METHOD AND APPARATUS FOR VEHICLE DAMAGE IDENTIFICATION

Cross-Reference to Related Applications

[0001] This international application is based upon and claims priority to Chinese Patent Application No. 201811089438.6, filed on September 18, 2018, the entire content of which is incorporated herein by reference.

Technical Field

[0002] Embodiments of the present specification relate to the field of machine learning, and more particularly to a method and an apparatus for vehicle damage identification using machine learning.

Technical Background

[0003] In a traditional vehicle insurance claim settlement scenario, an insurance company needs to send professional survey and damage assessment personnel to a scene of an accident to conduct an on-site survey and damage assessment, propose a vehicle's repair plan and compensation amount, take photos of the scene, and record the damage assessment photos for damage and compensation check by off-site inspectors. Due to the need for manual survey and damage assessment, the insurance company needs to invest a lot of labor costs and professional training costs. In terms of the experience of ordinary users, they have to wait for a manual surveyor to take pictures on the spot, a damage assessor to assess damage at a repair location, and a damage inspector to conduct a damage check during the claim settlement process. Therefore, the claim settlement period takes up to 1-3 days, the user waiting time is long, and the user experience is poor.

[0004] In view of the above, the inventors apply artificial intelligence and machine learning to the scene of damage assessment for a vehicle, and use the computer vision image identification technology in the field of artificial intelligence to automatically recognize, according to on-site damage pictures taken by ordinary users, the vehicle damage status reflected in the pictures, and automatically provide a repair plan. Therefore, manual surveys, damage assessments, and checks in the traditional vehicle insurance claim settlement scenario are not required, which can greatly reduce insurance company costs and improve the vehicle insurance claim settlement experience of users.

[0005] Moreover, the accuracy of damage identification in current intelligent damage assessment solutions needs to be further improved. Therefore, an improved solution is also desired for further optimizing vehicle damage detection results and improving identification accuracy.

Summary of the invention

[0006] Embodiments of the present specification provide a method and apparatus for vehicle damage identification, wherein a plurality of target detection models are used to perform detection on a to-be-identified picture and obtain a plurality of candidate bounding boxes, and then detection results are optimized based on association relationships between the candidate bounding boxes to improve the accuracy of damage identification.

[0007] According to a first aspect, a computer-implemented method for vehicle damage identification is provided, including: obtaining a vehicle damage picture to be identified ;inputting the vehicle damage picture into a plurality of pre-trained target detection models respectively, and obtaining corresponding detection results from the plurality of target detection models as a detection result set, wherein the detection result set includes candidate bounding boxes detected by the plurality of target detection models and category prediction results of the candidate bounding

boxes; for any first candidate bounding box of the candidate bounding boxes, determining an integrated feature vector of the first candidate bounding box according to at least an association relationship between the first candidate bounding box and other candidate bounding boxes, wherein the association relationship includes one or more of a bounding box content association relationship, a location association relationship, and a category prediction result association relationship; and separately inputting the integrated feature vectors corresponding to the candidate bounding boxes into a pre-trained classification model, and optimizing the detection result set according to output results of the classification model.

[0008] In an embodiment, the other candidate bounding boxes include a second candidate bounding box, and the association relationship includes a bounding box content association relationship. In this case, the integrated feature vector of the first candidate bounding box is determined by: obtaining a convolution feature vector of each candidate bounding box, wherein the convolution feature vector is used to represent a picture content feature in a corresponding bounding box; determining a content association feature between the first candidate bounding box and the second candidate bounding box based on a first convolution feature vector corresponding to the first candidate bounding box and a second convolution feature vector corresponding to the second candidate bounding box; and determining, according to at least the content association feature, the integrated feature vector of the first candidate bounding box.

[0009] In an embodiment, the content association feature between the first candidate bounding box and the second candidate bounding box is determined by: calculating a dot product of the first convolution feature vector and the second convolution feature vector, and determining the content association feature based on the dot product; and/or calculating a distance between the first convolution feature vector and the second convolution feature vector, and determining the

content association feature based on the distance, wherein the distance includes a cosine distance or a Euclidean distance.

[0010] In an embodiment, the other candidate bounding boxes include a second candidate bounding box, and the association relationship includes a location association relationship. In this case, the integrated feature vector of the first candidate bounding box is determined by: determining a location association feature between the first candidate bounding box and the second candidate bounding box based on a location parameter of the first candidate bounding box and a location parameter of the second candidate bounding box; and determining, according to at least the location association feature, the integrated feature vector of the first candidate bounding box.

