

METHOD, APPARATUS, DEVICE, AND STORAGE MEDIUM FOR DETECTING ROAD DISTRESS

FIELD OF THE DISCLOSURE

[0001] The invention relates to the fields of road engineering technology, particularly to a method, apparatus, device, and storage medium for detecting road distress. The determined distress detection result can evaluate the road safety and provide a reference for the subsequent maintenance of the road.

BACKGROUND

[0002] In the field of computer vision, initial research on deep learning primarily focused on Convolutional Neural Networks(CNNs). With technological advancements and increasing task demands, researchers have proposed object detection algorithms based on CNNs to achieve recognition and localization of specific targets in images. Each image may contain different types and sizes of targets to be detected. The goal of object detection algorithms is to output the types and precise location information of all targets to be detected as accurately as possible. In the field of computer vision, research tasks such as image segmentation and pose detection are extensions and expansions based on object detection research.

[0003] In research on intelligent positioning algorithms for hidden distresses in asphalt pavements, object detection models based on Region-based Convolutional Neural Networks can achieve effective localization of hidden distresses in road radar images. However, the localization

accuracy of this model for distresses still has room for improvement. Simultaneously, in experiments optimizing the base object detection model, it was found that the model has difficulty in locating small-target distresses, which severely limits the improvement of the model's localization effectiveness.

[0004] Therefore, in the existing technology, during the process of road distress detection, there exists the problem of low reliability of the distress detection results.

SUMMARY

[0005] In view of this, it is necessary to provide a method, apparatus, device, and storage medium for detecting road distress to solve the technical problem of low reliability of distress detection results existing in the process of road distress detection in the prior art.

[0006] To solve the above problems, the present invention provides a method for detecting road distress, comprises:

performing distress detection on a road radar image based on a fully trained large-target detection model and a fully trained small-target detection model, and obtaining a large-target detection result and a small-target detection result;

fusing the large-target detection result and the small-target detection result to determine a distress detection result of the road radar image;

wherein the large-target detection model includes a deformable convolutional layer, a dual-path detection head, and a smooth non-maximum suppression operation layer; and the small-target

detection model includes a multi-scale convolutional layer, a deformable ROI pooling layer, the dual-path detection head, a balanced line loss function, and the smooth non-maximum suppression operation layer.

[0007] In some possible embodiments, before performing distress detection on a road radar image based on a fully trained large-target detection model, and obtaining a large-target detection result, comprises:

acquiring a large-target distress detection dataset, the large-target distress detection dataset includes large-target road radar image samples and corresponding large-target distress results;

constructing an initial large-target detection model, inputting the large-target road radar image samples into the initial large-target detection model with large-target distress results as output, calculating sample intersection over union values, performing weight redistribution on the large-target distress detection dataset based on the sample intersection over union values, and performing iterative training on weight-adjusted large-target distress detection dataset until the fully trained large-target detection model is obtained.

[0008] In some possible embodiments, performing iterative training on the large-target distress detection dataset until the fully trained large-target detection model is obtained, comprises:

inputting the large-target road radar image samples into the initial large-target detection model, performing target division on the large-target road radar image samples according to the dual-path detection head to obtain large-target region of interest images;

performing irregular convolution processing on the large-target region of interest images according to the deformable convolutional layer to obtain large-target radar feature maps;

performing detection box screening on the large-target radar feature maps according to the smooth non-maximum suppression operation layer to obtain large-target distress detection boxes.

[0009] In some possible embodiments, before performing distress detection on a road radar image based on a fully trained small-target detection model, and obtaining a small-target detection result, comprises:

acquiring a small-target distress detection dataset, the small-target distress detection dataset includes small-target road radar image samples and corresponding small-target distress results;

constructing an initial small-target detection model, inputting the small-target road radar image samples into the initial small-target detection model with small-target distress results as output, calculating sample loss values, and performing iterative training on the small-target distress detection dataset where the sample loss values are less than a preset loss threshold until the fully trained small-target detection model is obtained.

[0010] In some possible embodiments, performing iterative training on the small-target distress detection data sample until the fully trained small-target detection model is obtained, comprises:

inputting the small-target road radar image samples into the initial small-target detection model, performing multi-scale convolution on the small-target road radar image samples according to the multi-scale convolutional layer to obtain multi-scale feature images;

performing irregular region division on the multi-scale feature images according to the deformable ROI pooling layer to obtain irregularly deformed region images;

performing target division on the irregularly deformed region images according to the dual-path detection head to obtain small-target region of interest images;

performing detection box screening on the small-target region of interest images according to the smooth non-maximum suppression operation layer to obtain small-target distress detection boxes;

performing loss value evaluation on the small-target distress detection boxes and the small-target distress results according to the balanced line loss function, and determining the initial small-target detection model when the loss value is minimized as the fully trained small-target detection model.

[0011] In some possible embodiments, acquiring a large-target distress detection dataset, comprises:

acquiring an initial distress detection dataset, the initial distress detection dataset includes road radar image data and corresponding distress annotation information;

filtering out road radar image data containing subsidence-type distresses and hidden-crack-type distresses as first large-target road radar image samples, and using the distress annotation information corresponding to the first large-target road radar image samples as first large-target distress results;

designating road radar image data containing loose-type distresses as second large-target road radar image samples, and determining that the distress annotation information corresponding to

the second large-target road radar image samples is empty.

