Research Advanced in Deep Learning Object Detection

Mingzhe Zhou
Northeastern University
Shenyang, China
20217639@stu.neu.edu.cn

Abstract: Object detection aims to locate and assign various predefined substances from images to corresponding classifications. Benefit from the rapid evolution of deep learning, target detection algorithms based on convolutional neural networks have achieved breakthroughs in both accuracy and efficiency. Based on the detailed literature research and analysis, in this paper, we make an integrated assess of the research progress of object detection. Specifically, we first introduce the existing representative algorithms from two-stage detection framework to one-stage detection framework. Then a battery of experiments are conducted to analyze the implementation of different detection algorithms on some common datasets. Finally, we summarize the main challenges and give an outlook on future research development for object detection.

Keywords—Object detection; Deep Learning; Anchor-based detection;

I. Introduction

With the persistent evolution of multimedia and Internet technologies, a large amount of visual data such as pictures and videos has been accumulated on the Internet, which contains a variety of useful and valuable information. For these massive visual data, the rapid and effective mining of the information has become an pressing issue to be solved, which has promoted the related research of computer vision algorithms. As a classic task in computer vision, object detection tends to distinguish and localize one or more objects (such as cars, faces, road animals, etc.) from images or videos. Compared with the ease with which humans can identify and locate specific types of objects, it was only in recent years that object detection algorithms benefited from the development of deep learning and artificial intelligence, and gradually matured. At present, object detection algorithms have been successfully and widely used in diverse fields and achieved exciting results, such as face detection, industrial product detection, AI driving, etc.

Early object detection algorithms transformed the detection problem into a classification problem, relying on designing handcrafted features and sliding windows to extract high-quality candidate regions. Usually,

There are some key steps in traditional approaches, such as preprocessing, window sliding, feature extraction, feature selection, feature classification and post-processing. Among them, the window size, sliding method and strategy have a great significance on the quality of feature extraction. Representative approaches mainly include V-J detection algorithm, HOG+SVM detection algorithm, DPM algorithm, etc. These approaches achieve good detection accuracy, but in most cases, they are only suitable for situations with obvious features and relatively simple image backgrounds. In our practical application, the situation is complex and changeable, and it is difficult to detect objects with general abstract features. In addition, traditional approaches use sliding windows for region selection, which has relatively high running time and cost [1].

As deep learning improves in such a rapid speed, detection algorithms have shifted from traditional approaches to more advanced techniques on the basics of deep convolutional neural networks (CNNs) [2]. It relies on a large amount of training data and autonomously learns the feature label of the object to be detected through a convolutional neural network. The advantages of strong learning ability (fitting any complex function), strong representation ability, and strong adaptability make it gradually occupy the mainstream position in object detection. According to how the candidate regions are produced in the whole network, the deep learning-based object detection framework can be further divided into two-stage and one-stage object detection algorithms. For the two-stage detection algorithm, the basic idea is similar to the traditional detection method. First, a certain number of object candidate regions are extracted, and then the convolutional neural network is used to distinguish and locate the candidate regions. In contrast to traditional approaches, the two-stage detection method uses convolutional features to replace handcrafted features, which significantly increases the feature expression ability of the object. In addition, the two-stage detection method can also dynamically fine-tune the position of the object prediction box through position box regression during training. Two-stage detectors usually require more time to extract all proposals, which usually have complex structures and lack global information, while one-stage detectors directly identify and localize objects through dense sampling, adopting predefined boxes of different scales and sizes to identify and localize objects. Locate the object.

