

[0057] Solutions of the embodiments of the disclosure will now be described in detail in combination with the accompanying drawings.

[0058] In the following descriptions, specific details such as a specific system structure, an interface, and a technology are set forth for purposes of illustration rather than limitation, so as to provide a thorough understanding of the embodiments of the disclosure.

[0059] In the disclosure, term “and/or” is only an association relationship describing associated objects and represents that three relationships may exist. For example, A and/or B may represent three conditions: i.e., independent existence of A, existence of both A and B and independent existence of B. In addition, the character “/” herein generally indicates that the contextual objects are in an "or" relationship. Furthermore, "multiple" herein means two or more than two. In addition, the term "at least one of" herein indicates any one of multiple or any combination of at least two of the multiple, for example, “including at least one of A, B and C” may indicate including any one or more elements selected from a set consisting of A, B and C.

[0060] In order to better understand the method for defect detection provided by the embodiments of the disclosure, the method for defect detection in a related art will be described below.

[0061] Taking the defect detection in power grid transmission lines as an example, China has more than one million kilometers of high-voltage transmission lines, which need regular inspection, conservation and maintenance. However, since most of the high-voltage transmission lines are erected in unmanned areas, the inspection of power grid is performed by taking aerial photography of the transmission lines by an aircraft and manual observation of aerial videos by naked eyes. In order to improve efficiency, a component defect detection solution based on image processing and machine learning is employed in the related art. However, this solution has a high omission rate and high requirements for photographed videos. In the related art, an Unmanned Aerial Vehicle (UAV) equipped with a High-Definition (HD) camera may be employed to realize the aerial photography of the transmission lines, so that the overall acquisition and photographing quality is improved and the cost is reduced. However, due to that defect

detection is still performed manually, the omission rate and the false alarm rate are still high. With the breakthrough of deep learning models in the field of computer vision, the performance indexes of basic tasks of computer vision have been greatly improved. However, all these models are more directed at tasks related to academic data sets, and there is no detection and recognition solution designed specifically for the scene of power inspection. Due to the fact that aerial images of the UAV are large in scale, and a self-explosion area of an insulator, a vibration damper, a bird's nest and so on that need to be detected during power grid inspection are all very small targets, there will be a lot of false detections and omission when abnormality defects are directly taken as detection targets, and the detection effect is very poor.

[0062] The above method for defect detection in the related art has the following problems: 1) the training process is cumbersome, and the omission rate and the false detection rate are high; 2) inference is performed using a central processing unit (CPU), and the test speed is slow; 3) the algorithm is not improved for a power scenario; and 4) only one kind of component or defect can be detected and recognized, such as insulator self-explosion, and a rust vibration damper.

[0063] Referring to FIG. 1a, FIG. 1a illustrates a flowchart of a method for defect detection according to embodiments of the disclosure. The method may include the following actions.

[0064] At S11: first object detection is performed on a to-be-detected image to obtain an object detection result. The object detection result includes a first class of object and first location information of the first class of object in the to-be-detected image.

[0065] In some embodiments of the disclosure, the to-be-detected image is firstly acquired before subjecting to the first object detection. Referring to FIG. 1b, FIG. 1b illustrates a state diagram of a method for defect detection according to embodiments of the disclosure. As illustrated in FIG. 1b, an image acquisition device acquires a to-be-detected image of various to-be-detected objects in a specific environment, for example, a to-be-detected image containing a first class of object, and then the image acquisition device transmits the to-be-detected image to a defect detection device. According to the method described in the embodiments of the disclosure, the defect detection device

performs defect detection on the to-be-detected image transmitted from the image acquisition device. The first class of object is a device that needs overhaul. In the embodiments of the disclosure, a device may become the first class of object as long as it is a real existing device. Taking traffic lights at an intersection, sewer manhole covers, and the like in a city as an example, the to-be-detected image which may contain the traffic lights or the manhole covers may be acquired through the image acquisition device, and then the to-be-detected image which may contain the traffic lights or the manhole covers is input to the defect detection device, so as to obtain the defect detection result of the traffic lights or the manhole covers.

[0066] For example, in a power grid transmission line, various devices in the line may need overhaul. In such a case, aerial images of various line devices in the line may be acquired by photography, such as an aerial image of an insulator string or a vibration damper, or an aerial image of a bird's nest. Next, the acquired aerial image is taken as the to-be-detected image and input into the defect detection device, and then the defect detection device may perform defect detection on the to-be-detected image according to the method described in the embodiments of the disclosure. The image acquisition device and the defect detection device may be integrated into one device, or may be multiple separate devices, which will not be specified here.

[0067] In some embodiments of the disclosure, before the to-be-detected image subjects the first object detection, an initial image may be scaled to a preset size. The preset size includes that the ratio of width to height is (1.1~1.7): 1 and/or that the side with smaller size is larger than 1000. In the embodiments of the disclosure, the ratio of width to height is 1.5: 1, and the minimum side is 1200. Then, the pixel values of the scaled initial image are compressed to a preset pixel value range. For example, the pixel value of the scaled initial image is compressed from 0-255 to 0-1. Then, the compressed initial image is normalized to obtain the to-be-detected image. The normalization includes normalizing the compressed initial image using the mean and variance (std) of an ImageNet data set, as illustrated in the following formula (1-1).

$$pixelvalue = \frac{(pixelvalue - mean)}{std} \quad (1-1)$$

- [0068] In the formula, “mean” represents the mean of the data set, “std” represents the variance of the data set, and “pixelvalue” represents the value of each pixel in the image.
- [0069] Of course, in some embodiments of the disclosure, other data sets, such as a Common Objects in Context (COCO) data set, may also be employed. The mode of normalization needs not to be limited to performing normalization by calling other data sets, but may be performed by other normally used modes of normalization. For example, the pixel values of all pixels in the initial image may be traversed, and a maximum value and a minimum value are recorded. The maximum value and the minimum value are used as parameters to normalize the initial image. A Sigmoid function may also be used to perform the normalization. Therefore, the embodiment of the disclosure does not specifically limit the mode of normalization. The acquired images are preprocessed as described above to unify image styles, so that the robustness of input can be improved.
- [0070] In the embodiment of the disclosure, the first object detection is performed on the to-be-detected image through a first area detection network of a defect detection model to obtain the object detection result.
- [0071] The defect detection model may be trained by the following operations.
- [0072] Firstly, a first sample set is acquired. The first sample set includes at least one sample image labeled with real defect information related to an object. The object here includes the first class of object, and of course may also include other classes of objects, such as a second class of object. Sample images in the first sample set may contain the first class of object and other classes of objects at the same time. The real defect information of the object may be information indicating that the object is normal and has no defect, or may be information indicating that the first class of object has a specific defect. That is to say, in the embodiments of the disclosure, the defect detection model may be trained using normal sample images and abnormal sample images containing defects, to obtain a defect detection model with strong adaptability. The sample images in the first sample set may be obtained by preprocessing original images. The preprocessing may be one or more of: scaling, color transformation, horizontal flip, vertical flip, rotation and tailoring. In some embodiments, in process of scaling the original image, the length and width of the original image are transformed according to a preset ratio. For

example, the ratio of length to width of the original image is scaled while keeping the ratio of 1.5: 1. Of course, in some embodiments of the disclosure, other aspect ratios may be kept, or the transformation may be made arbitrarily. Therefore, the aspect ratio of scaling the original image is not specified in the embodiments of the disclosure. The color transformation may be transformation of one or more of: brightness, saturation and chroma. For example, the brightness, saturation and chroma of the original image are all transformed, so that the same real defect information containing the same object is input into the defect detection model in different forms to train the defect detection model, so as to enhance the adaptability of the defect detection model and improve the accuracy of defect detection. In some embodiments, the above-mentioned modes of preprocessing may be used alone in the networks of different stages of the defect detection model, or may also be used in the networks of all stages.

[0073] Secondly, the sample images in the first sample set are detected by the defect detection model to obtain the defect detection result related to the object. The defect detection model may be trained by multiple graphics processors jointly, each graphics processor acquiring a respective different first sample set. The defect detection model is detected using multiple graphics processors so that the training speed is multiplied. In some embodiments, a batch standardization layer of the defect detection model in each graphics processor is synchronized, and the sample images are detected using the batch standardization layer synchronized defect detection model in each graphics processor, so as to obtain the defect detection result related to the object. One graphics processor collects the mean and variance of the batch standardization layers of the defect detection models in other graphics processors, so as to calculate a total mean and variance, and then this graphics processor returns the calculated mean and variance to the batch standardization layers of the defect detection models in the other graphics processors. The batch standardization layers of the defect detection models in multiple graphics processors are synchronized, so that a global first sample set may be used for normalization, which is equivalent to increasing batch size. Therefore, the influence of using multiple graphics processors is reduced.