[0012] In some possible embodiments, acquiring a small-target distress detection dataset, comprises:

acquiring an initial distress detection dataset, the initial distress detection dataset includes road radar image data and corresponding distress annotation information;

filtering out road radar image data containing loose-type distresses as first small-target road radar image samples, and using the distress annotation information corresponding to the first small-target road radar image samples as first small-target distress results;

designating road radar image data containing subsidence-type distresses and hidden-crack-type distresses as second small-target road radar image samples, and determining that the distress annotation information corresponding to the second small-target road radar image samples is empty.

[0013] In another aspect, the present invention also provides an apparatus for detecting road distress, comprises:

distress detection module, configured to perform distress detection on a road radar image based on a fully trained large-target detection model and a fully trained small-target detection model, and obtaining a large-target detection result and a small-target detection result;

distress detection result determination module, configured to fuse the large-target detection result and the small-target detection result to determine a distress detection result of the road radar image;

wherein the large-target detection model includes a deformable convolutional layer, a dual-path detection head, and a smooth non-maximum suppression operation layer; and the small-target detection model includes a multi-scale convolutional layer, a deformable ROI pooling layer, the dual-path detection head, a balanced line loss function, and the smooth non-maximum suppression operation layer.

[0014] In another aspect, the present invention also provides an electronic device, comprising memory and processor, among which,

the memory is used to store a program;

the processor coupled with the memory, is used to execute the program stored in the memory to implement the steps in the method for detecting road distress according to any of the aforementioned possible embodiment methods.

[0015] In another aspect, the present invention also provides a computer-readable storage medium, the computer-readable storage medium stores a program or instructions that, when executed by a processor, implement the steps of the method for detecting road distress according to the any of the aforementioned possible embodiment methods.

[0016] The beneficial effects of adopting the aforementioned embodiments are as follows: the method for detecting road distress provided by the present invention, by setting two models, a large-target detection model and a small-target detection model, to perform distress detection on road radar images, dual-path object detection is realized. By specifically controlling the structure and

corresponding functions of the models and fusing their detection results, the final detection accuracy is improved. By setting structures such as a multi-scale convolutional layer in the small-target detection model, the detection capability for small-target distresses is greatly enhanced, thereby ensuring the reliability of the final distress detection result.

BRIEF DESCRIPTION OF THE FIGURES

[0017] To describe the technical solutions in embodiments of the present invention more clearly, the following briefly introduces the accompanying drawings for describing the embodiments. Apparently, the accompanying drawings in the following description show merely some embodiments of the present invention, and those skilled in the art may still derive other accompanying drawings from these accompanying drawings without making creative efforts.

[0018] FIG. 1 is a flowchart of the method for detecting road distress provided by the present invention.

[0019] FIG. 2 is a flowchart of obtaining the large-target detection model provided by the present invention.

[0020] FIG. 3 is a flowchart of acquiring a large-target distress detection dataset provided by the present invention.

[0021] FIG. 4 is a flowchart of training large-target detection model provided by the present

invention.

[0022] FIG. 5 is a result diagram of generating a deformable convolution offset provided by the present invention.

[0023] FIG. 6 is a result diagram of sampling changes in deformable convolution provided by the present invention.

[0024] FIG.7 is a flowchart of obtaining the small-target detection model provided by the present invention.

[0025] FIG. 8 is a flowchart of training large-target detection model provided by the present invention.

[0026] FIG. 9 is a schematic structural diagram of multi-scale convolution provided by the present invention.

[0027] FIG. 10 is a result diagram of the main operation process of deformable ROI pooling provided by the present invention.

[0028] FIG. 11 is a schematic structural diagram of an apparatus for detecting road distress provided by the present invention.

[0029] FIG. 12 is a schematic structural diagram of the electronic device provided by the present invention.

DETAILED DESCRIPTION

[0030] The following describes preferred embodiments of the present invention in detail with reference to the accompanying drawings. The drawings form part of the present invention and, together with the embodiments of the present invention, are used to explain the principles of the present invention and are not intended to limit the scope of the present invention.

[0031] ROI (Region of Interest) is a common concept in image processing and object detection. It refers to a specific area of interest to the user in an image, usually containing the target or object we are interested in.

[0032] IOU (Intersection over Union) is a metric used to evaluate the performance of object detection algorithms. It measures the overlap between a predicted bounding box (a candidate box generated by the algorithm) and a ground truth bounding box (a real annotated box).

[0033] R-CNN (Region-based Convolutional Neural Network) is a deep learning model for object detection that achieves target localization and classification by combining region proposals and convolutional neural networks.

[0034] To solve the above problems, the present invention provides a method, apparatus, device, and storage medium for detecting road distress, which are described in detail below. **[0035]** FIG. 1 is a flowchart of the method for detecting road distress provided by the present invention. As shown in FIG. 1, the method for detecting road distress, comprises:

S101, performing distress detection on a road radar image based on a fully trained large-target

detection model and a fully trained small-target detection model, and obtaining a large-target detection result and a small-target detection result;

S102, fusing the large-target detection result and the small-target detection result to determine a distress detection result of the road radar image;

wherein, the large-target detection model includes a deformable convolutional layer, a dual-path detection head, and a smooth non-maximum suppression operation layer; and the small-target detection model includes a multi-scale convolutional layer, a deformable ROI pooling layer, the dual-path detection head, a balanced line loss function, and the smooth non-maximum suppression operation layer.