[0011] In an embodiment, the location association feature includes one or more of a center-to-center distance, an IoU (Intersection-over-Union) ratio, an area ratio, and an inclusion relationship.

[0012] In an embodiment, the other candidate bounding boxes include a second candidate bounding box, and the association relationship includes a category prediction result association relationship. In this case, the integrated feature vector of the first candidate bounding box is determined by: determining a prediction result association feature between the first candidate bounding box and the second candidate bounding box based on a first prediction result corresponding to the first candidate bounding box and a second prediction result corresponding to the second candidate bounding box; and determining, according to at least the prediction result association feature, the integrated feature vector of the first candidate bounding box.

[0013] In an embodiment, the first prediction result includes a first category and a first confidence, the second prediction result includes a second category and a second confidence, and

the prediction result association feature includes: whether the first category is the same as the second category, and a difference between the first confidence and the second confidence.

[0014] In an embodiment, the other candidate bounding boxes include one or more of the following: candidate bounding boxes having a center-to-center distance from the first candidate bounding box falling within a predetermined threshold, and a predetermined number of candidate bounding boxes ranked based on a center-to-center distance from the first candidate bounding box.

[0015] In an embodiment, the classification model includes one of a Gradient Boosting Decision Tree (GBDT) model, a Support Vector Machine (SVM) model, and a Deep Neural Network (DNN)-based classification model.

[0016] In an embodiment, the detection result set is optimized by: removing the first candidate bounding box from the detection result set when an output result of the classification model for the first candidate bounding box is an abnormal bounding box.

[0017] According to a second aspect, an apparatus for vehicle damage identification is provided, including: an obtaining unit configured to obtain a vehicle damage picture to be identified; a detecting unit configured to input the vehicle damage picture into a plurality of pre-trained target detection models respectively, and obtain corresponding detection results from the plurality of target detection models as a detection result set, wherein the detection result set includes candidate bounding boxes detected by the plurality of target detection models and category prediction results of the candidate bounding boxes; an integrating unit configured to, for any first candidate bounding box of the candidate bounding boxes, determine an integrated feature vector of the first candidate bounding box according to at least an association relationship between the first candidate bounding box and other candidate bounding boxes, wherein the association relationship includes one or more of a bounding box content association relationship, a location association relationship, and a category prediction result association relationship; and an

optimizing unit configured to separately input the integrated feature vectors corresponding to the candidate bounding boxes into a pre-trained classification model, and optimize the detection result set according to output results of the classification model.

[0018] According to a third aspect, a computer-readable storage medium is provided, having stored therein a computer program for causing a computer to perform the method of the first aspect when the computer program is executed in a computer.

[0019] According to a fourth aspect, a computing device is provided, including a memory and a processor, wherein the memory stores an executable code, and when the processor executes the executable code, the method of the first aspect is implemented.

[0020] According to the method and the apparatus provided by the embodiments of the present specification, on one hand, multiple target detection models are used at the same time to perform damage identification and detection on a vehicle damage picture, which makes generated detection results complement each other, and therefore reduces the possibility of missed detection to the greatest extent. Further, for the N candidate bounding boxes obtained by the plurality of target detection models, based on the association relationship between each candidate bounding box and other candidate bounding boxes, a connection is established between the candidate bounding boxes, and based on such a connection result, a trained classification model is further used to determine the accuracy of each candidate bounding box, thereby eliminating bounding boxes of false detection and further optimizing the damage identification result. In this way, the accuracy of damage identification can be comprehensively improved.

Brief Description of the Drawings

[0021] In order to make the technical solutions of embodiments of the present invention more clear, drawings required for the description of embodiments will be explained briefly as

follows. Clearly, the drawings used in the following description are merely example embodiments of the present invention.

[0022] FIG. 1 is a schematic diagram showing an implementation scenario of a method for vehicle damage identification according an embodiment;

[0023] FIG. 2 shows a flow chart of a method for vehicle damage identification according to an embodiment;

[0024] FIG. 3a shows an example of a vehicle damage picture according an embodiment;

[0025] FIG. 3b shows an example of a plurality of candidate bounding boxes detected by a plurality of target detection models according an embodiment;

[0026] FIG. 3c shows optimized detection results according an embodiment;

[0027] FIG. 4 illustrates the steps of determining an integrated feature vector according an embodiment; and

[0028] FIG. 5 shows a schematic block diagram of an identification apparatus according to an embodiment.