[0074] Finally, the real defect information and defect detection result are used to adjust a parameter of the defect detection model. In some embodiments, based on all

defect detection results and all real defect information, an average loss value of the defect detection model is determined, and then the parameter of the defect detection model is adjusted using the average loss value. Training the defect detection model synchronously by the multiple graphics processors enables the parameters of the trained defect detection model in all the graphics processors to be consistent. Compared with a single graphics processor, the training progress is accelerated by using the multiple graphics processors to synchronously train the defect detection model.

[0075] In some embodiments of the disclosure, before the sample images in the first sample set are detected by the defect detection model to obtain the defect detection result related to the object, the defect detection model may also be pre-trained by a second sample set, and/or the defect detection model may subject to preheating training using the first sample set. A learning rate used during the preheating training starts from a first learning rate lower than a preset learning rate and gradually increases to the preset learning rate. For example, if the preset learning rate is 0.6, a learning rate may be gradually increased from the first learning rate of 0.3 to 0.4 and then continuously increased to the preset learning rate of 0.6 during preheating training. The second sample set may be a public training set, for example, a common COCO training set and an ImageNet data set. Samples used during preheating training may be the same as those used in formal training, that is, the first sample set is still used during the preheating training. In some embodiments, after the parameters are initialized through pre-training of the second sample set, the first sample set is used for preheating training to continuously adjust the learning rate. In the process of continuously adjusting the learning rate through preheating training, it may be considered that the parameters of the defect detection model are further optimized, so that the training can be performed on the better parameters during formal training. In some embodiments, the first sample set used in the preheating training process may still be the first sample set obtained after the above preprocessing, which can also improve the adaptability of the defect detection model, so that the trained defect detection model has higher detection accuracy. In some embodiments, gradually increasing from the first learning rate to the preset learning rate linearly or exponentially may be considered as initializing the learning rate, so that the defect detection model can be trained from a better learning rate to cause a better training

effect.

[0076] The first area detection network in the embodiments of the disclosure includes a Faster Region with Convolutional Neural Networks (R-CNN) network, such as ResNet50. Of course, in other embodiments, other neural network models, such as a Single Shot MultiBox Detector (SSD) network model, may be used for the first object detection of the to-be-detected image, and the depth need not to be limited to ResNet50, but may also be ResNet101 or ResNet200, which are not specifically limited here.

[0077] The first area detection network used in the embodiments of the disclosure is a ResNet50 network. The to-be-detected image is input into a head network of the first area detection network, to extract a feature map of the to-be-detected image. The size of the feature map is one sixteenth of that of the to-be-detected image. Then candidate boxes are extracted from the feature map through Region Proposal Networks (RPN). The candidate boxes refer to areas where the first class of object may exist. Extracting a candidate box essentially means extracting the pixel values of four vertices of the candidate box. Then the candidate boxes and the feature map together pass through a pooling layer to obtain the feature map of each candidate box. This operation specifically includes that the information of four coordinates of the candidate box is sent into the pooling layer, and the coordinates of the four vertices of the candidate box are mapped to the feature map in the pooling layer. Then a feature vector of each candidate box is obtained through a fully connected layer. These feature vectors represent the feature information of the candidate boxes. Then the feature vectors are subjected to bounding box regression and are classified according to preset categories, to obtain the final object detection result. It is to be seen that the first object detection refers to detecting whether the to-be-detected image contains the first class of object. If the first class of object exists, the first location information of the first class of object in the to-be-detected image is further acquired. The confidence of the first class of object may also be further learned, that is, the probability that the detected first class of object is indeed a first class of object.

[0078] The object detection result includes the first class of object and the first location information of the first class of object in the to-be-detected image, and/or the probability that the candidate box belongs to the first class of object. In some

embodiments, the first location information includes the location information of the candidate box corresponding to the first class of object in the to-be-detected image. The first location information refers to the location information of the four vertices of the candidate box corresponding to the first class of object. The first class of object may also include a first subclass of object and a second subclass of object. The first subclass of object and the second subclass of object may not have a subordination relation with the first class of object, but may have the same attributes as the first class of object. For example, the first class of object may be a kind of device, while the first subclass of object may be another kind of device, and the second subclass of object may be other devices. The same attribute means that whether the device has a defect or not can be learned by taking pictures and then analyzing the pictures. The probability of being the first class of object may further include the probability of being the first subclass of object and/or the probability of being the second subclass of object.

[0079] In some embodiments of the disclosure, the object detection result further includes a second class of object and second location information of the second class of object in the to-be-detected image, and/or the probability of being the second class of object. The second class of object here may be an object that should not exist, and its existence is a defect. Of course, in some embodiments, if the to-be-detected image is obtained after processing the initial image, the object detection result obtained after the first object detection on the to-be-detected image includes the location information of the first class of object and/or the second class of object in the initial image. By further setting the second class of object, the defect status of multiple different target objects can be detected by the same algorithm, and there is no need to use multiple algorithms to detect the defects of different components or devices in the same image, which greatly simplifies the operation of defect detection.

[0080] By way of example, the initial image is an aerial image of a transmission line taken by a UAV, or a graphic frame cut from a high-resolution (HD) video. Firstly, the initial image is processed, which includes scaling the initial image to be an image of 1200*1800; then the pixel values of the scaled initial image are compressed to a preset pixel value range of 0-1; and next the compressed initial image is normalized using the mean and variance (std) of the ImageNet data set, to obtain the to-be-detected image.

Then, the to-be-detected image is input into the ResNet50 network to acquire the probability that the target object in each candidate box in the to-be-detected image corresponds to each preset target category, and the preset target category with the highest probability is selected as the category of the target object in the candidate box. The location information of each candidate box may also be output at the same time. In the transmission line, usual components include vibration dampers and insulator strings, and may also include an external defect, that is, the second class of object which should not exist and whose existence is a defect, namely a bird's nest. In the embodiments of the disclosure, the vibration damper and the insulator string are included in the first class of object, and the bird's nest is included in the second class of object. That is, after the to-be-detected image is input into the ResNet50, the output object detection result may include the vibration damper, the insulator or the bird's nest, as well as the corresponding location information in the to-be-detected image and relevant probabilities thereof. Here, the probability refers to the probability that a candidate box belongs to a vibration damper, an insulator string or a bird's nest.

[0081] In some embodiments of the disclosure, after the object detection result is obtained, the second class of object is determined as a third defect in the to-be-detected image, and the second class of object and the second location information are taken as the defect detection result related to the second class of object. With the second class of object being a bird's nest, as an example, if a bird's nest is detected in the to-be-detected image, the bird's nest is determined as the third defect in the to-be-detected image, and the second location information of the bird's nest in the to-be-detected image is acquired. Therefore, the final object detection result contains the bird's nest and the second location information of the bird's nest.

[0082] At S12: an image area containing the first class of object is acquired based on the first location information.

[0083] In the embodiments of the disclosure, based on the first location information, a target area corresponding to the first class of object is determined from the to-be-detected image. It specifically includes that an area corresponding to the first class of object is determined from the to-be-detected image, that is, the candidate box corresponding to the

first class of object is determined. Next, the target area is expanded outwards in the to-be-detected image by a preset multiple. The expanding mode includes that the center point of the target area is unchanged, and the length and width of the target area become 1.1-1.5 times of an original target area. In the embodiments of the disclosure, the length and width of the target area are changed to about 1.2 times of the original target area. If the boundary of the target area exceeds the boundary of the to-be-detected image, the part within the to-be-detected image is retained. Then, the expanded target area is extracted from the to-be-detected image to obtain the image area containing the first class of object. Specifically, the expanded target area is cut from the to-be-detected image.

[0084] Meanwhile, referring to FIG. 2 and FIG. 3, FIG. 2 illustrates a schematic diagram of a to-be-detected image in a method for defect detection according to embodiments of the disclosure. FIG. 3 illustrates a schematic diagram of an image area in a method for defect detection according to embodiments of the disclosure. As illustrated in FIG. 2 and FIG. 3, the to-be-detected image 1 includes a first class of object 100. The location information of the first class of object 100 is obtained through a first object detection network, and then an image area 110 containing the first class of object 100 is acquired based on the location information of the first class of object 100. In order to explain the relationship between the image area 110 and a candidate box 101 more clearly, the candidate box 101 of the first class of object 100 is illustrated in the image area in FIG. 3. Since the candidate box 101 of the first class of object 100 is expanded in a certain proportion in the to-be-detected image 1 when the image area 110 containing the first class of object 100 is acquired from the to-be-detected image 1, the acquired image area 110 is larger than the area where the original candidate box 101 is originally located.