[0036] It should be noted that the large-target model refers to a model used to detect relatively large-sized distresses, and the small-target model refers to a model used to detect relatively small-sized distresses.

[0037] Among them, road distress mainly includes subsidence-type distresses, hidden-crack-type distresses, and loose-type distresses. Subsidence-type distresses and hidden-crack-type distresses are generally large in size, while loose-type distresses are generally small in size. Therefore, subsidence-type distresses and hidden-crack-type distresses are defined as large-target distresses, and loose-type distresses are defined as small-target distresses.

[0038] It should also be noted that the large-target detection model and the small-target detection model are parallel in relationship. The order of detection by the two can be adjusted as needed,

which is not limited here.

[0039] In this embodiment, first, by setting two models, a large-target detection model and a small-target detection model, to perform distress detection on road radar images, dual-path object detection is realized. By specifically controlling the structure and corresponding functions of the models and fusing their detection results, the final detection accuracy is improved. By setting structures such as a multi-scale convolutional layer in the small-target detection model, the detection capability for small-target distresses is greatly enhanced, thereby ensuring the reliability of the final distress detection result.

[0040] As a preferred embodiment, in S101, before performing distress detection on the road radar image based on the fully trained large-target detection model and the fully trained small-target detection model, it is necessary to obtain the large-target detection model and the small-target detection model respectively.

[0041] As shown in FIG. 2, FIG. 2 is a flowchart of obtaining the large-target detection model provided by the present invention, comprising:

S201, acquiring a large-target distress detection dataset, the large-target distress detection dataset includes large-target road radar image samples and corresponding large-target distress results;

S202, constructing an initial large-target detection model, inputting the large-target road radar image samples into the initial large-target detection model with large-target distress results as output, calculating sample intersection over union values, performing weight redistribution on the

large-target distress detection dataset based on the sample intersection over union values, and performing iterative training on weight-adjusted large-target distress detection dataset until the fully trained large-target detection model is obtained.

[0042] It should be noted that an R-CNN module sharing parameters with the initial large-target detection model is set to calculate the sample IOU values. The R-CNN module only needs to perform forward calculation to obtain the sample IOU values.

[0043] In a specific embodiment, uniform sampling of training samples is performed based on the IOU value, that is, the weight corresponding to the sample IOU value is uniformly determined.

[0044] In this embodiment, by grouping training samples specifically based on the sample IOU value, redistribution of training samples is achieved. Since the IOU value can reflect the difficulty level of the training samples to some extent, the present invention can achieve equal treatment of samples with high difficulty and low difficulty, thereby improving the generalization ability of the model after training.

[0045] Specifically, data statistical research has found that in training samples, samples with IOU values between 0 and 0.05 account for a large proportion, while only a small number of samples are distributed in other IOU intervals. This uneven distribution phenomenon will cause the model to miss many difficult negative samples distributed outside the 0-0.05 IOU interval during random sampling. The balanced sampling method based on IOU values can divide sampling intervals according to the IOU value and perform uniform sampling within each IOU interval, thereby

making the sampling distribution closer to the difficult sample distribution. This enables the large-target detection model to collect more difficult samples during training than random sampling, which is beneficial for improving the training effect and generalization performance of the model.

[0046] Specifically, in S201, to acquire a large-target distress detection dataset, as shown in FIG.

3, FIG. 3 is a flowchart of acquiring a large-target distress detection dataset, comprising:

S301, acquiring an initial distress detection dataset, the initial distress detection dataset includes road radar image data and corresponding distress annotation information;

S302, filtering out road radar image data containing subsidence-type distresses and hidden-crack-type distresses as first large-target road radar image samples, and using the distress annotation information corresponding to the first large-target road radar image samples as first large-target distress results;

S303, designating road radar image data containing loose-type distresses as second large-target road radar image samples, and determining that the distress annotation information corresponding to the second large-target road radar image samples is empty.

[0047] In this embodiment, by adaptively deleting the loose-type distress information in the initial distress detection data set, information about small-target distresses is eliminated, thereby obtaining pure large-target distress sample information, which facilitates improving the training efficiency and pertinence of the large-target detection model in subsequent training.

[0048] In S202, to perform iterative training on the large-target distress detection dataset until the

fully trained large-target detection model is obtained, as shown in FIG. 4, FIG. 4 is a flowchart of training a large-target detection model provided by the present invention, comprising:

S401, inputting the large-target road radar image samples into the initial large-target detection model, performing target division on the large-target road radar image samples according to the dual-path detection head to obtain large-target region of interest images;

S402, performing irregular convolution processing on the large-target region of interest images according to the deformable convolutional layer to obtain large-target radar feature maps;

S403, performing detection box screening on the large-target radar feature maps according to the smooth non-maximum suppression operation layer to obtain large-target distress detection boxes.

