IMAGE ANALYZATION METHOD WITH UNRECOGNIZED OBJECT MEMORY MECHANISM

Abstract

An image analyzation method with an unrecognized object memory mechanism are provided. Initially, whether a target object belongs to a known category using an object recognition model is determined; if not, then the object memory mechanism is executed. The mechanism first obtains image feature information through a feature extraction algorithm, and generates an attention representative value by a feature classifier. The target objects are grouped according to the numerical range it belongs to and stored in corresponding specified directories. When the number of target images in a directory reaches a quantity threshold, a new category label is assigned to the target images and used as retraining data for the object recognition model. The image analyzation method classifies unrecognized objects and enables the object recognition model to recognize unrecognized objects through learning, without the need for manual labeling of unrecognized objects, thereby saving resources and cost.

Inventors:	**LIN; Wei-Lun (Taichung City, TW), WANG; Shao-Mao (Taichung City, TW)**
Applicant:	**FENG CHIA UNIVERSITY** (Taichung City, TW)
Family ID:	**1000007838657**
Assignee:	**FENG CHIA UNIVERSITY (Taichung City, TW)**
Appl. No.:	**18/581941**
Filed:	**February 20, 2024**

Publication Classification

Int. Cl.: **G06V10/44** (20220101); **G06T7/12** (20170101); **G06V10/764** (20220101); **G06V20/70** (20220101)

U.S. Cl.:

Background/Summary

BACKGROUND OF THE INVENTION

1. Field of the Invention

[0001] The present invention relates to an image analyzation method, and more particularly to an image analyzation method with an unrecognized object memory mechanism.

2. Description of the Related Art

[0002] The conventional image object detection technologies can generally be divided into two application directions. The first direction is to identify the categories of objects in an image; the other direction is to label or separate the areas belonging to the same object in the image.

[0003] In the applications of identifying object category in images, for instance, identification is performed by using a trained large-scale object recognition model to determine if there are known objects in the image, and marking the detected objects with detection boxes as the identification result. The generation of the large-scale object recognition model requires training the machine learning model with a large amount of labeled training image data, so that it can recognize objects in new images that match the labeled categories.

[0004] In the applications of labeling or separating objects in images, for example, labeling or separating is performed by using semantic segmentation algorithms to mark the pixel blocks of the same object in the image. Semantic segmentation usually differentiates pixel blocks from the background based on image properties such as texture, size, color, shape, etc. However, semantic segmentation algorithm cannot identify the category of objects, meaning it cannot determine the similarity of objects in multiple images or create associations among objects in different images.

[0005] In some combined applications, detection boxes from large-scale object recognition model and object masks from semantic segmentation algorithms are generated based on the same image. Furthermore, the Intersection Over Union (IOU) between the detection boxes and object masks is calculated to determine the accuracy of the object recognition by the large-scale object recognition model.

[0006] However, when an image to be processed contains any object that does not belong to any of the known categories of the trained large-scale object recognition model, the model cannot identify such objects. Training a large-scale object detection system to recognize unrecognized objects requires a large amount of labeled training data, and providing a large amount and high-quality training data is quite costly in terms of both expenses and workload. On the other hand, semantic segmentation algorithms can only separate objects from the background without identifying their categories. The aforementioned combined applications can only judge the accuracy of the large-scale object recognition model, and when the large-scale object recognition model fails to effectively identify objects, it still cannot further determine the category of object in the image.

[0007] In conclusion, conventional image analyzation technologies for object detection must be further improved in the aspect of recognizing unrecognized objects.

SUMMARY OF THE INVENTION

[0008] Given the challenge that existing image object detection technologies face in effectively recognizing the categories of new, untrained objects in input images, and the high cost of training models to recognize unrecognized objects, an objective of the present invention is to provide an image analyzation method with an unrecognized object memory mechanism. To achieve the foregoing objective, the image analyzation method includes: [0009] reading a target image, wherein the target image includes a target object; [0010] determining whether the target object is a

known category object through an object recognition model; [0011] if the target object does not belong to a known category, executing an object memory mechanism, comprising: [0012] generating an image feature information of the target object through a feature extraction algorithm; [0013] performing a feature classification process on the image feature information to produce an attention representative value, and determining a numerical range to which the target object belongs; [0014] storing the target image in a specified directory corresponding to the numerical range in an image database; [0015] determining whether a number of at least one stored image in the specified directory exceeds a quantity threshold; [0016] if the number of the at least one stored image exceeds the quantity threshold, assigning a new category label to the at least one stored image, and inputting the at least one stored image after labeling, as training data into the object recognition model in a retraining process; wherein the object recognition model recognizes the target object after the retraining process.

[0017] The image analyzation method with an unrecognized object memory mechanism of the invention generates the attention representative value by extracting and classifying features of the target object in the target image, and stores the target image in different specified directories based on the numerical range of the attention representative value, thus performing the clustering based on similar features of the target object. When the number of stored images in a specified directory reaches a quantity threshold, a new category label is assigned to the stored images with similar features, which are then used as training data for retraining the object recognition model. Through such process, the object recognition model can recognize objects in new input images having features similar to the target object. This method provides an object recognition model, which originally could not recognize a new type of object, with a mechanism similar to human memory and learning. After training, the object recognition model is able to recognize new types of objects based on their image features, without the need for manual labeling, thus saving manpower and cost.