[0085] In some embodiments of the disclosure, if the detected image is obtained by performing scaling, pixel value compression and normalization on the initial image, acquiring the image area containing the first class of object based on the first location information refers to determining the target area corresponding to the first class of object in the initial image based on the first location information of the first class of object in the initial image. That is, the location of the candidate box corresponding to the first class of object in the initial image is determined, and the target area in the initial image is

expanded outwards by a preset multiple; and the target area having been expanded outwards is extracted from the initial image, so as to obtain the image area containing the first class of object. Of course, since the to-be-detected image corresponds to the initial image, in some other embodiments of the disclosure, even if the detected image is obtained after a series of processing including scaling, pixel values compression and normalization is performed on the initial image, the specific operation of acquiring the image area containing the first class of object based on the first location information may still include that the image area containing the first class of object is extracted from the to-be-detected image. Therefore, the image area containing the first class of object may be acquired from the to-be-detected image or the initial image, which is not specifically limited here.

[0086] Therefore, by expanding the target area outwards in the to-be-detected image by the preset multiple, some background information can be retained after the corresponding target area containing the first class of object is acquired, and the accuracy of defect detection on the first class of object can be improved.

[0087] In some embodiments of the disclosure, after the image area containing the first class of object is acquired, it needs to preprocess the image area, that is, to adjust the size of the image area. The aspect ratio of the image area may be adjusted to 1:1. If the image area is obtained from the initial image, the image area needs to be normalized. If the image area is obtained from a normalized to-be-detected image, the normalization may not be performed here.

[0088] For example, based on the first location information of the insulator string in the to-be-detected image, the candidate box corresponding to the insulator string in the to-be-detected image is determined. The candidate box is expanded outwards by 1.2 times in the to-be-detected image, and then the expanded candidate box is extracted from the to-be-detected image to obtain the image area containing the insulator.

[0089] At S13, defect detection is performed on the image area to obtain a defect detection result related to the first class of object.

[0090] The way of performing defect detection on the image area includes at least one

of following. One way is to perform second object detection on the image area to obtain a first defect detection result. Herein, the first defect detection result includes information of at least one first defect existing on the first class of object. The other way is to classify the image area to obtain a second defect detection result. Herein, the second defect detection result includes information of a second defect that the first class of object has. The second object detection refers to further object detection performed on the first class of object. What is obtained through the first object detection is the location information of the first class of object, but the defect information of the first class of object is not directly detected; therefore, targeted defect detection is performed on the first class in the second object detection. Classifying the image area refers to classifying the first class of object in the image area according to preset defect categories, and the probability information that the first class of object belongs to each defect category is calculated.

[0091] The image area containing the first class of object is detected in various ways, and different detection modes may be employed for different first class of objects, so that the defect detection is more targeted, and the detection result is more accurate. Alternatively, different detection modes may be employed for the same first class of object, and then the results obtained by the different defect detection modes may be combined to obtain final defect information of the to-be-detected image, so that the defect detection result is more accurate.

[0092] In response to that the first class of object is a first subclass of object, the action of performing second object detection on the image area to obtain the first defect detection result is executed. In response to that the first class of object is a second subclass of object, the action of classifying the image area to obtain the second defect detection result is executed.

[0093] By employing different defect detection modes for different target objects, the defect detection is more flexible. In comparison with that the same detection mode is employed for all classes of objects, the scheme of employing different defect detection modes for different target objects is more targeted, enabling the defect detection result to be more accurate.

[0094] For example, the insulator string is preset as the first subclass of object, and the

vibration damper is preset as the second subclass of object. Therefore, when the first class of object is detected as the insulator string, the second object detection is performed on the image area corresponding to the insulator string to obtain the first defect detection result. When the first class of object detected in a to-be-detected picture is a vibration damper, the image area corresponding to the vibration damper is classified to obtain the second defect detection result. When it is detected that the first class of object includes both an insulator string and a vibration damper, the second object detection is performed on the image area corresponding to the insulator string and the image area corresponding to the vibration damper is classified. In some embodiments, the defect detection on both the insulator string and the vibration damper may be performed simultaneously, that is, the second object detection and the classification are performed simultaneously. Of course, in some embodiments, the second object detection and the image area classification may also be performed respectively on the same first subclass of object, and then the results obtained by the two processing modes are combined to finally obtain the first defect detection result of the first subclass of object.

[0095] The first defect detection result includes the location information of each first defect and a first probability of the first defect. The location information of the first defect refers to the location of the first defect in the image area containing the first class of object. The probability of the first defect refers to the confidence of the first defect obtained by detection, that is, the probability that the first defect is indeed a first defect. For example, the second object detection is performed on the image area containing the insulator string, and it is found that insulator self-explosion may occur at a location A on the insulator string. The defect of insulator self-explosion is a first defect, while the location A on the insulator string is second location information of insulator self-explosion, and the confidence of insulator self-explosion is the first probability of being insulator self-explosion.

[0096] In the embodiments of the disclosure, when the first class of object is the first subclass of object, the detection network used in second object detection on the image area is a second area detection network. The second area detection network has a smaller depth than the first area detection network. For example, when the first area detection network is ResNet50, the second area detection network may be ResNet18. Both the

network model used in the second object detection and the network model used in the first object detection are Faster R-CNN networks, so the specific process of the second object detection is not repeated here. The difference from the first object detection is that in the first object detection, the to-be-detected image that may contain different first class of objects will be detected, whether the to-be-detected image contains the first class of object is determined and the first location of the first class of object is acquired, while in the second object detection, only the image area where the first subclass of object is located is detected, specifically including that the probability that the first subclass of object has a defect is detected.

[0097] For example, after the image area containing the insulator string is input into the second area detection network ResNet18, whether the insulator self-explosion occurs in the insulator string will be detected. If yes, the output will be the location information of the insulator self-explosion in the insulator string and the probability information of self-explosion.

[0098] In the embodiments of the disclosure, after the second object detection is performed on the image area to obtain the first defect detection result, the first defect detection result may be processed. Processing the first defect detection result includes that a defect area of each first defect on the to-be-detected image is determined according to the location information of the first defect. Since the location information obtained by the second object detection is the location information of the first defect on the image area containing the first class of object, the defect area of the first defect on the to-be-detected image needs to be found by means of mapping. Then, preset processing is performed on the defect area. The preset processing includes at least one of the following operations: overlap suppression is performed on an overlapping area between different defect areas, and for each first defect, the first defect is filtered away in response to that the first probability of the first defect meets a first filtering condition. The operation of performing overlap suppression on the overlapping area between the different defect areas may further include that non-maximum suppression is performed on the overlapping area between the different defect areas. When the area of the overlapping areas reaches a preset value, the first probabilities of the overlapping defect areas are acquired, and the respective probabilities are compared to each other, so that the defect

area with higher probability is retained and the defect area with lower probability is discarded. The first filtering condition is that the first probability is lower than a first probability threshold. That is, when the probability of a first defect is lower than the first probability threshold, the defect area corresponding to this first defect is eliminated. At the same time, the location information and the first probability of the first defect are deleted from the first defect detection result, that is, the final first defect detection result is determined based on only the defect area obtained after preset processing.

[0099] By acquiring the probability of the first defect in the first subclass of object and the location information of the first defect in the image area containing the first subclass of object, finding the area corresponding to the first defect in the to-be-detected image based on the location information, and performing overlap suppression on an overlapping part, one first defect corresponds to one respective defect area, thus reducing the problem of subsequently processing one defect for multiple times and improving the accuracy of final output. By eliminating an area meeting the filtering condition, the number of areas to be processed subsequently can be reduced and the processing efficiency can be improved.

[00100] In some embodiments of the disclosure, by eliminating areas with low defect probabilities, many unnecessary processing areas are reduced, and the processing efficiency is improved. Overlap suppression is performed by means of non-maximum suppression, so that the defect areas with higher defect probabilities can be retained, and the accuracy of defect detection can also be guaranteed while the processing efficiency is improved.