[0049] It should be noted that the dual-path detection head is a dual-path detection head including parallel convolutional layers and fully connected (FC) layers. Generally, the R-CNN structure uses dual FC layers for classification and regression tasks. However, by comparing two-stage object detection algorithms, it is found that convolutional layers are more suitable for regression tasks than FC layers, while FC layers are more suitable for classification tasks. Because FC layers have higher spatial sensitivity, they can better distinguish between global targets and local targets in classification tasks. However, FC layers have poor robustness in regression tasks, while convolutional layers can achieve better results than FC layers in regression tasks. Therefore, the present invention effectively improves the detection accuracy of the model by improving the dual-path detection head.

[0050] The deformable convolutional layer refers to replacing the traditional convolution operations in the original Backbone of the base model with deformable convolution. To illustrate the difference between the deformable convolutional layer and the original traditional convolution, first, the traditional convolution operation is defined as:

$$M_{out}[l] = \sum_{c \in \mathbf{C}} w[c] \cdot M_{in}[l + c]$$

[0051] Wherein, M_{in} is the original image input to the traditional convolutional layer, M_{out} is the feature map output after the traditional convolution operation, l is the position parameter in the feature map, $\mathbf{C}$ is the convolution matrix, c is the position parameter in the convolution matrix, and w is the sampling weight.

[0052] By adding an offset to the traditional convolution kernel, uniform convolution becomes irregular convolution. The deformable convolution is obtained as:

$$Q_{out}[l] = \sum_{c \in \mathbf{C}} w[c] \cdot Q_{in}[l + c + \Delta c]$$

[0053] Wherein, Q_{in} is the original image input to the deformable convolution, is the feature map output after the deformable convolution operation, l is the position parameter in the feature map, $\mathbf{C}$ is the convolution matrix, c is the position parameter in the convolution matrix, w is the sampling weight, and Δc is the offset added to the traditional convolution kernel for deformable convolution.

[0054] After applying the offset to the traditional convolution kernel, the original uniform convolution becomes irregular convolution with a variable shape. The convolutional network that generates the offset also participates in the backpropagation optimization during model training. The offset generated after training will better match the characteristics of the distress targets in the image. Among them, as shown in FIG. 5, FIG. 5 is a result diagram of generating a deformable convolution offset provided by the present invention. As shown in FIG. 6, FIG. 6 is a result diagram of sampling changes in deformable convolution provided by the present invention. In FIG. 6, the light gray dots represent the positions before offset, the dark black dots represent the positions after offset, and the arrows represent the offset vectors.

[0055] The smooth non-maximum suppression operation layer refers to replacing the traditional Non-Maximum Suppression (NMS) operation with smooth non-maximum suppression (Soft-NMS). First, when predicting target detection boxes, far more prediction boxes than the number of targets are output. However, each target should have only one real detection box. Therefore, it is necessary to remove redundant prediction boxes and retain the prediction box that best fits the target as the final target detection box. The effect of the traditional NMS operation is:

$$s_i = \begin{cases} s_i, & IOU(M, b_i) < N_i \\ 0, & IOU(M, b_i) \geq N_i \end{cases}$$

[0056] NMS retains the detection box with the highest prediction score and filters out detection boxes whose overlap degree with it exceeds the IOU threshold, thereby achieving the removal of

redundant detection boxes.

[0057] However, although NMS is simple and effective, it has the following problems in complex object detection scenarios: when targets are densely distributed in the detection scene or two targets are close, NMS may cause the detection box of one target to suppress the only detection box of its adjacent target, resulting in missed detection of targets and reducing the final detection accuracy of the model. Therefore, Soft-NMS operation is used to avoid this situation.

[0058] Specifically, when Soft-NMS filters redundant detection boxes, it no longer directly removes detection boxes exceeding the IOU threshold. Instead, it uses a Gaussian function to reduce their prediction scores, specifically:

$$\begin{cases} s_i = s_i G[IOU(M, b_i)], \forall b_i \notin D \\ G(x) = e^{-\frac{x^2}{\sigma}} \end{cases}$$

[0059] Wherein, s_i is the prediction score of the detection box, M is the detection box with the highest current prediction score, b_i is the detection box to be filtered, D is the final result after Soft-NMS, and $G(x)$ is the Gaussian function.

[0060] Obviously, detection boxes closer to M will have their scores reduced more, while detection boxes farther away from M will have their scores reduced less. Therefore, after using Soft-NMS, many detection boxes that would be mistakenly deleted by NMS can be retained, thereby reducing target missed detection and improving the detection effectiveness of the model.

[0061] In this embodiment, by using the dual-path detection head to perform target division on the

large-target road radar image samples, the accuracy of classification results is ensured while regression stability is improved; by using the deformable convolutional layer to perform irregular convolution processing on the large-target ROI images, it can better adapt to the needs of distress detection on images; by using the smooth non-maximum suppression operation layer to screen detection boxes on the large-target radar feature maps, the reliability of the final target detection boxes can be better ensured; in summary, by adaptively adjusting some structures in the initial large-target detection model, it can better adapt to the needs of large-target distress detection and improve detection accuracy.