Focusing on the three main technical frameworks of the two-stage and one-stage detection approaches introduced above, this paper first analyzes and summarizes the research progress and stages of object detection algorithms based on deep learning. Second, we detail general datasets for object detection and the result of different algorithms on mainstream datasets as well. Finally, the future potential development in the field of object detection is prospected

II. Two-stage Object Detection Algorithms

Though have fully development, the traditional approach, however, facing the problem of the low accuracy. Therefore to deal with it, Girshick et al. come up a deep learning-based object detection algorithm R-CNN [3]. As the Figure 1 shown, Here are the key process in R-CNN algorithm: 1) Use selective search (selective search) algorithm from the image to be detected, which contains the object to be detected; 2) Normalize the candidate area, that is, scale all candidate boxes into Fixed size (227×227 pixels); 3) To obtain a settled-length feature vector, a deep convolutional neural network is applied for extract feature from the from the recommend region; 4) Finally, the feature vector is delivered to the support vector machine (SVM). The vector is classified to get the category information, and then the category information is sent to the fully connected network for regression, and the corresponding position coordinate information is obtained. R-CNN successfully introduced the CNN network into object detection for the first time, which leads to a huge increase in a detection rate, from 35.1% to 53.7% (based on PASCAL VOC), opening up a new path for the detection mission.

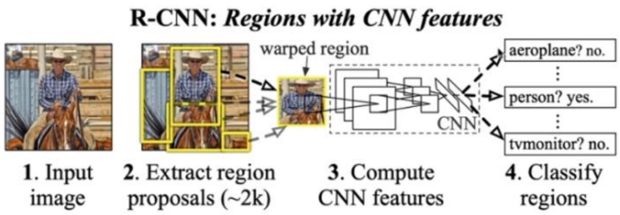

Fig. 1. Framework of R-CNN algorithm [4]

Though the R-CNN algorithm has laid the basic framework of the two-stage object detection algorithm based on deep learning **Error! Reference source not found.**, it still has the shortcomings of large computational load and slow running speed. Specifically, most of the 2k candidate regions are highly overlapping, but these redundant candidate regions will be input into the CNN network to extract features, which restricts the running speed of the algorithm in a large extent. In addition, training and testing cannot be run end-to-end, and candidate region generation, feature extraction, classification, and regression rely on additional information storage and reading. Finally, the candidate regions must undergo deformation operations before being input to the CNN network for feature extraction, which may lead to the loss of image information, which in turn affects the final detection accuracy.

To solve the low detection efficiency of R-CNN, He et al. and Girshick et al. respectively come up spatial pyramid pooling net (SPPNet) and Fast R-CNN algorithms based on R-CNN, respectively. Instead of sending all the candidate regions into the deep convolutional neural network, these two approaches sent the image to the deep network, after which map all the candidate regions in certain layer. These two types of approaches greatly improve the detection efficiency, and also improve the detection accuracy on the PASCAL VOC 2007 dataset from 66% to 70%. An additional Spatial Pyramid Pooling layer was added to the last layer of the R-CNN in SPP network, which converts the input convolutional features into 16×256, 4×256 and 1×256 dimensional feature vectors, respectively, The output of one layer is pooled to obtain an output of fixed length 5376. Through the above operations, SPP can obtain candidate regions and features of the entire image with only one convolution, which greatly reduces the difficulty and computation time. In addition, the image input to the convolutional network does not need to be of fixed size, reducing the deformation loss of image distortion. However, in the SPP algorithm, convolutional feature extraction and full connection, SVM is independent. Therefore, the training loss of the SVM cannot be loaded into the convolutional layer before the SPP layer, so weather the efficiency will be improved is uncertain. In addition, since its core algorithm still uses the R-CNN framework, it still needs a lot of disk resources to run. Similar to SPP, the Fast RCNN network also directly convolves the entire image, and uses ROI pooling to transform the size of the features, avoiding multiple repeated calculations for the same repeated unpacking. Fast R-CNN training speed is about 9 times faster than SPP-Net and about 3 times faster: testing speed is 214 times faster than R-CNN and 11 times faster than SPP-Net. Its mAP on the VOC 2012 dataset is about 66%.