Detailed Description

[0029] The solutions provided in the present specification will be described below with reference to the accompanying drawings.

[0030] In order to identify damage in a picture for vehicle damage assessment, a method commonly used in the industry is to determine the damaged components in a picture and their degree of damage by comparing with massive historical databases to obtain a similar picture. However, the accuracy of damage identification in such a way is not ideal.

[0031] In addition, some target detection models are trained by means of sample labeling. These target detection models are trained to detect vehicle damage as a detection object for damage

identification. Reflection and stain interference are two major hindrances in the process of damage identification. In the result of the identification, the true damage is often detected correctly, and at the same time, a few reflections or stains are detected as damage, and thus false detection occurs.

[0032] In order to improve the accuracy of damage identification, in the embodiments disclosed in the present specification, a solution for optimizing damage detection results is creatively provided. FIG. 1 is a schematic diagram showing an implementation scenario of an embodiment disclosed in the present specification. As shown in FIG. 1, after a vehicle damage picture is obtained, the picture is input to a plurality of target detection models. The plurality of target detection models independently perform damage identification and output respective damage detection results. In general, the damage detection result includes a group of identified damage bounding boxes and corresponding category prediction results. Then, multiple groups of damage bounding boxes output by the plurality of target detection models are used as candidate bounding boxes, and an association relationship between the candidate bounding boxes is determined, thereby establishing a connection between the candidate bounding boxes. According to an embodiment, the association relationship may include one or more of the following: an association relationship about bounding box content, an association relationship about location, and an association relationship about category prediction results. Based on such associations, a trained classification model is used to optimize the damage detection results to obtain more accurate damage detection results. The specific implementation process of optimizing the damage detection results is described below.

[0033] FIG. 2 shows a flow chart of a method for vehicle damage identification according to an embodiment. The method may be performed by any apparatus, device, platform, and device cluster having computing and processing capabilities. As shown in FIG. 2, the method includes: step 21: obtaining a vehicle damage picture to be identified; step 22: inputting the vehicle damage

picture into a plurality of pre-trained target detection models respectively, and obtaining groups of corresponding detection results from the plurality of target detection models, wherein the groups of detection results constitute a detection result set, and the detection result set includes N candidate bounding boxes detected by the plurality of target detection models and category prediction results of the candidate bounding boxes; step 23, for any first candidate bounding box of the N candidate bounding boxes, determining an integrated feature vector of the first candidate bounding box according to an association relationship between the first candidate bounding box and other candidate bounding boxes, wherein the association relationship at least includes one or more of the following: a bounding box content association relationship, a location association relationship, and a category prediction result association relationship; and step 24, separately inputting the integrated feature vectors corresponding to the candidate bounding boxes into a pre-trained classification model, and optimizing the detection result set according to output results of the classification model. The manner in which the above various steps are performed will be described below.

[0034] First, in step 21, a vehicle damage picture is obtained. It can be understood that the picture may be a picture of a vehicle damage scene taken by an ordinary user, and is a picture to be subject to damage identification. FIG. 3a shows an example of a vehicle damage picture. The picture is an unprocessed on-site picture taken by an ordinary user.

[0035] Next, in step 22, the vehicle damage picture is input into a plurality of pre-trained target detection models respectively, and groups of corresponding detection results are obtained from the plurality of target detection models.

[0036] It can be understood that the target detection models are used to identify a specific target object from the picture and classify the target object. In order to identify the vehicle damage by using target detection models, a large number of labeled pictures with labeled damage objects

may be used as training samples to train the target detection models, so that the trained target detection models may be used to identify damage objects from pictures.

[0037] Various target detection models have been provided based on various network structures and various detection algorithms. For example, a one-stage detection model may directly determine the category probability and position coordinates of a target object from a picture, that is, it may directly identify the target object. Typical examples of one-stage detection models include Single Shot Detector (SSD) models and You Only Look Once (YOLO) models. Also for example, a two-stage detection model first generates a candidate region, also called a region of interest (ROI), in the picture, and then performs target identification and bounding box regression in the candidate region. Typical examples of two-stage detection models include Region-Convolutional Neural Network (R-CNN) models, Fast R-CNN models, Faster R-CNN models, and so on. Other target detection models have also been provided.

[0038] A plurality of target detection models may be selected from the various target detection models mentioned above, and target detection is performed on the vehicle damage picture in step 22. The selected plurality of target detection models may each conduct detection on the vehicle damage picture in parallel and independently.