[00101] When the first class of object is the second subclass of object, the second defect detection result includes the second probability that the first class of object has the second defect. In some embodiments, the second defect here refers to whether the second subclass of object has a defect or not. The second probability is the probability that the second subclass has a defect. Of course, in other embodiments, the second defect may include multiple different sub-defects. The second probability is the probability that the second subclass of object has each sub-defect. The process of classifying the image area includes that the image area is classified to obtain the probability that the first class of

object has each preset defect, that is, the probability that the second subclass of object has each preset defect. For example, when the first class of object is the second subclass of object, such as a vibration damper, the image area containing the vibration damper is classified to obtain the probabilities that the vibration damper is a normal vibration damper, a rust vibration damper, a torsional vibration damper, a damaged vibration damper and a fell-off vibration damper. For example, the probability of being a normal vibration damper is 0.1, the probability of being a rust vibration damper is 0, the probability of being a torsional vibration damper is 0.2, the probability of being a damaged vibration damper is 0.5, and the probability of being a fell-off vibration damper is 0. At this time, it may be concluded that the probability that the vibration damper has the preset defect, namely being a damaged vibration damper, is the highest. In some embodiments of the disclosure, it may be determined that the vibration damper is a damaged vibration damper by this probability. In the embodiments of the disclosure, after the probability that the first class of object has each preset defect is obtained, the probabilities of all preset defects are added to obtain a second probability that the first class of object has the second defect. That is, the four defect probabilities of the vibration damper are added together, and the probability that the vibration damper has a defect is learned to be 0.7. In such a case, the second defect is that the vibration damper has a defect, and the second probability is the probability that the vibration damper has a defect. The probabilities of various preset defects are combined, to determine the defect probability of the first class of object. The possibilities that the first class of object has various preset defects are widely considered, so that the defect probability of the first class of object is more accurate.

[00102] During classification of the image area, the network used is a classification network. The classification network may also have a smaller depth than the first area detection network. In such a case, the classification network may also be ResNet18, but it may also be a network with another depth. For example, when the first area detection network is ResNet101, the classification network may also be ResNet50. Of course, in some embodiments, the depth of the classification network may be the same as that of the first area detection network. Therefore, the selection of the classification network is not specifically limited here.

[00103] Considering properties of different defects of different components, detection is performed using different network models by segments, so that the problem of high omission rate in a single network detection or classification device can be alleviated to a certain extent. At the same time, the network model applied to the second object detection has a smaller depth than the network model applied to the first object detection, so that the amount of calculation during the second object detection can be reduced and the detection efficiency can be improved.

[00104] In some embodiments of the disclosure, after the defect detection result is obtained, for each defect in the defect detection result, the defect is filtered away in response to that the defect meets a second filtering condition. Here, the defect detection result may be a combination of the first defect detection result and the second defect detection result as well as the defect detection result related to the second class of object. The location information of each defect in the defect detection result is used to obtain the size of the defect, and for each defect in the defect detection result, the defect is filtered away in response to that the size of the defect does not meet a preset size condition. For example, the defect area of insulator self-explosion, or the area of the bird's nest or the area where an abnormal vibration damper is located that has a too large or too small size is filtered away. Alternatively, the defects whose probability is lower than a second probability threshold in the defect detection result are filtered away. Specifically, different thresholds may be set for different classes of objects respectively. For example, the threshold is set to 0.3 for the first subclass of object of the first object, is set to 0.2 for the second subclass of object, is set to 0.2 for the second class of object, and the like. Then, the defects in the defect detection result are filtered according to preset thresholds. Of course, the size and probability of the defects in the defect detection result may also be filtered at the same time to reduce the number of defects in the defect detection result. Finally, the information of each defect in the defect detection result is output according to the sequence of the probabilities of defects in the final defect detection result from high to low. The output defect information includes the location information and probability of the defect in the to-be-detected image, or the location information and corresponding probability of the defect in the initial image are directly output.

[00105] In some embodiments of the disclosure, the form of output may include

outputting the locations of the defects in the to-be-detected image and/or in the initial image and probabilities thereof in the form of characters, or directly identifying all the defects in the to-be-detected image and/or the initial image by means of labeling and marking their respective probabilities in identification boxes, or directly outputting the image area corresponding to each defect in the to-be-detected image and/or the initial image and marking the probability of the defect in the image area.

[00106] The defects in the defect detection result are filtered, so that the filtered defect detection result contains as less information as possible which belongs to normal situations or excludes obviously unreasonable situations, thus improving the rationality of the defect detection result and reducing the false detection rate.

[00107] According to the above-mentioned solution, the location of the first class of object in the to-be-detected image is firstly detected, and then defect detection is performed on the image area containing the first class of object. Since the defect detection is performed directly on the image area where the object is located, the defect detection area is reduced, and the detection omission rate and false detection rate can be reduced.

[00108] Those skilled in the art may understand that, in the above method of detailed description, the writing sequence of various actions does not mean a strict execution sequence or constitutes any limitation on the implementation process, and the specific execution sequence of actions should be determined by its function and possible internal logic. An executive subject of the method for defect detection may be an apparatus for defect detection, for example, the method for defect detection may be executed by terminal device or a server or other processing device. Herein, the terminal device may be User Equipment (UE), mobile device, a user terminal, a terminal, a cellular phone, a cordless phone, Personal Digital Assistant (PDA), a handheld device, a computing device, a vehicle-mounted device, a wearable device, etc. In some embodiments, the method for defect detection may be implemented in a manner that a processor calls computer readable instructions stored in a memory.

[00109] Hereinafter, an exemplary application of the embodiments of the disclosure in an actual application scenario will be described. The embodiments of the disclosure

provide a method for detecting multiple categories of defects in a power grid transmission line based on deep learning. A deep learning model is used to perform two-stage detection and classification on large-scale input pictures. Through the method, abnormal defects such as insulator self-explosion detection, vibration damper defect detection and bird's nest detection may be performed jointly. Referring to FIG. 4a, the implementation flow of the method mainly includes the following functional modules: an RGB image input module 410, an image preprocessing module 420, a primary component detection module 430, a secondary vibration damper classification and recognition module 440, a secondary insulator self-explosion detection module 450, a self-explosion detection post-processing module 460 and an abnormal defect screening module 470. Each module is described in detail below.

- [00110] 1) The RGB image input module: this module may acquire data output shot by different UAV cameras, and obtain aerial RGB images or HD videos photographed by the UAV. For a HD video, it may be divided into image frames to obtain the corresponding RGB images.

- [00111] 2) The image preprocessing module: as illustrated in FIG. 4b, this module may perform uniform processing on the RGB image to obtain a to-be-input image for defect detection. The processing includes: a) image scaling 421: the input HD RGB image is scaled to be an image of 1200×1800; b) pixel value compression 422: the pixel value of the scaled image is compressed from 0-255 to 0-1; and c) pixel value normalization 423: the mean and variance std of an ImageNet data set are used to normalize the pixel value compressed image according to the aforementioned formula (1-1).

- [00112] 3) The primary component detection module: the input of this module is the normalized RGB image, and three types of primary components i.e., a bird's nest, an insulator string and a vibration damper are detected through an object detection network Faster R-CNN, so as to obtain three candidate areas i.e., an insulator string area, a vibration damper area and a bird's nest area, and the corresponding score of each candidate area. The main network structure of the module is as illustrated in FIG. 4c. A ResNet50 head network 431 is used as a backbone network of the detection network, to extract a C4 layer feature map 432. The flow of performing primary component detection

by this module includes that: the RGB image passes through the ResNet50 head network 431 to obtain the C4 layer feature map 432, and the size of the feature map is 1/16 of the input image. The ResNet50 head network 431 is composed of at least one convolution layer, at least one normalization layer and a downsampling layer which are stacked continuously, and performs forward operation in the way of layer-skipping connection. The C4 layer feature map 432 passes through an RPN to extract candidate boxes 433, and an extracted candidate box is an area where an object may exist. The candidate boxes and the C4 layer feature map together pass through an RoI pooling layer 434 to obtain a feature map of each candidate area, and then pass through at least one fully connected layer 435 to obtain a feature vector of each candidate area. The feature vectors of the candidate areas represent the feature information of the candidate areas. Finally, the feature vector of each candidate area is subjected to box coordinate regression 436 and multi-category classification 437 to obtain a final detection result, and the original-picture location areas and corresponding probability values of the bird's nest, the insulator string and the vibration damper are output. The bird's nest area is directly sent to the abnormal defect screening module, the insulator string area is sent to the secondary self-explosion detection module for self-explosion detection, and the vibration damper area is sent to the secondary vibration damper classification and recognition module for vibration damper category classification.