[0062] Further, as shown in FIG. 7, FIG. 7 is a flowchart of obtaining the small-target detection model provided by the present invention, comprising:

S701, acquiring a small-target distress detection dataset, the small-target distress detection dataset includes small-target road radar image samples and corresponding small-target distress results;

S702, constructing an initial small-target detection model, inputting the small-target road radar image samples into the initial small-target detection model with the small-target distress results as output, calculating sample loss values, and performing iterative training on the small-target distress detection dataset where the sample loss values are less than a preset loss threshold until the fully trained small-target detection model is obtained.

[0063] In a specific embodiment, training samples can be divided into difficult samples and easy samples based on detection difficulty. Making the model pay more attention to difficult samples

during training helps improve the model's generalization performance. The original R-CNN structure of the base model uses a random sampling method, which cannot guarantee the number of difficult samples after sampling.

[0064] Therefore, to increase the number of difficult samples collected by R-CNN, the present invention replaces random sampling with the Online Hard Example Mining (OHEM) method. OHEM regards samples that produce larger loss values as difficult samples, sorts the samples according to the loss value, and selects the top difficult samples by loss value, thereby ensuring the number of difficult samples after sampling.

[0065] In a specific embodiment, the OHEM sampling method uses an R-CNN module sharing parameters with the original R-CNN module to calculate the sample loss. This R-CNN module only needs to perform forward calculation to obtain the loss value and does not need backpropagation. After calculating the loss values, the training samples are sorted in ascending order of loss value. After NMS, the 256 samples with the largest loss values are selected for training the original R-CNN module. After OHEM sampling, the small-target detection model allows difficult samples to receive more training iterations, which helps improve the generalization performance of the small-target detection model.

[0066] In this embodiment, by using the sample loss value as a grouping criterion to redistribute training samples, it can effectively combine the characteristics of small targets to select more comprehensive samples specifically, thereby improving the training effect of the model and

ensuring the reliability of the trained small-target detection model.

[0067] As a preferred embodiment, in S701, similar to the process of obtaining the large-target distress detection dataset, to acquire the small-target distress detection dataset, first, acquiring an initial distress detection dataset, the initial distress detection dataset includes road radar image data and corresponding distress annotation information; then, filtering out road radar image data containing loose-type distresses as first small-target road radar image samples, and using the distress annotation information corresponding to the first small-target road radar image samples as first small-target distress results; finally, designating road radar image data containing subsidence-type distresses and hidden-crack-type distresses as second small-target road radar image samples, and determining that the distress annotation information corresponding to the second small-target road radar image samples is empty.

[0068] In this embodiment, by adaptively deleting the annotation information of subsidence-type distresses and hidden-crack-type distresses in the initial distress detection data set, information about large-target distresses is cleared, thereby obtaining pure small-target distress sample information, which facilitates improving the training efficiency and pertinence of the small-target detection model in subsequent training.

[0069] As a preferred embodiment, in S702, to perform iterative training on the small-target distress detection dataset until the trained small-target detection model is obtained, as shown in FIG. 8, FIG. 8 is a flowchart of training large-target detection model provided by the present

invention, comprising:

S801, inputting the small-target road radar image samples into the initial small-target detection model, performing multi-scale convolution on the small-target road radar image samples according to the multi-scale convolutional layer to obtain multi-scale feature images;

S802, performing irregular region division on the multi-scale feature images according to the deformable ROI pooling layer to obtain irregularly deformed region images;

S803, performing target division on the irregularly deformed region images according to the dual-path detection head to obtain small-target region of interest images;

S804, performing detection box screening on the small-target region of interest images according to the smooth non-maximum suppression operation layer to obtain small-target distress detection boxes;

S805, performing loss value evaluation on the small-target distress detection boxes and the small-target distress results according to the balanced line loss function, and determining the initial small-target detection model when the loss value is minimized as the fully trained small-target detection model.

[0070] The multi-scale convolutional layer refers to replacing the intermediate convolutional layers of the residual structure with multi-scale convolution to enhance the multi-scale representation capability of the network. Specifically, by replacing the ResNet-50 (intermediate convolutional layers) originally used in the Backbone of the base model with Res2Net (multi-scale

convolution), it is beneficial to improve the feature extraction capability of the small-target detection model for small targets. The basic structure of Res2Net is shown in FIG. 9, FIG. 9 is a schematic structural diagram of multi-scale convolution provided by the present invention. Res2Net's fusion of multi-scale convolution methods can make the output features have different receptive fields and contain feature information of the target at different scales.

[0071] The deformable ROI pooling layer refers to replacing the original ROI Align operation in the ROI alignment module of the base model with deformable ROI pooling. Deformable ROI pooling enables the aligned ROI features to incorporate correlation information from other regions, which is beneficial for the small-target detection model to detect small-size distress targets.

[0072] The ordinary ROI pooling operation divides the ROI region into 7×7 grids and outputs the aligned ROI features through max pooling on the divided sub-region blocks. Deformable ROI pooling adds an offset to each sub-region block of the 7×7 division, turning the uniformly divided region into an irregular deformable region. The deformable ROI pooling operation is defined as:

$$R_{out}(i,j) = \sum_{b \in BIN(i,j)} MAX[f_R(b + \Delta b_{i,j})]$$

[0073] Wherein, f_R is the feature image matrix input to the deformable ROI pooling layer, R_{out} is the ROI feature matrix output by the deformable ROI pooling layer, i and j represent the horizontal and vertical coordinates in the matrix, $BIN(i,j)$ is the uniformly divided sub-region block matrix, b is the element value in the sub-region block matrix, and $\Delta b_{i,j}$ is the offset added to the

sub-region block by the deformable ROI pooling.