[0039] It can be understood that each target detection model outputs a group of detection results. Generally, the detection result output by a target detection model includes a bounding box containing the target object and the category of the target object in the bounding box. When a target detection algorithm is applied to damage identification, the region bounding box of the damage object and the damage category of the damage object will be obtained. Correspondingly, the plurality of target detection models output groups of corresponding detection results, and the groups of detection results constitute a detection result set. Therefore, the detection result set includes a total of N damage bounding boxes detected by the plurality of target detection models,

and corresponding damage category predictions. FIG. 3b schematically shows a plurality of damage bounding boxes included in the detection result set and corresponding damage categories (the number at the upper left corner of each bounding box) output by the plurality of target detection models.

[0040] Next, the damage bounding boxes detected by the target detection models are used as candidate bounding boxes, and the N candidate bounding boxes are further optimized.

[0041] In step 23, for the N candidate bounding boxes, an association relationship between each candidate bounding box and other candidate bounding boxes is determined, thereby establishing connections among the bounding boxes for optimizing the results. An example of a candidate bounding box, referred to herein as a first candidate bounding box, is described below for explanation. It should be understood that the first candidate bounding box here is any one of the N candidate bounding boxes, wherein the wording "first" is only for differentiation, and has no other limiting effect.

[0042] For any first candidate bounding box described above, in step 23, an integrated feature vector of the first candidate bounding box is determined according to at least the association relationship between the first candidate bounding box and other candidate bounding boxes. In one embodiment, the integrated feature vector of the first candidate bounding box includes a first part and a second part, the first part is a feature part associated only with the candidate bounding box itself, and the second part is a feature part obtained on the basis of association relationship with other candidate bounding boxes. In another embodiment, the integrated feature vector may include the second part only.

[0043] According to the embodiments of the present specification, the association relationship may be an association relationship about the content of the bounding box, an

association relationship about the location of the bounding box, and an association relationship about the category prediction result of the bounding box.

[0044] More specifically, in one embodiment, the association relationship includes a bounding box content association relationship, that is, an association relationship between image features within bounding boxes. In such a case, the association relationship in content between the first candidate bounding box and the other candidate bounding boxes may be determined, and the integrated feature vector of the first candidate bounding box is determined accordingly. For clarity and simplicity of the description, it is assumed that the other candidate bounding boxes include a second candidate bounding box, and the first candidate bounding box and the second candidate bounding box are used as an example below to describe the manner in which step 23 is performed.

[0045] FIG. 4 illustrates the steps of determining an integrated feature vector, i.e., the sub-steps of step 23, in one embodiment. As shown in FIG. 4, first in step 231, a convolution feature vector of each candidate bounding box is obtained, and the convolution feature vector is used to represent the picture content feature in a corresponding bounding box.

[0046] It can be understood that most of the current target detection models are convolutional neural network (CNN) based models. In a CNN, convolution processing is first performed on the to-be-identified picture to obtain a convolution feature map, which reflects the abstract feature of the picture content. The target detection models then perform target classification and bounding box regression based on the convolution feature map. Accordingly, each of the candidate bounding boxes output by the target detection models may correspond to a sub-region in the convolution feature map. By performing feature extraction on each sub-region, a convolution feature vector corresponding to each candidate bounding box may be obtained. It is assumed that the first candidate bounding box corresponds to a first convolution feature vector F1 and the second candidate bounding box corresponds to a second convolution feature vector F2.

[0047] Next, in step 232, based on the first convolution feature vector F1 and the second convolution feature vector F2, an association feature A_{1-2} between the first candidate bounding box and the second candidate bounding box with respect to the picture content is determined.

[0048] In one embodiment, the difference between the first convolution feature vector F1 and the second convolution feature vector F2 may be calculated, and based on the difference, the association feature A_{1-2} between the first candidate bounding box and the second candidate bounding box with respect to the picture content is determined.

[0049] In another embodiment, a dot product between the first convolution feature vector F1 and the second convolution feature vector F2 may be calculated, and based on the dot product result, the association feature A_{1-2} between the first candidate bounding box and the second candidate bounding box with respect to the picture content is determined.

[0050] Specifically, assuming that F1= (a1, a2, …, an); F2= (b1, b2, …, bn),

[0051] according to the above example, the content association feature A_{1-2} may be defined as:

[0052] A_{1-2} = (a1*b1, a2*b2, …, an*bn).

[0053] The content association feature A_{1-2} so calculated is a vector.