- [00113] 4) The secondary vibration damper classification and recognition module: this module takes the vibration damper area detected by the primary component detection module as input, mainly performs, through a ResNet18 classification network, multi-classification of abnormality classes on the vibration damper area cut from an original image (the classified categories also include a normal category), and finally integrates the probabilities of defect categories to obtain a normal vibration damper and vibration dampers with different abnormalities. According to the form of the vibration damper, the module may divide the vibration damper into the following five categories: a normal vibration damper, a rust vibration damper, a torsional vibration damper, a damaged vibration damper and a fell-off vibration damper, among which the last four categories are defect categories. As illustrated in FIG. 4d, the flow of performing vibration damper classification and recognition by the module includes four sub-flows: cutting the

vibration damper area from the original image 441, image preprocessing 442, ResNet18 multi-classification 443, and defect category probability integration 444. For the vibration damper area having a probability exceeding a first specific threshold that is detected by the primary component detection module, firstly, the vibration damper area is cut out from the original image. Then, image preprocessing is performed on each sub-image of cut out vibration damper area to keep the scale consistent. Afterwards, each sub-image is input into the ResNet18 classification network for classification of the five categories, and the probability values corresponding to five categories of vibration damper form are calculated by softmax. Finally, probabilities of defect categories are integrated, that is, four probability values of being the vibration damper defects are summed as the probability that the vibration damper has a defect.

- [00114] 5) The secondary insulator self-explosion detection module: this module takes the insulator string area detected by the primary component detection module as input, and mainly performs, through the Faster R-CNN detection network, the self-explosion area detection on each sub-image of insulator string area, so as to obtain an insulator that may contain self-explosion. The main network structure of the module is as illustrated in FIG. 4e. A ResNet18 head network 451 is used as the backbone network of the detection network, which may reduce the calculation amount. Firstly, the insulator string areas whose probabilities exceed a second specific threshold are cut out, and each sub-image obtained is preprocessed to keep the scale consistent, so as to obtain the RGB images for detection. Then, each RGB image passes through the ResNet18 head network 451 to detect the self-explosion area of the insulator string, so as to obtain a C4 layer feature map 452. The C4 layer feature map 452 passes through the RPN to extract candidate boxes 453, and an extracted candidate box is an area where insulator string self-explosion may exist. The candidate boxes and the C4 layer feature maps together pass through an RoI pooling layer 454 to obtain the feature map of each candidate area, and then pass through at least one fully connected layer 455 to obtain the feature vector of each candidate area. Finally, the feature vector of each candidate area is subjected to box coordinate regression 456 and multi-category classification 457 to obtain the final detection result, and the self-explosion area of the insulator string and the corresponding probability value are output.

- [00115] 6) The self-explosion detection post-processing module: this module maps a self-explosion detection result in the sub-image of insulator string area back to be a result in an original image, and performs overlapping suppression on the self-explosion area having been detected for multiple times repeatedly, so as to obtain the location information of the insulator self-explosion area in the original image. The processing procedure of the module may be divided into the following three actions: a) coordinates of the self-explosion area in the sub-image of insulator string area are mapped back to an area in the original image; b) an overlapping self-explosion area is suppressed; and c) a self-explosion area box with a lower score is filtered away.
- [00116] 7) The abnormal defect screening module: this module screens the detection and recognition results of the bird's nest area, the abnormal vibration damper and the insulator self-explosion to obtain a final abnormal defect detection result. The processing procedure of the module may be divided into the following three actions: a) the bird's nest area and the self-explosion area are filtered according to their sizes, and oversized and undersized detection boxes are eliminated; b) for the bird's nest, the insulator self-explosion and the abnormal vibration damper, according to the thresholds of their respective categories, the results lower than the corresponding thresholds are eliminated; and c) all the retained detection results are overall sorted as output results according to the scores from high to low.
- [00117] In the embodiments of the disclosure, aerial high-resolution images shot by the UAV are used as input, and a lightweight deep learning detection and classification model is introduced to recognize several defect anomalies such as insulator self-explosion, vibration damper and bird's nest attachment jointly, which has the following beneficial effects. 1) In a power grid defect detection algorithm in the related art, image processing and machine learning methods are mainly used, and hand-crafted features need to be constructed for objects transcendentally, which has poor robustness and low accuracy and is easily affected by the shooting environment. In the embodiments of the disclosure, a deep learning model is based on, and features are autonomously learned by networks, so that the application scenarios are wider, the robustness is stronger and the accuracy is higher. 2) In the embodiments of the disclosure, a two-stage detection-detection and detection-classification algorithm proposed in combination with a circuit

patrol scenario is applicable to the aerial HD images shot by the UAV, and the defect characteristic that the insulator self-explosion relies on the detected insulator string is fully considered, which may solve the problem of a poor effect in detection and recognition of HD images when a single network detection or classification device is used in the related art. 3) In the related art, one can only detect and recognize the defects of one component, and it is inefficient to repeatedly input the original image into different algorithms to detect various defects. In the embodiments of the disclosure, detection and recognition of several defects are performed in one algorithm, so that the detection efficiency is greatly improved on the premise of ensuring accuracy.

[00118] Referring to FIG. 5, FIG. 5 illustrates a schematic structural diagram of an apparatus for defect detection according to embodiments of the disclosure. The apparatus 30 for defect detection includes a first object detection module 31, an image area acquisition module 32 and a defect detection module 33. The first object detection module 31 is configured to perform first object detection on a to-be-detected image to obtain an object detection result, the object detection result including a first class of object and first location information of the first class of object in the to-be-detected image. The area acquisition module is configured to acquire, based on the first location information, an image area containing the first class of object. The defect detection module 33 is configured to perform defect detection on the image area to obtain a defect detection result related to the first class of object.

[00119] According to the above-mentioned solution, the location of the first class of object in the to-be-detected image is firstly detected, and then defect detection is performed on the image area containing the first class of object. Since the defect detection is performed directly on the image area where the object is located, the defect detection area is reduced, and the detection omission rate and false detection rate can be reduced.

[00120] The apparatus 30 for defect detection further includes a preprocessing module (not shown).

[00121] In some embodiments of the disclosure, before the first object detection module 31 performs the first object detection on the to-be-detected image to obtain the

object detection result, the preprocessing module is configured to scale an initial image to obtain an image of a preset size, compress a pixel value of the scaled initial image to a preset pixel value range, and normalize the compressed initial image to obtain the to-be-detected image.

[00122] According to the above-mentioned solution, by preprocessing the acquired images to obtain images of a uniform style, the input robustness can be improved.

[00123] In some embodiments of the disclosure, the first object detection module 31 is configured to perform the first object detection on the to-be-detected image to obtain the object detection result includes that: the first object detection is performed on the to-be-detected image using a first area detection network of a defect detection model, to obtain the object detection result. The first object detection module 31 is configured to perform the second object detection on the image area to obtain the first defect detection result includes that: the second object detection is performed on the image area using a second area detection network of the defect detection model, to obtain the first defect detection result. The second area detection network has a smaller depth than the first area detection network. The first object detection module 31 is configured to classify the image area to obtain the second defect detection result includes that: the image area is classified using a classification network of the defect detection model, to obtain the second defect detection result.

[00124] According to the above-mentioned solution, by considering properties of different defects of different components, detection is performed using different network models by segments, so that the problem of high omission rate in a single network detection or classification device can be alleviated to a certain extent. At the same time, the network model applied to the second object detection has a smaller depth than the network model applied to the first object detection, so that the amount of calculation during the second object detection can be reduced and the detection efficiency can be improved

[00125] In some embodiments of the disclosure, the image area acquisition module 32 is configured to acquire the image area containing the first class of object based on the first location information includes that: a target area corresponding to the first class of

object in the to-be-detected image is determined, based on the first location information; the target area is expanded outwards in the to-be-detected image by a preset multiple; and the expanded target area is extracted from the to-be-detected image to obtain the image area containing the first class of object.

[00126] According to the above-mentioned solution, by expanding the target area outwards in the to-be-detected image by the preset multiple, some background information can be retained after the corresponding target area containing the first class of object is acquired, and the accuracy of defect detection on the first class of object can be improved.

[00127] In some embodiments of the disclosure, the defect detection module 33 is further configured to perform the defect detection on the image area to obtain the defect detection result related to the first class of object includes at least one of the following. Second object detection is performed on the image area to obtain a first defect detection result, the first defect detection result including information of at least one first defect existing on the first class of object. The image area is classified to obtain a second defect detection result, the second defect detection result including information of a second defect that the first class of object has.

[00128] According to the above-mentioned solution, the image area containing the first class of object is detected in various ways, and different detection modes may be used for different first class of objects respectively, so that the defect detection is more targeted, the obtained defect detection result is more accurate. Alternatively, different detection modes may be employed for the same first class of object, and then the results obtained by the different defect detection modes may be combined to obtain final defect information of the to-be-detected image, so that the defect detection result is more accurate.