Although Fast R-CNN increases the detection speed, a large part of the time is still spent on the selective search of candidate regions. Therefore, Ren et al. come up a faster R-CNN algorithm, adding a regional proposal network (RPN) on the basis of fast R-CNN. This network extracts candidate regions by setting anchors of different scales, and replaces the traditional candidate region production approaches such as selective search.et. It also makes the end-to-end training come true, at the same time improves the operating speed of the network. As Figure 2 shows, the Faster R-CNN network makes up of four parts: Anchor Generation layer, Region Proposal layer, ROI Pooling layer, Classification layer. In this layer, the convolution layer is used to extract the features of the whole input image and generate the output of the feature map; RPN network for extracting candidate regions, whose input is a feature mapping obtained through a convolutional layer and whose output is diverse candidate regions; ROI pooling layer plays an important role in changing inputs of various sizes into settled-length outputs. The classification and regression layer is used to determine the label to which the candidate region belongs and the exact location of the object in the image. Unlike Fast R-CNN, all steps in Faster R-CNN are under a separate deep learning framework. This innovative approach leads to a huge increase in both speed and the detection accuracy (PASCAL VOC 2007 dataset), the latter one even reached to 73.2%

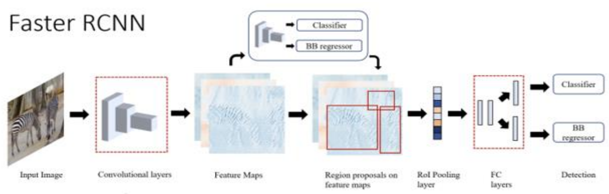

Fig. 2. Framework of Faster R-CNN algorithm

With the continuous evolution of deep learning, however, all the algorithm built on Faster R-CNN could be easily affected by the complexity of the basic network, the number of candidate boxes, the complexity of the classification, and so on. Around some problems existing in Faster R-CNN, many scholars have carried out a series of improvement work. The representative work is as follows:

(1) Mask R-CNN. Faster R-CNN rounds the size of the feature map when performing down sampling and RoI pooling. This approach has little effect on the classification task, but has a serious effect on the accuracy of the detection task position frame. For this reason, He et al. Mask R-CNN is come up, which does not use the rounding operation for the feature map size change in the network, but fills the pixels at non-integer positions through bilinear difference. This makes the

downstream feature map to the upstream feature map without position error, which not only improves the object detection effect, but also enables the algorithm to meet the accuracy requirements of the semantic segmentation task.

(2) RFCN. Dai et al. realized that the network layer behind ROI pool is no longer translation-invariant, and the number of layers behind ROI pool directly affects detection efficiency. Therefore, a region-based complete convolutional network (RFCN) is proposed to solve this problem. This location-sensitive scoring map eliminates the region of interest judgment sub-network and uses a location-sensitive ROI pooling layer to directly distinguish pool results, improving detection accuracy to 80.5% on the PASCAL VOC 2007 dataset.

(3) Soft-NMS and Softer-NMS [6]. The NMS algorithm is an essential post-processing step in most object detection approaches, but the threshold of NMS is hard to determine, which will delete boxes that should not be deleted when it is set too small while will increase the false detection rate when it is set too large. To this end, Bodla et al. come up the Soft-NMS algorithm. Unlike the NMS algorithm, which directly deletes all detection frames with an intersection over union (IOU) greater than the threshold, Soft-NMS sets a new confidence threshold. The lower the score, the more detection frames with a final score greater than the confidence threshold can be retained, which can improve the recall rate of the target detection algorithm. In order to further improve the prediction accuracy of the object position, He et al. proposed the softer NMS algorithm, which uses a new bounding box regression loss called KL loss (Kullback Leibler, KL) to simultaneously learn the shape variables and position deltas of the bounding box. At the same time, Softer-NMS uses KL Loss on the Soft-NMS algorithm based on weight averaging; finally, Softer-NMS algorithm increases the detection accuracy of Faster R-CNN based on VGG-16 from 23.6% on the MS COCO dataset to 29.1%.

III. ONE-STAGE OBJECT DETECTION ALGORITHMS

Although the two-stage object detection algorithm has achieved high accuracy in recognition, it has a large amount of network layers and nodes, and the complex calculation speed makes it difficult to achieve real-time object detection. To this end, the single-stage stage framework came into being, which cancels the candidate region generation and fine-tuning in the two-stage approaches, and directly regresses the objectbox based on the feature map.