[0054] In still another embodiment, a dot product between F1 and F2 may be calculated, and based on the dot product result, the association feature A_{1-2} between the first candidate bounding box and the second candidate bounding box with respect to the picture content is determined as A_{1-2}=a1*b1+a2*b2+…+an*bn, and the content association feature A_{1-2} so calculated is a scalar.

[0055] In another embodiment, a distance D between the first convolution feature vector F1 and the second convolution feature vector F2 may be calculated, and based on the distance D,

the content association feature is determined. More specifically, the cosine distance or Euclidean distance between F1 and F2 may be calculated as the above-described distance D.

[0056] In other embodiments, other mathematical operations may also be performed on the basis of the two convolution feature vectors to obtain a corresponding association feature between the first candidate bounding box and the second candidate bounding box with respect to the picture content.

[0057] Then, in step 233, the integrated feature vector of the first candidate bounding box is determined according to at least the content association feature $A_{1\text{-}2}$. It can be understood that the above content association feature $A_{1\text{-}2}$ represents the similarity in content between the first candidate bounding box and a certain non-specific other candidate bounding box (the second candidate bounding box). When the content association feature is a vector, it may be used as a part of the integrated feature vector of the first candidate bounding box; when the content association feature is a scalar, it may be used as a vector element of the integrated feature vector of the first candidate bounding box.

[0058] The process of calculating the content association feature is described above in connection with a certain other candidate bounding box (the second candidate bounding box). It can be understood that the process may be applied to calculating a content association feature between the first candidate bounding box and any of the other candidate bounding boxes.

[0059] As described above, in addition to the association with respect to the picture content, the association relationship between bounding boxes may also include a location association relationship. The following description takes the first candidate bounding box and the second candidate bounding box as an example again.

[0060] In one embodiment, to determine the integrated feature vector of the first candidate bounding box, a location parameter of the first candidate bounding box and a location parameter

of the second candidate bounding box are first obtained. It can be understood that detection results output by the target detection models will include the location parameter of each candidate bounding box. In general, the location parameter of a candidate bounding box is often in the form of (x, y, w, h), where (x, y) represents the center coordinates of the candidate bounding box, w represents the width of the bounding box, and h represents the height of the bounding box. Therefore, the location parameter of the first candidate bounding box and the location parameter of the second candidate bounding box may be easily extracted from the detection result set.

[0061] Then, a location association feature between the first candidate bounding box and the second candidate bounding box is determined on the basis of the location parameter of the first candidate bounding box and the location parameter of the second candidate bounding box.

[0062] According to different embodiments, the location association feature includes one or more of the following: a center-to-center distance, an intersection over union (IoU) ratio, an area ratio, an inclusion relationship, and so on. Here, the IoU ratio represents a ratio of the area of overlap of two candidate bounding boxes to the area of union of the two candidate bounding boxes; the inclusion relationship may be that the first candidate bounding box includes/excludes the second candidate bounding box, or the second candidate bounding box includes/excludes the first candidate bounding box, and so on.

[0063] The integrated feature vector of the first candidate bounding box is then determined according to at least the location association feature. Similarly, the location association feature, determined as above, between the first candidate bounding box and the second candidate bounding box may be used as a part of the vector element of the integrated feature vector of the first candidate bounding box.

[0064] The process of calculating the location association feature is described above in connection with a certain other candidate bounding box (the second candidate bounding box), and

it can be understood that the process may be applicable to calculating a location association feature between the first candidate bounding box and any of the other candidate bounding boxes.

[0065] Furthermore, according to an embodiment, the association relationship between bounding boxes may further include a category prediction result association relationship. The following description takes the first candidate bounding box and the second candidate bounding box as an example again.

[0066] In an embodiment, in order to determine the integrated feature vector of the first candidate bounding box, firstly, the category prediction results of the first candidate bounding box and the second candidate bounding box are obtained and respectively denoted as a first prediction result and a second prediction result, and based on the first prediction result and the second prediction result, a prediction result association feature between the first candidate bounding box and the second candidate bounding box is then determined.

[0067] In general, the category prediction result at least includes a predicted target category. For example, in FIG. 3b, there is a number at the upper left corner of each candidate bounding box indicating the predicted target category, such as 12, indicating scratch damage. Other numbers indicate other types of damage. Correspondingly, the first prediction result includes a first category, and the second prediction result includes a second category; and the prediction result association feature determined based on this may include a comparison result of the first category and the second category, that is, “the same” or “different”.