[00129] In some embodiments of the disclosure, the defect detection module 33 is configured to perform the defect detection on the image area to obtain the defect detection result related to the first class of object includes that: in response to that the first class of object is a first subclass of object, the action of performing second object detection on the image area to obtain the first defect detection result is executed; or in

response to that the first class of object is a second subclass of object, the action of classify the image area to obtain the second defect detection result is executed.

[00130] According to the above-mentioned solution, by employing different defect detection modes for different target objects, the defect detection is more flexible. In comparison with that the same detection mode is employed for all classes of objects, the scheme of employing different defect detection modes for different target objects is more targeted, enabling the defect detection result to be more accurate.

[00131] In some embodiments of the disclosure, the first defect detection result includes location information of each of the at least one first defect and a first probability of the first defect. The defect detection module 33 is further configured to perform the following after the second object detection is performed on the image area to obtain the first defect detection result: for each of the at least one first defect, determining a defect area of the first defect in the to-be-detected image according to the location information of the first defect, and performing overlap suppression on an overlapping area between different defect areas; or for each of the at least one first defect, filtering away the first defect in response to that the first probability of the first defect meets a first filtering condition.

[00132] According to the above-mentioned solution, by acquiring the probability of the first defect in the first subclass of object and the location information of the first defect in the image area containing the first subclass of object, finding the area corresponding to the first defect in the to-be-detected image based on the location information, and performing overlap suppression on an overlapping part, one first defect corresponds to one respective defect area, thus reducing the problem of subsequently processing one defect for multiple times and improving the accuracy of final output. By eliminating an area meeting the filtering condition, the number of areas to be processed subsequently can be reduced and the processing efficiency can be improved.

[00133] In some embodiments of the disclosure, the first filtering condition is that the first probability is lower than a first probability threshold; and/or the operation of performing overlap suppression on the overlapping area between the different defect areas may include that: non-maximum suppression is performed on the overlapping area

between the different defect areas.

[00134] According to the above-mentioned solution, by eliminating areas with low defect probabilities, many unnecessary processing areas are reduced, and the processing efficiency is improved. Overlap suppression is performed by means of non-maximum suppression, so that the defect areas with higher defect probabilities can be retained, and the accuracy of defect detection can also be guaranteed while the processing efficiency is improved.

[00135] In some embodiments of the disclosure, the second defect detection result may include a second probability that the first class of object has the second defect. The defect detection module is configured to classify the image area to obtain the second defect detection result includes that: the image area is classified to obtain a probability that the first class of object has each of preset defects; and probabilities of the preset defects are added to obtain the second probability that the first class of object has the second defect

[00136] According to the above-mentioned solution, the probabilities of various preset defects are combined, to determine the defect probability of the first class of object. The possibilities that the first class of object has various preset defects are widely considered, so that the defect probability of the first class of object is more accurate.

[00137] In some embodiments of the disclosure, the first object detection module 31 is configured to perform the first object detection on the to-be-detected image to obtain the object detection result includes that: the first object detection is performed on the to-be-detected image using a first area detection network of a defect detection model, to obtain the object detection result. The defect detection module 33 is configured to perform the second object detection on the image area to obtain the first defect detection result includes that: the second object detection is performed on the image area using a second area detection network of the defect detection model, to obtain the first defect detection result. The second area detection network has a smaller depth than the first area detection network. The defect detection module 33 is configured to classify the image area to obtain the second defect detection result includes that: the image area is classified using a classification network of the defect detection model, to obtain the second defect detection result.

[00138] According to the above-mentioned solution, by considering properties of different defects of different components, detection is performed using different network models by segments, so that the problem of high omission rate in a single network detection or classification device can be alleviated to a certain extent. At the same time, the network model applied to the second object detection has a smaller depth than the network model applied to the first object detection, so that the amount of calculation during the second object detection can be reduced and the detection efficiency can be improved.

[00139] In some embodiments of the disclosure, the object detection result further includes a second class of object and second location information of the second class of object in the to-be-detected image. After performing the first object detection on the to-be-detected image to obtain the object detection result, the first object detection module 31 is configured to: determine the second class of object as a third defect in the to-be-detected image, and take the second class of object and the second location information as a defect detection result related to the second class of object

[00140] According to the above-mentioned solution, by setting the second class of object that should not exist as the third defect, one defect detection model can detect multiple kinds of objects, which enhances the adaptability of the defect detection model.

[00141] In some embodiments of the disclosure, after obtaining the defect detection result, the defect detection module 33 is further configured to: for each defect in the defect detection result, filter away the defect in response to that the defect meets a second filtering condition, to obtain a final defect detection result of the to-be-detected image.

[00142] According to the above-mentioned solution, by filtering away defects that do not meet the condition, the obviously unreasonable defects are reduced, and the final defect detection result is more accurate.

[00143] In some embodiments of the disclosure, the defect detection module 33 is configured to: for each defect in the defect detection result, filter away the defect in response to that the defect meets the second filtering condition includes at least one of following. For each defect in the defect detection result, a size of the defect is obtained

using the location information of the defect in the defect detection result, and the defect in the defect detection result is filtered away in response to that the size of the defect does not meet a preset size condition. For each defect in the defect detection result, the defect is filtered away in response to that a probability of the defect is lower than a second probability threshold.

[00144] According to the above-mentioned solution, by filtering away defects of which a size or probability does not meet the conditions, the obviously unreasonable defects are reduced, and the final defect detection result is more accurate.

[00145] In some embodiments of the disclosure, after obtaining the final defect detection result of the to-be-detected image, the defect detection module 33 is further configured to: output information of each defect in the final defect detection result in a sequence of probabilities of all defects in the final defect detection result from high to low.

[00146] According to the above-mentioned solution, by outputting the information of defects in the final defect detection result according to the sequence of the probabilities of the defects from high to low, an output result is neater and further observation is facilitated.

[00147] The apparatus for defect detection includes a training module (not shown). The training module is configured to train the defect detection model.

[00148] In some embodiments of the disclosure, the first object detection and the defect detection are executed by a defect detection model; and the defect detection model is trained at least by: acquiring a first sample set including at least one sample image labeled with real defect information related to an object, the object including a first class of object; detecting, by the defect detection model, the at least one sample image in the first sample set, to obtain the defect detection result related to the object; and adjusting a parameter of the defect detection model using the real defect information and the defect detection result related to the object.

[00149] According to the above-mentioned solution, the accuracy of the defect detection model trained in the above mode is higher, so that the detection result obtained

through the first object detection and defect detection are executed to make is more accurate.

[00150] In some embodiments of the disclosure, the defect detection model is trained by a plurality of graphics processors jointly, each graphics processor acquiring a respective different first sample set. The training module is configured to detect, by the defect detection model, the at least one sample image in the first sample set to obtain the defect detection result related to the object includes that: a batch standardization layer of the defect detection model in each graphics processor is synchronized, and the at least one sample image is detected using the defect detection model with the synchronized batch standardization layer in each graphics processor, so as to obtain the defect detection result related to the object. Adjusting the parameter of the defect detection model using the real defect information and the defect detection result related to the object includes: determining an average loss value of the defect detection model based on all defect detection results and all real defect information; and adjusting the parameter of the defect detection model using the average loss value.

[00151] According to the above-mentioned solution, by synchronizing batch standardization layers of the defect detection models on the multiple graphics processors, a global first sample set may be used for normalization, which is equivalent to increasing batch size, thus reducing the influence of using multiple graphics processors. The training speed is increased by training the defect detection models through the multiple graphics processors.

[00152] In some embodiments of the disclosure, before detecting, by the defect detection model, the at least one sample image in the first sample set to obtain the defect detection result related to the object, the training module is further configured to perform at least one of following. The defect detection model is pre-trained using a second sample set. Preheating training is performed on the defect detection model using the first sample set. A learning rate used in the preheating training starts from a first learning rate lower than a preset learning rate and gradually increases to the preset learning rate. Each of the at least one sample image in the first sample set is obtained by at least following: preprocessing an original image to obtain the sample image. The preprocessing includes

one or more of: scaling, color transformation, horizontal flip, vertical flip, rotation and tailoring.

[00153] Therefore, by pre-training the defect detection model using the second sample set before formal training of the defect detection model so that the parameters of the defect detection model are initialized, or by initializing the learning rate before formal training, and by preprocessing the original images, the adaptability of the defect detection model is improved.

[00154] Referring to FIG. 6, FIG. 6 illustrates a schematic structural diagram of an electronic device according to embodiments of the disclosure. The electronic device 40 includes a memory 41 and a processor 42. The processor 42 is configured to execute program instructions stored in the memory 41 so as to implement the actions of the method for defect detection embodiments. In an implementation scenario, the electronic device 40 may include, but is not limited to, a microcomputer and a server. In addition, the electronic device 40 may also include mobile devices such as a notebook computer and a tablet computer, which is not limited here.