To address the poor efficiency of the two-stage target detection algorithm, YOLO (You Only Look Once) emerged [7]. This algorithm realizes feature extraction, candidate box classification and regression directly in the same branchless deep convolutional network. By this way, not only the structure become more simple, the speed also improved. The fps increased from 7 frames/s of Faster R-CNN to 45 frames/s. This significant increase makes it possible for the computer to deal with some real-time detection. The network structure of YOLO is shown in Figure 3. S×S represents the number of grids that the initial image is divided into, S represents the length or width of the image is equally divided into S parts. The core of YOLO is in the last two layers. There is a 4096-dimensional fully connected layer, which is then fully connected to a 7×7×30 dimensional tensor is a 4096-dimensional fully connected layer, which is then fully connected to a 7×7×30 dimensional tensor, followed by a convolutional layer. The whole process does not need to determine the middle candidate area first, where a separate network can complete the determination of the category and the regression of the position.

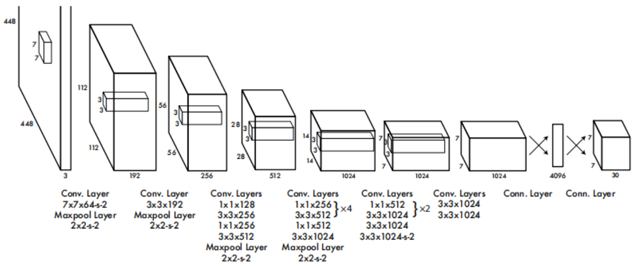

Fig. 3. Network structure of YOLO algorithm

YOLO has a faster detection speed, accordingly, it is not accurate enough in object localization, and its low recall rate was relatively low, so its detection accuracy is low. To solve this problem, Redmon et al. further comes up with the YOLOV2 algorithm [8], which combines batch normalization, high resolution classifier, direct object frame location detection (location prediction), multi-scale training (Multi-scale) and other operations to prefect the detection precision of the detection program, and finally improve the detection accuracy of the program. With the help of the new network, the precision varied from 66. 4% of YOLOv1 to 78. 6% (PASCAL VOC 2007). In addition, YOLOv2 also specifically trains a Darknet-19 network consisting of 19 convolutional layers and 5 maximum pooling layers as the backbone network of the model to extract features and decrease the computational effort of the model.

On the basis of YOLOv2, YOLOv3 [9] uses a newly designed Darknet-53 residual network combined with feature pyramid networks (FPN) for multi-scale mixed prediction to further improve detection accuracy and speed. The basic idea of YOLOv3 is to first use the feature extraction network to gain a feature map in a particular size, and then divide the input image into a corresponding number of grid units. If the centric coordinate of the real object falls on a grid cell, the object is predicted by that grid cell, because each grid cell forecasts a settled number of bounding boxes (using the K-means clustering algorithm in YOLOv2 (K -means) to obtain 3 bounding boxes with different initial sizes), and finally select the bounding box with the largest IOU with the ground truth to predict the object. Compared with Darknet-19 of YOLOv2, the pooling layer that changes the size of the feature map in Darknet-53 of YOLOv3 is basically implemented by the convolution layer, which reduces the computational load of the model. Secondly, the residual blocks in the ResNet network are introduced to deal with the gradient problem caused by overmuch layers of the straight-tube network structure contained in YOLOv2. ResNet's residual structure makes it less difficult to train deep networks, so the network can be made up to 53 layers to improve detection accuracy. These changes make YOLOv3 use 1/3 of the time to achieve an accuracy comparable to SSD. YOLOv4 is the masterpiece of the YOLO network. YOLOv4 [10] improves on the previous network in all aspects: mosaic data enhancement, CIOU loss, FSFOFT-NMS, SPP-Net, CSP Net, and the introduction of CBAM attention mechanism. With the addition of these algorithms, the map on the COCO dataset reaches 43.5% and the speed reaches an amazing 65 FPS, which is a major milestone for the YOLO series.