[0068] In some cases, the predicted target categories are organized in a hierarchical structure. For example, category A refers to a scratch, where A1, A2, and A3 represent mild, medium, and severe scratches, respectively. In this case, the prediction result association feature between the first candidate bounding box and the second candidate bounding box may further include an upper category to which both the first category and the second category belong.

[0069] In one embodiment, the target detection models also output predicted confidences together with the predicted categories of the candidate bounding boxes. That is, the first prediction result includes a first category and a first confidence, and the second prediction result includes a second category and a second confidence. In such a case, the prediction result association feature between the first candidate bounding box and the second candidate bounding box may include the likeness or unlikeness of the first category and the second category, and the difference between the first confidence and the second confidence.

[0070] The process of calculating an association feature is described above in connection with the first candidate bounding box and a specific other candidate bounding box (the second candidate bounding box), and the association feature may include one or more of a content association feature, a location association features, and a prediction result association feature. Such an association feature may be used as a basis for determining the integrated feature vector of the first candidate bounding box.

[0071] With regard to the other candidate bounding boxes described above, different options are also possible in different embodiments. In one embodiment, for the first candidate bounding box, an association relationship between the first candidate bounding box and each of the other candidate bounding boxes of the N candidate bounding boxes may be separately determined, based on which the integrated feature vector is determined. For example, four target detection models are used and a total of 100 candidate bounding boxes (N=100) are detected. Then, for any first candidate bounding box among the detected candidate bounding boxes, for example, candidate bounding box 1, 99 groups of association features between the candidate bounding box and the other 99 candidate bounding boxes can be separately calculated, and the 99 groups of association features are joined or integrated together to obtain an integrated feature vector of the first candidate bounding box.

[0072] In one embodiment, in order to make each obtained integrated feature vector have the same length, the association features may be counted and combined after separately calculating each group of association features with respect to each of the other candidate bounding boxes. For example, the average value of each group of association features is calculated, and the maximum value and minimum value are obtained, etc., and these statistical values are used as vector elements of the integrated feature vector.

[0073] In another embodiment, a part of all other candidate bounding boxes may be selected to calculate the association features. In one example, candidate bounding boxes having a center-to-center distance from the first candidate bounding box falling within a predetermined threshold are selected, and association features between these candidate bounding boxes and the first candidate bounding box are then calculated. In another example, a predetermined number of candidate bounding boxes ranked based on a center-to-center distance from the first candidate bounding box are selected. In other words, a predetermined number (e.g., 10) of candidate bounding boxes having the closest center-to-center distances from the first candidate bounding box are selected, and association features between these candidate bounding boxes and the first candidate bounding box are then calculated. In this way, only the association relationship between the first candidate bounding box and some candidate bounding boxes around it is considered, and the candidate bounding boxes which are far away are ignored, thereby reducing the calculation amount and improving efficiency.

[0074] As described above, the integrated feature vector of the first candidate bounding box is determined according to the association relationship between the first candidate bounding box and the other candidate bounding boxes. It can be understood that, for each of the N candidate bounding boxes, the above-described process may be performed, so that the integrated feature vector corresponding to each candidate bounding box is obtained.

[0075] Referring back to FIG. 2, in step 24, the integrated feature vectors corresponding to the candidate bounding boxes are separately input into a pre-trained classification model, and the detection result set is optimized according to output results of the classification model.

[0076] The classification model may implement classification prediction by using multiple model structures and model algorithms. For example, the classification model may be a Gradient Boosting Decision Tree (GBDT) model, a Support Vector Machine (SVM) model, a Deep Neural Network (DNN) based model, or the like.

[0077] In one embodiment, the classification model is pre-trained by positive samples including, for example, a plurality of damage regions labeled as true damage, and negative samples including a plurality of damage regions labeled as true damage and at least one region labeled as non-damage, so that the classification model learns, from the positive samples, the characteristics of the association features among the multiple true damage regions, and learns, from the negative samples, characteristics of the association features between the non-damage region and the true damage regions.

[0078] As described above, in step 22, a plurality of target detection models are used to identify the vehicle damage picture to obtain preliminary detection results. These target detections also require the use of a large number of labeled pictures as a training sample set for training. These labeled pictures are manually labeled to indicate true damage objects. Usually, the training sample set also includes some negative samples, that is, pictures of non-damage regions. Therefore, the classification model may reuse the training sample set for training, that is, obtain therefrom a plurality of damage regions labeled as true damage as positive samples, obtain a plurality of damage regions labeled as true damage and at least one region labeled as non-damage as negative samples, and the classification model is trained by using the positive and negative samples. This is especially suitable for the training of a two-category model.