[0074] As shown in FIG. 10, FIG. 10 is a result diagram of the main operation process of deformable ROI pooling provided by the present invention. The offset of the sub-region block is generated by adding an FC layer after the ROI features output by the ordinary ROI pooling layer. During model training, the weight parameters of the FC layer also participate in the backpropagation optimization process.

[0075] The small-target detection model is mainly used to detect small-size distress targets, and its detection accuracy is greatly affected by the target position. Therefore, the small-target detection model has higher requirements for the accuracy of target position regression. On the other hand, since the detection object of the small-target detection model is only the loose-type distress category, its requirements for classification tasks are relatively low. Based on the above analysis, it is necessary to increase the weight of the regression loss in the comprehensive loss calculation of the small-target detection model, i.e., increase the weighting coefficient before the regression loss. However, if the weight of the regression loss is directly increased, gradient imbalance between difficult and easy samples will occur.

[0076] Therefore, in the original R-CNN structure of the base model, the present application replaces the regression loss calculation using the SmoothL1Loss function with the BalancedL1Loss function (balanced line loss function).

[0077] Among them, after differentiating the SmoothL1Loss function, the gradient change during

backpropagation can be obtained as:

$$\frac{\partial SmoothL1}{\partial x} = \begin{cases} x, & \text{if } |x| < 1 \\ 1, & \text{otherwise} \end{cases}$$

[0078] Wherein, *SmoothL1* is the SmoothL1 loss function, and x represents the regression error value.

[0079] Training samples with regression error values greater than 1 can be regarded as difficult samples, and the propagation gradient value they produce is constant at 1. Training samples with regression error values less than 1 can be regarded as easy samples, and the propagation gradient they produce is related to x . After increasing the weighting coefficient of the regression loss, the increased gradient due to the larger coefficient for difficult samples will be much larger than that for easy samples, leading to gradient imbalance between difficult and easy samples, which is detrimental to the improvement of model training effectiveness. Moreover, the detection objects of the small-target detection model are mostly small-target difficult samples, which will exacerbate the gradient imbalance between difficult and easy samples.

[0080] The BalancedL1Loss function is:

$$L_{BalancedL1}(x) = \begin{cases} \frac{\alpha}{b}(b|x| + 1) \ln(b|x| + 1) - \alpha|x|, & \text{if } |x| < 1 \\ \gamma|x| + C, & \text{otherwise} \end{cases}$$

[0081] Wherein, α is 0.5, γ is 1.5, b is used to ensure function continuity, and the value of b needs to satisfy the equation $\alpha \ln(b + 1) = \gamma$.

[0082] After differentiating the BalancedL1Loss function, the gradient change during backpropagation can be obtained as:

$$\frac{\partial L_{\text{BalancedL1}}}{\partial x} = \begin{cases} \alpha \ln(b|x| + 1), & \text{if } |x| < 1 \\ \gamma, & \text{otherwise} \end{cases}$$

[0083] By replacing the original SmoothL1 loss function with the BalancedL1 loss function, the gradient for easy samples will increase faster, making the gradient contribution of difficult and easy samples balanced after the increase in the regression loss weighting coefficient, thereby improving the aforementioned gradient imbalance problem and enhancing the training effectiveness of the small-target detection model.

[0084] In this embodiment, by using the multi-scale convolutional layer to perform multi-scale convolution on the small-target road radar image samples, the dimensionality of feature information can be increased; by using the deformable ROI pooling layer to perform irregular region division on the multi-scale feature images, adaptive adjustment of region division is achieved, which better meets actual needs; by using the dual-path detection head to perform target division on the irregularly deformed region images, the accuracy of classification results is ensured while regression stability is improved; by using the smooth non-maximum suppression operation layer to screen detection boxes on the small-target ROI images, the reliability of the final target detection boxes can be better ensured; by using the balanced line loss function to evaluate the loss value between the small-target distress detection boxes and the small-target distress results, the

aforementioned gradient imbalance problem can be improved, and the training effectiveness of the small-target detection model can be enhanced; in summary, by adaptively adjusting some structures in the initial small-target detection model, it can better adapt to the needs of small-target distress detection and improve detection accuracy.

[0085] Through the above methods, by setting two models, a large-target detection model and a small-target detection model, to perform distress detection on road radar images, dual-path object detection is realized; by specifically controlling the structure and corresponding functions of the models and fusing their detection results, the final detection accuracy is improved; by setting structures such as a multi-scale convolutional layer in the small-target detection model, the detection capability for small-target distresses is greatly enhanced, thereby ensuring the reliability of the final distress detection result.

[0086] To solve the above problems, the present invention also provides an apparatus for detecting road distress. As shown in FIG. 11, FIG. 11 is a schematic structural diagram of an apparatus for detecting road distress provided by the present invention. The apparatus for detecting road distress 1100 comprises:

distress detection module 1101, configured to perform distress detection on a road radar image based on a fully trained large-target detection model and a fully trained small-target detection model, and obtaining a large-target detection result and a small-target detection result;

distress detection result determination module 1102, configured to fuse the large-target detection

result and the small-target detection result to determine a distress detection result of the road radar image;

wherein the large-target detection model includes a deformable convolutional layer, a dual-path detection head, and a smooth non-maximum suppression operation layer; and the small-target detection model includes a multi-scale convolutional layer, a deformable ROI pooling layer, the dual-path detection head, a balanced line loss function, and the smooth non-maximum suppression operation layer.