[00155] Specifically, the processor 42 is configured to control the processor itself and the memory 41 to implement the actions in any of the method for defect detection embodiments. The processor 42 may also be called a Central Processing Unit (CPU). The processor 42 may be an integrated circuit chip with signal processing capabilities. The processor 42 may also be a general-purpose processor, a Digital Signal Processor (DSP), an Application Specific Integrated Circuit (ASIC), a Field-Programmable Gate Array (FPGA) or other programmable logic devices, a discrete gate or transistor logic device, or a discrete hardware component. The general-purpose processor may be a microprocessor or the processor may be any conventional processor, etc. Furthermore, the processor 42 may be jointly implemented by the integrated circuit chip.

[00156] According to the above-mentioned solution, the location of the first class of object in the to-be-detected image is firstly detected, and then defect detection is performed on the image area containing the first class of object. Since the defect detection is performed directly on the image area where the object is located, the defect detection area is reduced, and the omission rate and false detection rate may be reduced.

- [00157] Referring to FIG. 7, FIG. 7 illustrates a schematic structural diagram of a computer readable storage medium according to embodiments of the disclosure. The computer readable storage medium 50 stores program instructions 501 capable of being executed by a processor. The program instructions 501 are configured to implement the actions in the embodiments of the method for defect detection.
- [00158] The embodiments of the disclosure further provide a computer program product. When executed by a processor, the computer program product implements any method of the embodiments. The computer program product may be implemented by hardware, software or a combination thereof. In some embodiments of the present disclosure, the computer program product is embodied as a computer storage medium. In other embodiments of the disclosure, the computer program product is embodied as a software product, such as a Software Development Kit (SDK).
- [00159] According to the above-mentioned solution, the location of the first class of object in the to-be-detected image is firstly detected, and then defect detection is performed on the image area containing the first class of object. Since the defect detection is performed directly on the image area where the object is located, the defect detection area is reduced, and the omission rate and false detection rate may be reduced.
- [00160] In some embodiments, the functions or included modules of the apparatus provided by the embodiments of the disclosure may be configured to execute the method described in the above method embodiments, and its implementation may refer to the descriptions of the above method embodiments for brevity, which is not repeated here.
- [00161] The above descriptions of various embodiments tend to emphasize the differences between the various embodiments, and their similarities may be referred to each other for brevity, which is not repeated here.
- [00162] In the embodiments provided in the disclosure, it should be understood that the disclosed methods and apparatuses may be implemented in other ways. For example, the apparatus embodiment described above is only illustrative, and for example, division of the modules or units is only division in logic functions, and other division manners may be employed during practical implementation. For example, units or components may be

combined or integrated into another system, or some characteristics may be neglected or not executed. In addition, coupling or direct coupling or communication connection between displayed or discussed components may be indirect coupling or communication connection, implemented through some interfaces, device or units, and may be electrical and mechanical or in other forms.

[00163] In addition, function units in each embodiment of the disclosure may be integrated into a processing unit, or each unit may exist independently, or two or more than two units may be integrated into one unit. The integrated unit may be implemented in a hardware form and may also be implemented in form of software functional unit.

[00164] When being implemented in form of software functional unit and sold or used as an independent product, the integrated unit may be stored in a computer-readable storage medium. Based on such an understanding, the technical solutions of the disclosure essentially, or the part contributing to the related art, or all or part of the technical solutions may be implemented in a form of a software product. The computer software product is stored in a storage medium and includes several instructions for instructing a computer device (which may be a personal computer, a server, or a network device) or a processor to execute all or some of the actions of various implementations and methods described in the disclosure. The above-mentioned storage medium includes: various media capable of storing program codes such as a USB flash disk, a mobile hard disk, a Read-Only Memory (ROM), a Random Access Memory (RAM), a magnetic disk, or an optical disk.

INDUSTRIAL PRACTICABILITY

[00165] Embodiments of the disclosure provide a method for defect detection and a related apparatus, a device, and a storage medium. The method includes that: first object detection is performed on a to-be-detected image to obtain an object detection result, the object detection result including a first class of object and first location information of the first class of object in the to-be-detected image; an image area containing the first class of object is acquired based on the first location information; and defect detection is performed on the image area to obtain a defect detection result related to the first class of object. According to the above-mentioned solution, according to the method for defect

detection provided by the embodiments of the disclosure, the defect detection is performed on the to-be-detected image, which can reduce the defect detection omission rate and false detection rate.

CLAIMS

1. A method for defect detection, comprising:

performing first object detection on a to-be-detected image to obtain an object detection result, the object detection result comprising a first class of object and first location information of the first class of object in the to-be-detected image;

acquiring, based on the first location information, an image area containing the first class of object; and

performing defect detection on the image area to obtain a defect detection result related to the first class of object.

2. The method of claim 1, wherein performing the defect detection on the image area to obtain the defect detection result related to the first class of object comprises at least one of:

performing second object detection on the image area to obtain a first defect detection result, the first defect detection result comprising information of at least one first defect existing on the first class of object; or

classifying the image area to obtain a second defect detection result, the second defect detection result comprising information of a second defect that the first class of object has.

3. The method of claim 2, wherein performing the defect detection on the image area to obtain the defect detection result related to the first class of object comprises:

in response to that the first class of object is a first subclass of object, executing the action of performing second object detection on the image area to obtain the first defect detection result; or

in response to that the first class of object is a second subclass of object, executing the action of classifying the image area to obtain the second defect detection result.

4. The method of claim 2 or 3, wherein the first defect detection result comprises location information of each of the at least one first defect and a first probability of the first defect; and after performing the second object detection on the image area to obtain the first defect detection result, the method further comprises at least one of following:

for each of the at least one first defect, determining a defect area of the first defect in the to-be-detected image according to the location information of the first defect, and performing overlap suppression on an overlapping area between different defect areas; or

for each of the at least one first defect, filtering away the first defect in response to that the first probability of the first defect meets a first filtering condition.

5. The method of claim 4, wherein at least one of following is satisfied:

the first filtering condition is that the first probability is lower than a first probability threshold; or

performing overlap suppression on the overlapping area between the different defect areas comprises:

performing non-maximum suppression on the overlapping area between the different defect areas.

6. The method of any one of claims 2-5, wherein the second defect detection result comprises a second probability that the first class of object has the second defect; and classifying the image area to obtain the second defect detection result comprises:

classifying the image area to obtain a probability that the first class of object has each of preset defects; and

adding probabilities of the preset defects to obtain the second probability that the first class of object has the second defect.

7. The method of any one of claims 2-6, wherein performing the first object detection on the to-be-detected image to obtain the object detection result comprises:

performing the first object detection on the to-be-detected image using a first area detection network of a defect detection model, to obtain the object detection result;

wherein performing the second object detection on the image area to obtain the first defect detection result comprises:

performing the second object detection on the image area using a second area detection network of the defect detection model, to obtain the first defect detection result, wherein the second area detection network has a smaller depth than the first area detection network; and

wherein classifying the image area to obtain the second defect detection result comprises:

classifying the image area using a classification network of the defect detection model, to obtain the second defect detection result.

8. The method of any one of claims 1-7, wherein the object detection result further comprises a second class of object and second location information of the second class of object in the to-be-detected image; and after performing the first object detection on the to-be-detected image to obtain the object detection result, the method further comprises:

determining the second class of object as a third defect in the to-be-detected image, and taking the second class of object and the second location information as a defect detection result related to the second class of object.

9. The method of any one of claims 1-8, after obtaining the defect detection result, further comprising:

for each defect in the defect detection result, filtering away the defect in response to that the defect meets a second filtering condition, to obtain a final defect detection result of the to-be-detected image.

10. The method of claim 9, wherein for each defect in the defect detection result, filtering away the defect in response to that the defect meets the second filtering condition comprises at least one of:

for each defect in the defect detection result, obtaining a size of the defect using the location information of the defect in the defect detection result, and filtering away the defect in the defect detection result in response to that the size of the defect does not meet a preset size condition; or

for each defect in the defect detection result, filtering away the defect in response to that a probability of the defect is lower than a second probability threshold.

11. The method of claim 9 or 10, after obtaining the final defect detection result of the to-be-detected image, further comprising:

outputting information of each defect in the final defect detection result in a sequence of probabilities of all defects in the final defect detection result from high to low.

12. The method of any one of claims 1-11, wherein acquiring the image area containing the first class of object based on the first location information comprises:

determining, based on the first location information, a target area corresponding to the first class of object in the to-be-detected image;

expanding the target area outwards in the to-be-detected image by a preset multiple;
and

extracting the expanded target area from the to-be-detected image to obtain the image area containing the first class of object.