Another outstanding algorithm for single-level target detection is SSD (separate shot multi Box detector) [11]. Liu et al. come up the SSD algorithm in 2016, which makes a great balance between detection speed and accuracy. SSD takes VGG as the basic skeleton and improves it by adding additional convolutional layers to obtain deeper feature information. In the last few layers of convolution, SSD uses anchoring methods to extract candidate frames for feature mapping at each scale, and determines the type and location of objects based on the candidate frames obtained by anchoring at different scales. Compared with Faster R-CNN, the anchors in SSD are scattered into different feature maps, and multi-scale features are used for multi-scale requirements. The mAP of SSD is 79.8% on PASCAL VOC 2007, 78.5% on PASCAL VOC 2012, and 28.8% on MS-COCO, achieving a good balance of detection speed and accuracy.

SSD constructs feature pyramids for detecting objects at different scales. The Conv4 layer with a feature stride of 8 is used to detect small objects, and the Conv8 layer with a feature stride of 64 is used to detect large objects. This allows small objects to not lose too much positional information in the shallow layers, while large objects can also be well localized and recognized in the deep layers. However, the small object features generated by shallow layers lack sufficient semantic information, resulting in poor small object detection performance. In response to the above problems, FSSD (Feature Fusion Separate Shot multi-box Detector, FSSD) adds a lightweight and efficient feature fusion module on the basis of traditional SSD [12]. FSSD first defines the framework of the feature fusion module and extracts the key factors that affect the performance of feature fusion. The FSSD feature fusion module firstly performs projection splicing of features of different scales in different layers, and then uses the batch normalization layer to normalize the feature values. Some down sampling blocks are then appended to generate new feature pyramids, which are fed back to the multi-box detector to produce the final detection-heavy results. Based on the above architecture, FSSD has a great improvement in performance compared to traditional SSD, but the loss in speed is small, especially for small objects. In addition, FSSD also outperforms many advanced VGGNet-based object detectors, and the feature fusion module also outperforms FPN in object detection. DSSD (Deconvolutional Separate Shot Detector) is one of the most famous improvements in SSD algorithms [13]. It replaces the VGG network in SSD with Resnet-101, which enhances the feature extraction ability; it also uses a deconvolution layer to add a lot of contextual information. An important enhancement of DSSD is its effectiveness in small object detection. However, the Resnet-101 network is too deep, making it slightly slower than SSD.

IV. EXPERIMENTS

A. Datasets

Data is one of the major elements of artificial intelligence development, where any research is inseparable from the support of data. Two of the most famous dataset are Pascal VOC and MS-COCO. The first one includes two different versions, VOC2007 and VOC2012. VOC2007 has 5,000 images and more than 12,000 labeled objects, while VOC2012 includes 11,000 images, more than 27,000 labeled objects and 20 classes of objects, adding semantic segmentation tasks and action recognition. MS-COCO is one of the most complex datasets, including 91 common objects found in nature, with more than 2000,000 numbers of images with every separate of them in 3.5 sorts including multiple perspectives.

B. Evaluation indicators

The accuracy, completion rate, recall, average accuracy, and mean accuracy (mAP) are the most common way to test the performance metrics in target detection. The accuracy rate represents the scales of the specimen of the classification pair in all samples, and is the ratio of the number of samples that correctly predict the class to the total number of samples. The precision means the rate of the number of correctly identified positive samples among the identified positive samples. Recall refers to the rate of the number of correctly identified positive samples in all positive samples in the test set. The average precision mAP is the main performance index used in the target detection algorithm, which refers to the area under the P-R curve. The P-R curve shows the trade-off between precision and recall on the classifier. The points on the P-R curve are the recall of the model and precision under a certain threshold. The P-R curve is calculated by changing the threshold from Generated by moving high to low. Mean Precision (mAP) is calculated by first calculating the mean precision (AP) for each class, and then calculating the mean of the APs.