[0087] The apparatus for detecting road distress 1100 provided by the above mentioned embodiments can implement the technical solutions described in the embodiments of the method for detecting road distress. The specific embodiment principles of the above mentioned modules or units can be referred to the corresponding content in the embodiments of the method for detecting road distress, and will not be repeated here.

[0088] As shown in FIG.12, the present invention also provides an electronic device 1200. The electronic device 1200 includes a processor 1201, a memory 1202, and a display 1203. FIG. 12 only shows some components of the electronic device 1200. It should be understood that it is not required to implement all the shown components, and more or fewer components can be implemented alternatively.

[0089] The processor 1201 can be a Central Processing Unit (CPU), a micro processor, or other data processing chips in some embodiments. It is used to run the program code stored in the

memory 1202 or process data, such as the foot reflex zones identification method of the present invention.

[0090] In some embodiments, the processor 1201 may be a single server or a server group. The server group may be centralized or distributed. In some embodiments, the processor 1201 may be local or remote. In some embodiments, the processor 1201 may be implemented on a cloud platform. In one embodiment, the cloud platform may include a private cloud, a public cloud, a hybrid cloud, a community cloud, a distributed cloud, an internal cloud, a multi-cloud, or any combination thereof.

[0091] The memory 1202 can be an internal storage unit of the electronic device 1200 in some embodiments, such as a hard disk or memory of the electronic device 1200. In other embodiments, the memory 1202 can also be an external storage device of the electronic device 1200, such as a plug - in hard disk, a Smart Media Card (SMC), a Secure Digital (SD) card, a Flash Card, etc. equipped on the electronic device 1200.

[0092] Further, the memory 1202 can include both the internal storage unit and the external storage device of the electronic device 1200. The memory 1202 is used to store the application software installed on the electronic device 1200 and various types of data.

[0093] In some embodiments, the display 1203 can be an LED display, a liquid crystal display, a touch type liquid crystal display, an OLED (Organic Light-Emitting Diode) touch device, etc. The display 1203 is used to display the information of the electronic device 1200 and a visual user

interface. The components 1201-1203 of the electronic device 1200 communicate with each other through the system bus.

[0094] In some embodiments of the present invention, when the processor 1201 executes program for detecting road distress in the memory 1202, the following steps can be implemented:

performing distress detection on a road radar image based on a fully trained large-target detection model and a fully trained small-target detection model, and obtaining a large-target detection result and a small-target detection result;

fusing the large-target detection result and the small-target detection result to determine a distress detection result of the road radar image;

wherein, the large-target detection model includes a deformable convolutional layer, a dual-path detection head, and a smooth non-maximum suppression operation layer; and the small-target detection model includes a multi-scale convolutional layer, a deformable ROI pooling layer, the dual-path detection head, a balanced line loss function, and the smooth non-maximum suppression operation layer.

[0095] It should be understood that when the processor 1201 executes the program for detecting road distress in the memory 1202, in addition to the above mentioned functions, it can also implement other functions. For specific details, please refer to the descriptions of the relevant method embodiments above.

[0096] Furthermore, the embodiments of the present invention do not specifically limit the type of

the electronic device 1200 mentioned. The electronic device 1200 can be portable electronic devices such as mobile phones, tablet computers, personal digital assistants (PDAs), wearable devices, laptop computers, etc. Exemplary embodiments of portable electronic devices include, but are not limited to, portable electronic devices equipped with operating systems such as iOS, Android, Microsoft, or others. The above - mentioned portable electronic devices can also be other portable electronic devices. It should also be understood that in some other embodiments of the present invention, the electronic device 1200 may not be a portable electronic device, but a desktop computer with a sensitive touch surface (such as a touch panel).

[0097] Correspondingly, the embodiments of the present invention also provide a computer-readable storage medium. The computer-readable storage medium is used to store programs or instructions that can be read by a computer. When the programs or instructions are executed by a processor, they can implement the steps or functions of the method for detecting road distress provided by the above mentioned method embodiments.

[0098] It is understood by those skilled in the art that all or part of the process to implement the above embodiments may be accomplished by instructs the relevant hardware (e.g. processor, controller, etc.) through a computer program that may be stored in a computer readable storage medium. Among them, the computer readable storage medium is disk, optical disc, read-only storage memory or random storage memory.

[0099] The above method and apparatus for determining deflection basin parameters, electronic

device, medium and product provided by the invention are introduced in detail. The principle and embodiment mode of the invention are described in this paper with specific examples. The above embodiment is only used to help understand the method and its core idea of the invention. At the same time, for the technical personnel in the field, according to the idea of the invention, there will be changes in the specific mode of embodiment and scope of application. In summary, the content of this specification should not be understood as a limitation of the invention.