[0018] Other objectives, advantages and novel features of the invention will become more apparent from the following detailed description when taken in conjunction with the accompanying drawings.

Description

BRIEF DESCRIPTION OF THE DRAWINGS
[0019] FIGS. **1**A to **1**B are flowcharts of the image analyzation method with an unrecognized object memory mechanism according to the invention.
[0020] FIG. **2** is a block diagram of the image analyzation device with an unrecognized object memory mechanism.
DETAILED DESCRIPTION OF THE INVENTION
[0021] With reference to FIGS. **1**A, **1**B and FIG. **2**, the image analyzation method with an unrecognized object memory mechanism **100** of the present invention is executed by an image analyzation device **10** with an unrecognized object memory mechanism. The image analyzation method **100** includes steps S**101** to S**105**. The image analyzation device **10** includes a processor **11** and a storage device **12**, where the storage device **12** is used to store at least one stored image, with each stored image under a specified directory corresponding to different numerical ranges.
[0022] The processor **11**, for example, is a central processing unit (CPU), an image processing unit (GPU), a digital signal processor (DSP), etc.; the storage device **12**, for example, is read-only memory (ROM), random-access memory (RAM), a solid-state drive (SSD), flash memory, etc. However, the invention is not limited to these.
[0023] With reference to FIG. **1**A, the steps S**101** to S**104** of the image analyzation method with an unrecognized object memory mechanism are further explained below.

[0024] In step S**101**, the processor **11** receives a target image that contains a target object.

[0025] More specifically, the processor **11** has an input/output interface for receiving information and outputting computational results. The target object, for example, is identified by the processor performing a semantic segmentation process on the received target image, classifying multiple pixels within the target image, producing at least one object mask, and defining the pixels within the range of the object mask as the target object. Preferably, the object mask generated by the semantic segmentation process is an irregular mask.

[0026] The semantic segmentation process produces object masks based on features such as texture, color, edges, shape, and size in the target image. Preferably, the semantic segmentation model is any one of the following: a Segment Anything Model (SAM), a Hybrid Gene Algorithm (HGA) model, and a Mask Region-Based Convolutional Neural Networks (R-CNN) model, but the invention is not limited to these.

[0027] In step S**102**, the processor **11** inputs the target object into an object recognition model to determine whether the target object is a known category object. The object recognition model is a model pre-trained with multiple labeled data and capable of recognizing at least one category of object based on features such as texture, color, edges, shape, and size in images. When the input target object belongs to one of the object categories recognizable by the object recognition model, the object recognition model outputs corresponding object category information, as in step S**103**. If the input target object does not belong to any of the recognizable categories by the object recognition model, the object recognition model outputs a recognition failure information. Hence, the processor determines that the target object does not belong to any known category, and thus executes the object memory mechanism, as in step S**104**. The purpose of the object memory mechanism is to enable the object recognition model to understand, learn, or remember the features of the target object or similar objects, so that the object memory mechanism is able to recognize similar objects upon input of new target images and determine the new target images to be of the same category as the target object.

[0028] With reference to FIG. **1**B, the object memory mechanism includes steps S**1041** to S**1046**, which will be further explained in detail below.

[0029] In step S**1041**, the processor **11** inputs the target object from the target image into a feature extraction algorithm to generate image feature information. The feature extraction algorithm may include at least one of a Principal Components Analysis (PCA), deep learning neural network autoencoders, Scale-Invariant Feature Transform (SIFT), etc., but the invention is not limited to these.

[0030] In step S**1042**, the processor **11** generates an attention representative value through a feature classification process on the image feature information and determines the numerical range of the attention representative value. More specifically, this step first generates an attention value through a feature classifier, and then converts the attention value into an attention representative value through a normalization function. For example, the normalization function could be a SoftMax function, with its domain ranging from 0 to 1. Hence, the attention representative value is a real number between to 1, represented as a percentage of the attention representative value. The feature classifier, for instance, could be any one of a Convolutional Neural Network (CNN) model, a Long-Short Term Memory Neural Network (LSTM) model, or an attention mechanism of a Transformer model. Preferably, the feature classifier is a self-attention mechanism of a Transformer model.

[0031] Before determining the numerical range of the attention representative value, the domain of the normalization function is divided into multiple numerical ranges based on a preset number of intervals, which can be determined based on system design, preset image types, or similarity requirements. For example, if the preset number of intervals is 10, then the domain 0% to 100% (0 to 1) is divided into intervals of 10% each, with each numerical range having an upper limit and a lower limit, e.g., 0% to 17 9.999% (first interval), 10% to 19.999% (second interval), . . . 90% to

99.999% (tenth interval). When the attention representative value is between an upper limit and a lower limit of one of these intervals, it is determined to belong to that numerical range. For instance, if the attention representative value is 56%, it belongs to the 50% to 59.999% (sixth interval).