C. Performance analysis

As presented in Figure 4 and Figure 5, we report various results on the VOC2007 and COCO dataset and several conclusions can be drawn as follow: (1) In the continuous improvement of the two-stage algorithm, its mAP in the VOC test set is getting higher and higher. mAP can reach 78% in the Faster R-CNN, which is about 19.5% improvement compared with the initial R-CNN. (2) Among the single-stage algorithms, the mAP of DSSD is the highest in the VOC test set, which can reach about 81.5%. In the COCO dataset, the mAP generally decreases compared to VOC, and the highest one is YOLOv3, whose mAP is about 43.5%. (3) With regard to accuracy, the two-stage algorithm is stronger in contrast to the single-stage algorithm. However, the single-stage algorithm is faster, with SSD (300) even reaching 46 fps, which is a qualitative improvement over the R-CNN algorithm with a few fps.

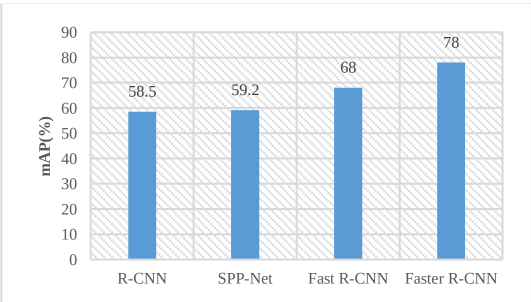

Fig. 4. Single-stage algorithms in performance test plots

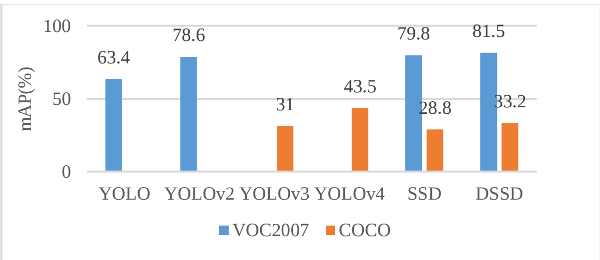

Fig. 5. Two-stage algorithms in performance test plots

V. DISCUSSION

Although the object detection problem has gone an comprehensive and rapid development process in the past research and has been widely used in many fields, the existing algorithms still have many fluxes, and its future research trends are briefly described below from several aspects.

(1) Lightweight real-time object detection. In the past, traget detection algorithms accounted for about tens to hundreds of Mb of memory ranging. While today's frontier fields, such as autonomous driving, smart cameras, face tracking recognition, etc., all impose the requirement of lightweight object detection algorithms with high accuracy and real-time. In recent researches, it is mainly achieved by light-weighting the backbone network or improving the performance of classification networks, such as MobileNetV2-SSDLite, Tiny-DSOD and ThunderNet .

(2) Weakly supervised detection problem. Most of the current mainstream algorithms are built on strongly supervised learning and rely on manually labeled data, which is less efficient. The access to large datasets in certain fields, such as medicine, has even become a limiting factor for their research. The main difficulties in its development are imprecise training labels, background noise interference, insufficient training samples and other problems.

(3) Small object detection: In some specific cases, the recognized object accounts for a small proportion of the whole, sometimes even only a few pixels, and the ability of the machine to recognize this type of object is generally low. Potential applications in this area are, for example, medical cell detection. At this stage, small object detection is generally achieved by coalescing high-resolution features and high-dimensional features in low-resolution images and by oversampling the images containing small objects .

(4) 3D object detection: In practical applications, 3D detection has more far-reaching significance for the progress of some fields, such as remote sensing mapping, military survey, biomedical detection, etc., because it contains more information about the length, width, and height of the object. At present, there are mainly fusion of image data and point cloud data, and only point cloud data as input in two directions. The representative algorithms are MMF, F-Point Net and GS3D, respectively. Although the development of this field has been relatively complete, but there are still exist some difficult to break through the bottleneck, such as perspective projection, light and other problems caused by noise, occlusion, etc., in the future there is still great potential for development.

VI. CONCLUSION

Following the development of technical design ideas, this paper summarizes the current representative algorithms in deep learning-based target detection research from two-stage target detection framework and single-stage target detection framework, and compares and analyzes common data sets and related algorithms in Experimental results on mainstream datasets. At the same time, focusing on the issues of lightweight real-time object detection, weakly supervised detection problem, small object detection and 3D object detection, we further analyze and prospect the future development direction of this research field.