CLAIMS

What is claimed is:

1. A method for detecting road distress, comprises:

performing distress detection on a road radar image based on a fully trained large-target detection model and a fully trained small-target detection model, and obtaining a large-target detection result and a small-target detection result;

fusing the large-target detection result and the small-target detection result to determine a distress detection result of the road radar image;

wherein, the large-target detection model includes a deformable convolutional layer, a dual-path detection head, and a smooth non-maximum suppression operation layer; and the small-target detection model includes a multi-scale convolutional layer, a deformable ROI pooling layer, the dual-path detection head, a balanced line loss function, and the smooth non-maximum suppression operation layer.

2. The method for detecting road distress of claim 1, before performing distress detection on a road radar image based on a fully trained large-target detection model, and obtaining a large-target detection result, comprises:

acquiring a large-target distress detection dataset, the large-target distress detection dataset includes large-target road radar image samples and corresponding large-target distress results;

constructing an initial large-target detection model, inputting the large-target road radar image

samples into the initial large-target detection model with large-target distress results as output, calculating sample intersection over union values, performing weight redistribution on the large-target distress detection dataset based on the sample intersection over union values, and performing iterative training on weight-adjusted large-target distress detection dataset until the fully trained large-target detection model is obtained.

3. The method for detecting road distress of claim 2, performing iterative training on the large-target distress detection dataset until the fully trained large-target detection model is obtained, comprises:

inputting the large-target road radar image samples into the initial large-target detection model, performing target division on the large-target road radar image samples according to the dual-path detection head to obtain large-target region of interest images;

performing irregular convolution processing on the large-target region of interest images according to the deformable convolutional layer to obtain large-target radar feature maps;

performing detection box screening on the large-target radar feature maps according to the smooth non-maximum suppression operation layer to obtain large-target distress detection boxes.

4. The method for detecting road distress of claim 1, before performing distress detection on a road radar image based on a fully trained small-target detection model, and obtaining a small-target detection result, comprises:

acquiring a small-target distress detection dataset, the small-target distress detection dataset

includes small-target road radar image samples and corresponding small-target distress results;
constructing an initial small-target detection model, inputting the small-target road radar image samples into the initial small-target detection model with the small-target distress results as output, calculating sample loss values, and performing iterative training on the small-target distress detection dataset where the sample loss values are less than a preset loss threshold until the fully trained small-target detection model is obtained.

5. The method for detecting road distress of claim 4, performing iterative training on the small-target distress detection data sample until the fully trained small-target detection model is obtained, comprises:

inputting the small-target road radar image samples into the initial small-target detection model, performing multi-scale convolution on the small-target road radar image samples according to the multi-scale convolutional layer to obtain multi-scale feature images;

performing irregular region division on the multi-scale feature images according to the deformable ROI pooling layer to obtain irregularly deformed region images;

performing target division on the irregularly deformed region images according to the dual-path detection head to obtain small-target region of interest images;

performing detection box screening on the small-target region of interest images according to the smooth non-maximum suppression operation layer to obtain small-target distress detection boxes;

performing loss value evaluation on the small-target distress detection boxes and the small-target

distress results according to the balanced line loss function, and determining the initial small-target detection model when the loss value is minimized as the fully trained small-target detection model.

6. The method for detecting road distress of claim 2, acquiring a large-target distress detection dataset, comprises:

acquiring an initial distress detection dataset, the initial distress detection dataset includes road radar image data and corresponding distress annotation information;

filtering out road radar image data containing subsidence-type distresses and hidden-crack-type distresses as first large-target road radar image samples, and using the distress annotation information corresponding to the first large-target road radar image samples as first large-target distress results;

designating road radar image data containing loose-type distresses as second large-target road radar image samples, and determining that the distress annotation information corresponding to the second large-target road radar image samples is empty.

7. The method for detecting road distress of claim 4, acquiring a small-target distress detection dataset, comprises:

acquiring an initial distress detection dataset, the initial distress detection dataset includes road radar image data and corresponding distress annotation information;

filtering out road radar image data containing loose-type distresses as first small-target road radar image samples, and using the distress annotation information corresponding to the first small-

target road radar image samples as first small-target distress results;

designating road radar image data containing subsidence-type distresses and hidden-crack-type distresses as second small-target road radar image samples, and determining that the distress annotation information corresponding to the second small-target road radar image samples is empty.

8. An apparatus for detecting road distress , comprises:

distress detection module, configured to perform distress detection on a road radar image based on a fully trained large-target detection model and a fully trained small-target detection model, and obtaining a large-target detection result and a small-target detection result;

distress detection result determination module, configured to fuse the large-target detection result and the small-target detection result to determine a distress detection result of the road radar image;

wherein the large-target detection model includes a deformable convolutional layer, a dual-path detection head, and a smooth non-maximum suppression operation layer; and the small-target detection model includes a multi-scale convolutional layer, a deformable ROI pooling layer, the dual-path detection head, a balanced line loss function, and the smooth non-maximum suppression operation layer.

9. An electronic device, comprising memory and processor, among which,

the memory is used to store a program;

the processor coupled with the memory, is used to execute the program stored in the memory to

implement the steps in the method for detecting road distress according to any one of claims 1 to 7.

10. A computer-readable storage medium, the computer-readable storage medium stores a program or instructions that, when executed by a processor, implement the steps of the method for detecting road distress according to any one of claims 1 to 7.