13. The method of any one of claims 1-12, before performing the first object detection on the to-be-detected image to obtain the object detection result, further comprising:

scaling an initial image to a preset size;

compressing a pixel value of the scaled initial image to a preset pixel value range; and

normalizing the compressed initial image to obtain the to-be-detected image.

14. The method of any one of claims 1-13, wherein the first object detection and the defect detection are executed by a defect detection model; and the defect detection model is trained at least by:

acquiring a first sample set, wherein the first sample set comprises at least one sample image labeled with real defect information related to an object, and the object comprises a first class of object;

detecting, by the defect detection model, the at least one sample image in the first sample set, to obtain the defect detection result related to the object; and

adjusting a parameter of the defect detection model using the real defect information and the defect detection result related to the object.

15. The method of claim 14, wherein the defect detection model is trained by a plurality of graphics processors jointly, each graphics processor acquiring a respective different first sample set;

wherein detecting, by the defect detection model, the at least one sample image in the first sample set to obtain the defect detection result related to the object comprises:

synchronizing a batch standardization layer of the defect detection model in each graphics processor, and detecting the at least one sample image using the defect detection model with the synchronized batch standardization layer in each graphics processor, so as to obtain the defect detection result related to the object; and

wherein adjusting the parameter of the defect detection model using the real defect information and the defect detection result related to the object comprises:

determining an average loss value of the defect detection model based on all defect detection results and all real defect information; and

adjusting the parameter of the defect detection model using the average loss value.

16. The method of claim 14 or 15, before detecting, by the defect detection model, the at least one sample image in the first sample set to obtain the defect detection result related to the object, further comprising at least one of following:

pre-training the defect detection model using a second sample set; or

performing preheating training on the defect detection model using the first sample set, wherein a learning rate used in the preheating training starts from a first learning rate lower than a preset learning rate and gradually increases to the preset learning rate;

wherein each of the at least one sample image in the first sample set is obtained by at least following:

preprocessing an original image to obtain the sample image, wherein the preprocessing comprises one or more of: scaling, color transformation, horizontal flip, vertical flip, rotation and tailoring.

17. An apparatus for defect detection comprising:

a first object detection module, configured to perform first object detection on a to-be-detected image to obtain an object detection result, the object detection result comprising a first class of object and first location information of the first class of object in the to-be-detected image;

an image area acquisition module, configured to acquire, based on the first location information, an image area containing the first class of object; and

a defect detection module, configured to perform defect detection on the image area to obtain a defect detection result related to the first class of object.

18. The apparatus of claim 17, wherein the defect detection module is further configured to perform at least one of:

performing second object detection on the image area to obtain a first defect detection result, the first defect detection result comprising information of at least one first defect existing on the first class of object; or

classifying the image area to obtain a second defect detection result, the second defect detection result comprising information of a second defect that the first class of object has.

19. The apparatus of claim 18, wherein the defect detection module is further configured to:

in response to that the first class of object is a first subclass of object, execute the action of performing second object detection on the image area to obtain the first defect detection result; or

in response to that the first class of object is a second subclass of object, execute the action of classifying the image area to obtain the second defect detection result.

20. The apparatus of claim 18 or 19, wherein the first defect detection result comprises location information of each of the at least one first defect and a first probability of the first defect; and after performing the second object detection on the image area to obtain the first defect detection result, the defect detection module is further configured to perform at least one of following:

for each of the at least one first defect, determining a defect area of the first defect in the to-be-detected image according to the location information of the first defect, and performing overlap suppression on an overlapping area between different defect areas; or

for each of the at least one first defect, filtering away the first defect in response to that the first probability of the first defect meets a first filtering condition.

21. The apparatus of claim 20, wherein at least one of following is satisfied:

the first filtering condition is that the first probability is lower than a first probability threshold; or

the defect detection module is further configured to: perform non-maximum suppression on the overlapping area between the different defect areas.

22. The apparatus of any one of claims 18-21, wherein the second defect detection result comprises a second probability that the first class of object has the second defect; and the defect detection module is further configured to:

classify the image area to obtain a probability that the first class of object has each of preset defects; and

add probabilities of the preset defects to obtain the second probability that the first class of object has the second defect.

23. The apparatus of any one of claims 18-22, wherein:

the first object detection module is further configured to: perform the first object detection on the to-be-detected image using a first area detection network of a defect detection model, to obtain the object detection result;

the defect detection module is further configured to: perform the second object detection on the image area using a second area detection network of the defect detection model, to obtain the first defect detection result, wherein the second area detection network has a smaller depth than the first area detection network; and classify the image area using a classification network of the defect detection model to obtain a second defect detection result.

24. The apparatus of any one of claims 17-23, wherein the object detection result further comprises a second class of object and second location information of the second class of object in the to-be-detected image; and after performing the first object detection on the to-be-detected image to obtain the object detection result, the defect detection module is further configured to:

determine the second class of object as a third defect in the to-be-detected image, and take the second class of object and the second location information as a defect detection result related to the second class of object.

25. The apparatus of any one of claims 17-24, wherein after obtaining the defect detection result, the defect detection module is further configured to:

for each defect in the defect detection result, filter away the defect in response to that the defect meets a second filtering condition, to obtain a final defect detection result of the to-be-detected image

26. The apparatus of claim 25, wherein the defect detection module is further configured to perform at least one of following:

for each detect in the defect detection result, obtain a size of the defect using the location information of the defect in the defect detection result, and filter away the defect in the defect detection result in response to that the size of the defect does not meet a preset size condition; or

for each detect in the defect detection result, filter away the defect in response to that a probability of the defect is lower than a second probability threshold.

27. The apparatus of claim 25 or 26, wherein after obtaining the final defect detection result of the to-be-detected image, the defect detection module is further configured to:

output the information of each defect in the final defect detection result in a sequence of probabilities of all defects in the final defect detection result from high to low.

28. The apparatus of any one of claims 17-27, wherein the image area acquisition module is further configured to:

determine, based on the first location information, a target area corresponding to the first class of object in the to-be-detected image;

expand the target area outwards in the to-be-detected image by a preset multiple; and

extract the expanded target area from the to-be-detected image to obtain the image area containing the first class of object.

29. The apparatus of any one of claims 17-28, further comprising a preprocessing module, configured to perform the following actions before the first object detection is performed on the to-be-detected image to obtain the object detection result:

scaling an initial image to a preset size;

compressing a pixel value of the scaled initial image to a preset pixel value range; and

normalizing the compressed initial image to obtain the to-be-detected image.

30. The apparatus of any one of claims 17-29, further comprising a training module, configured to:

acquire a first sample set, wherein the first sample set comprises at least one sample image labeled with real defect information related to an object, and the object comprises a first class of object;

detect, by the defect detection model, the at least one sample image in the first sample set, to obtain the defect detection result related to the object; and

adjust a parameter of the defect detection model using the real defect information and the defect detection result related to the object.

31. The apparatus of claim 30, wherein the defect detection model is trained by a plurality of graphics processors jointly, each graphics processor acquiring a respective different first sample set; and the training module is further configured to:

synchronize a batch standardization layer of the defect detection model in each graphics processor, and detecting the at least one sample image using the defect detection model with the synchronized batch standardization layer in each graphics processor, so as to obtain the defect detection result related to the object;

determine an average loss value of the defect detection model based on all defect detection results and all real defect information; and

adjust the parameter of the defect detection model using the average loss value.

32. The apparatus of claim 30 or 31, wherein before detecting, by the defect detection model, the at least one sample image in the first sample set to obtain the defect detection result related to the object, the training module is further configured to:

pre-train the defect detection model using a second sample set; or

perform preheating training on the defect detection model using the first sample set, wherein a learning rate used in the preheating training starts from a first learning rate lower than a preset learning rate and gradually increases to the preset learning rate;

wherein each of the at least one sample image in the first sample set is obtained by at least following:

preprocessing an original image to obtain the sample image, wherein the preprocessing comprises one or more of: scaling, color transformation, horizontal flip, vertical flip, rotation and tailoring.

33. An electronic device, comprising:

a memory, and

a processor,

wherein the processor is configured to execute program instructions stored in the memory to implement the method according to any one of claims 1-16.

34. A computer readable storage medium having stored thereon program instructions that, when executed by a processor, cause the processor to implement the method of any one of claims 1-16.

35. A computer program product comprising a non-transitory computer readable storage medium having stored thereon computer program that, when read and executed by a computer, implements the method of any one of claims 1-16.