[0032] In step S**1043**, the processor **11** stores the target image in the specified directory corresponding to the determined numerical range in the storage device **12** as a stored image. In the aforementioned example, where the preset number of intervals is 10, there are ten directories corresponding to the first to the tenth intervals, and the target image with an attention representative value of 56% stored in the sixth directory. This step stores target images with attention representative values of the same numerical range in the same specified directory. Target objects in the stored images within the same specified directory, having attention representative values in the same numerical range, are defined as belonging to the same category of objects.

[0033] In step S**1044**, the processor **11** determines whether the number of at least one stored image in the specified directory exceeds a quantity threshold. In step S**1045**, when the number of stored images in the specified directory exceeds the quantity threshold, a new category label is assigned to the at least one stored image containing the target image, and the labeled stored images are inputted to the object recognition mode as training data in a retrain process of the object recognition model. The retraining process enables the object recognition model to recognize the target object and similar object image after the retraining. Conversely, in step S**1046**, when the number of stored images in the specified directory does not exceed the quantity threshold, the process for this target image is concluded. When a new target image is received, steps S**101** to S**104** are executed again.

[0034] In step S**1044**, when the number of stored images in the specified directory accumulates to a sufficient amount to serve as training data for the object recognition model's quantity threshold, the processor **11** assigns a new category label to these stored images. In step S**1045**, the stored images are used as training data in the retraining process of the object recognition model. If the number of target images in the specified directory has not yet exceeded the quantity threshold, the storage device **12** continues to accumulate stored images that have been clustered in each specified directory.

[0035] In summary, the image analyzation method with an unrecognized object memory mechanism according to the present invention provide a learning and memory mechanism for unrecognized objects that are originally unrecognizable by the object recognition model. Target images are grouped based on the numerical range of the attention value generated by the feature classifier. Grouped stored images that are accumulated by a certain amount are labeled and utilized as training data to enable the object recognition model to "learn and remember" the image features of such new category objects. Consequently, the model becomes capable of recognizing similar target objects in the future, achieving a computer-automated learning and memory mechanism similar to how humans learn and remember unrecognized objects in images.

[0036] Even though numerous characteristics and advantages of the present invention have been set forth in the foregoing description, together with details of the structure and function of the invention, the disclosure is illustrative only. Changes may be made in detail, especially in matters of shape, size, and arrangement of parts within the principles of the invention to the full extent indicated by the broad general meaning of the terms in which the appended claims are expressed.

Claims

1. An image analyzation method with an unrecognized object memory mechanism, the method comprising: reading a target image, wherein the target image includes a target object; determining whether the target object is a known category object through an object recognition model; if the target object does not belong to a known category, executing an object memory mechanism, comprising: generating an image feature information of the target object through a feature

extraction algorithm; performing a feature classification process on the image feature information to produce an attention representative value, and determining a numerical range to which the target object belongs; storing the target image in a specified directory corresponding to the numerical range in an image database; determining whether a number of at least one stored image in the specified directory exceeds a quantity threshold; and if the number of the at least one stored image exceeds the quantity threshold, assigning a new category label to the at least one stored image, and inputting the at least one stored image after labeling, as training data into the object recognition model in a retraining process; wherein the object recognition model recognizes the target object after the retraining process.

2. The image analyzation method with the unrecognized object memory mechanism as claimed in claim 1, further comprising: performing a semantic segmentation process on the target image to classify multiple pixels therein, producing at least one object mask; and defining a pixel collection within a range of the object mask in the target image as the target object.

3. The image analyzation method with the unrecognized object memory mechanism as claimed in claim 2, wherein the semantic segmentation process is performed by any one of the following: a segment anything model (SAM), a hybrid gene algorithm (HGA) model, and a mask region-based convolutional neural network (R-CNN) model.

4. The image analyzation method with the unrecognized object memory mechanism as claimed in claim 2, wherein: the object mask is an irregular mask.

5. The image analyzation method with the unrecognized object memory mechanism as claimed in claim 1, wherein the feature classification process comprises: generating an attention value through a feature classifier; and converting the attention value into an attention representative value through a normalization function.

6. The image analyzation method with the unrecognized object memory mechanism as claimed in claim 5, further comprising: dividing the domain of the normalization function into a preset number of numerical ranges based on a preset number of intervals, each numerical range having an upper limit and a lower limit, and the numerical ranges being non-overlapping; wherein, when the attention representative value is between the upper and lower limits of one of the numerical ranges, determining that the attention representative value belongs to the one of the numerical ranges.

7. The image analyzation method with the unrecognized object memory mechanism as claimed in claim 1, wherein: when the object recognition model outputs an object category information as a result, the target object is a known category object.

8. The image analyzation method with the unrecognized object memory mechanism as claimed in claim 5, wherein the feature classifier is any one of the following: a convolutional neural network model, a long short-term memory neural network model, and an attention mechanism of a transformer model; and the normalization function is a Softmax function.