[0079] The trained classification model performs classification prediction on a corresponding candidate bounding box based on the input integrated feature vector. Typically, the classification result may be a two-category result, for example, whether the corresponding candidate bounding box is a true and accurate damage bounding box. In an embodiment, the classification result may also be a multi-category result. For example, category 1 corresponds to bounding box location and damage category prediction being both accurate, category 2 corresponds to the bounding box including damage but the damage category prediction being inaccurate, and category 3 corresponds to a non-damage bounding box, and so on.

[0080] Based on the output result of the classification model, the detection results may be optimized.

[0081] In one embodiment, if the classification result indicates that the damage category prediction of a candidate bounding box is inaccurate (for example, corresponding to category 2 in the multi-category result), the candidate bounding box may be labeled as a pending region and then sent to a manual platform for re-checking.

[0082] In another example, a further model algorithm may be employed to correct the candidate bounding box that is inaccurately predicted. For example, the prediction result of the candidate bounding box is corrected by referring to a prediction result of another candidate bounding box closest to the candidate bounding box and having an accurate prediction indicated by a classification result. For example, the center-to-center distances between candidate bounding box 1 and candidate bounding boxes 2, 3 and 4 are within a preset range, and the distances are relatively close. The damage category prediction results of candidate bounding boxes 2, 3, and 4 are all category 12-scratch damage, and the classification results of the candidate bounding boxes, given by the classification model, indicate that the prediction results are accurate. However, the prediction result of candidate bounding box 1 is category 11-deformation, and the classification

result of the candidate bounding box, given by the classification model, indicates the prediction result is inaccurate. Considering the continuity of the damage object, the prediction result of the candidate bounding box 1 may be corrected by referring to the prediction results of candidate bounding boxes 2, 3 and 4, thereby optimizing the detection result set.

[0083] In another embodiment, the classification result output by the classification model based on the integrated feature vector of the first candidate bounding box indicates that the first candidate bounding box is an abnormal bounding box, i.e., not a true and accurate damage bounding box. Then, the first candidate bounding box may be removed from the detection result set to optimize the damage detection results.

[0084] In an embodiment, for the multiple candidate bounding boxes shown in FIG. 3b, according to step 23 to step 24, the association features between each candidate bounding box and other candidate bounding boxes are calculated and input into the classification model to determine the accuracy of the detection results of the candidate bounding boxes. Assuming that the classification model output shows that the rightmost candidate bounding box is an abnormal bounding box, the rightmost candidate bounding box may be removed from the detection results to obtain the optimized damage detection results shown in FIG. 3c. It may be found that the rightmost candidate bounding box is actually light reflected from the ground, while the other candidate bounding boxes contain continuous scratches. Therefore, the association relationship with other candidate bounding boxes, calculated for the rightmost candidate bounding box, exhibits abnormal features, which are then embodied in the integrated feature vector. Thus, based on the integrated feature vector, the classification model gives a classification result that the candidate bounding box is abnormal.

[0085] In view of the above-described damage identification process, it is concluded that on one hand, multiple target detection models are used at the same time to perform damage

identification and detection on a vehicle damage picture, which makes generated detection results complement each other, and therefore reduces the possibility of missed detection to the greatest extent. Further, for the N candidate bounding boxes obtained by the plurality of target detection models, based on the association relationship between each candidate bounding box and other candidate bounding boxes, connections are established among the candidate bounding boxes, and based on such connection results, a trained model is further used to determine the accuracy of each candidate bounding box, thereby eliminating bounding boxes of false detection and further optimizing the damage identification results. In this way, the accuracy of damage identification can be comprehensively improved.

[0086] According to an embodiment, an apparatus for vehicle damage identification is provided. FIG. 5 shows a schematic block diagram of an identification apparatus according to an embodiment. It can be understood that the apparatus may be implemented by any apparatus, device, platform, and device cluster having computing and processing capabilities. As shown in FIG. 5, the apparatus 500 includes: an obtaining unit 51 configured to obtain a vehicle damage picture to be identified; a detecting unit 52 configured to input the vehicle damage picture into a plurality of pre-trained target detection models respectively, and obtain groups of corresponding detection results from the plurality of target detection models, wherein the groups of detection results constitute a detection result set, and the detection result set includes N candidate bounding boxes detected by the plurality of target detection models and category prediction results of the candidate bounding boxes; an integrating unit 53 configured to, for any first candidate bounding box of the N candidate bounding boxes, determine an integrated feature vector of the first candidate bounding box according to at least an association relationship between the first candidate bounding box and other candidate bounding boxes, wherein the association relationship at least includes one or more of the following: a bounding box content association relationship, a location association

relationship, and a category prediction result association relationship; and an optimizing unit 54, configured to separately input the integrated feature vectors corresponding to the candidate bounding boxes into a pre-trained classification model, and optimize the detection result set according to output results of the classification model.

[0087] The following description is made in conjunction with the embodiments in which the other candidate bounding boxes include a second candidate bounding box.

[0088] According to an embodiment, the association relationship includes a bounding box content association relationship. In such a case, the integrating unit 54 may further include (not shown): a convolution feature obtaining module configured to obtain a convolution feature vector of each candidate bounding box, wherein the convolution feature vector is used to represent a picture content feature in a corresponding bounding box; an association feature determining module configured to determine a content association feature between the first candidate bounding box and the second candidate bounding box based on a first convolution feature vector corresponding to the first candidate bounding box and a second convolution feature vector corresponding to the second candidate bounding box; and an integrated feature determining module, configured to determine an integrated feature vector of the first candidate bounding box according to at least the content association feature.

[0089] Further, in one embodiment, the association feature determining module may be configured to perform at least one of the following: calculating a dot product of the first convolution feature vector and the second convolution feature vector, and determining the content association feature based on a dot product result; and calculating the distance between the first convolution feature vector and the second convolution feature vector, and determining the content association feature based on the distance, wherein the distance includes a cosine distance or a Euclidean distance.

[0090] According to an embodiment, the association relationship includes a location association relationship. In such a case, the integrating unit 53 may be configured to: determine a location association feature between the first candidate bounding box and the second candidate bounding box based on a location parameter of the first candidate bounding box and a location parameter of the second candidate bounding box; and determine, according to at least the location association feature, the integrated feature vector of the first candidate bounding box.

[0091] Further, in one embodiment, the location association feature includes one or more of the following: a center-to-center distance, an IoU (Intersection-over-Union) ratio, an area ratio, and an inclusion relationship.

[0092] According to an embodiment, the association relationship includes a category prediction result association relationship. In such a case, the integrating unit 53 may be configured to: determine a prediction result association feature between the first candidate bounding box and the second candidate bounding box based on a first prediction result corresponding to the first candidate bounding box and a second prediction result corresponding to the second candidate bounding box; and determine, according to at least the prediction result association feature, the integrated feature vector of the first candidate bounding box.

[0093] Further, in one embodiment, the first prediction result includes a first category and a first confidence, the second prediction result includes a second category and a second confidence, and the prediction result association feature includes: whether the first category is the same as the second category, and a difference between the first confidence and the second confidence.

[0094] In one embodiment, the other candidate bounding boxes include one or more of the following: candidate bounding boxes having a center-to-center distance from the first candidate bounding box falling within a predetermined threshold, and candidate bounding boxes having a

center-to-center distance from the first candidate bounding box ranked within a predetermined number.

[0095] In one embodiment, the classification model includes one of the following: a GBDT model, an SVM model, and a DNN based classification model.

[0096] In one embodiment, the optimizing unit 54 is configured to remove the first candidate bounding box from the detection result set when the output result of the classification model for the first candidate bounding box is an abnormal bounding box.

[0097] In this way, vehicle damage can be efficiently and accurately identified through the apparatus 500.

[0098] According to an embodiment, a computer-readable storage medium is provided, having stored therein a computer program for causing a computer to perform the method as illustrated in FIG. 2 when the computer program is executed in a computer.

[0099] According to an embodiment, a computing device is provided, including a memory and a processor, wherein the memory stores executable codes, and when the processor executes the executable codes, the method illustrated in FIG. 2 is implemented.

[0100] Those skilled in the art should appreciate that in one or more examples described above, the functions described may be implemented by hardware, software, firmware, or any combination thereof. When implemented by software, the functions may be stored in a computer-readable medium or transmitted as one or more instructions or codes on a computer-readable medium.

[0101] The specific embodiments described above explain the objectives, technical solutions, and beneficial effects of the present invention. It should be understood that the above description is only specific embodiments of the present invention, and is not intended to limit the scope of the present invention. Any modifications, equivalents, improvements, and the like made

on the basis of the technical solutions of the present invention should be included in the scope of the present invention.