REFERENCES

[1] XM Bao, SQ Wang. A review of deep learning-based target detection algorithms[J]. Sensors and Microsystems,2022,41(04):5-9.DOI:10.13873/J.1000-9787(2022)04-0005-05.

[2] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton . ImageNet Classification with Deep Convolutional Neural Networks public history 24 May 2017

[3] Girshick R , Donahue J , Darrell T , et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[J]. IEEE Computer Society, 2013.

[4] Gowda S R , Arava L , Zhan X , et al. 3D Nanoporous Nanowire Current Collectors for Thin Film Microbatteries[J]. Nano Letters, 2012, 12(3):1198. [3] Yanya Wang. A review of Two-Stage-based target detection algorithms[J]. Journal of the Hebei Academy of Sciences,2022,39(02):14-22.DOI:10.16191/j.cnki.hbkx.2022.02.003.

[5] YY Wang. A review of Two-Stage-based target detection algorithms [J]. Journal of the Hebei Academy of Sciences, 2022, 39(2):9.

[6] Bodla N , Singh B , Chellappa R , et al. Soft-NMS -- Improving Object Detection With One Line of Code[J]. 2017.

[7] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection. Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016. 779–788.

[8] Redmon J , Farhadi A . YOLO9000: Better, Faster, Stronger[C]// IEEE Conference on Computer Vision & Pattern Recognition. IEEE, 2017:6517-6525.

[9] Redmon J , Farhadi A . YOLOv3: An Incremental Improvement[J]. arXiv e-prints, 2018.

[10] Bochkovskiy A , Wang C Y , Liao H . YOLOv4: Optimal Speed and Accuracy of Object Detection[J]. 2020.

[11] W Liu , Anguelov D , Erhan D , et al. SSD: Single Shot MultiBox Detector[J]. 2015.

[12] L Wei , Dragomir A , Dumitru E , et al. SSD: Single Shot MultiBox Detector[C]// European Conference on Computer Vision. Springer, Cham, 2016.

[13] CY Fu , W Liu, Ranga A , et al. DSSD : Deconvolutional Single Shot Detector[J]. 2017.

2022 International Conference on Computer Graphics, and Information Visualization (CGIV 2022) Sanya, China

NOTIFICATION OF ACCEPTANCE

Dear Author(s):

On behalf of the 2022 International Conference on Computer Graphics, and Information Visualization (CGIV 2022), we're glad to inform you that your paper:

Paper ID: CGIV-155

Paper Title: Research Advanced in Deep Learning Object Detection

Author(s): Mingzhe Zhou

has been Accepted!

CGIV 2022 aims to bring researchers, engineers and students to the areas of Computer Graphics, Modeling and simulation, Information Visualization, and will provide an international forum for sharing the most advanced research results, experiences and original research contributions on related topics.

All accepted papers will be published by IEEE Conference Publishing Services, and will be submitted to EI Compendex, Thomson ISTP and CNKI databases.

2022 International Conference on Computer

Graphics, and Information Visualization

2022计算机图形与信息可视化国际会议

2022年06月30日

ICCGIV 2022 计算机图形与信息可视化国际会议 Organizing Committee 组织委员会

Engineering Village

IEEE

2022 International Conference on Computer Graphics, and Information Visualization (CGIV 2022) Sanya, China

Reviewer Comments for Article CGIV-155		
GENERAL ASSESSMENT	Originality of the Manuscript	☐Poor ☐Fair ☑Good ☐Excellent
	Significance of the Manuscript	☐Poor ☐Fair ☑Good ☐Excellent
	Organization of the Manuscript	☐Poor ☐Fair ☑Good ☐Excellent
	Completeness of the Manuscript	☐Poor ☐Fair ☑Good ☐Excellent
	Clarity of Presentation	☐Poor ☐Fair ☑Good ☐Excellent
TITLE	☑Accurately reflects content ☐Acceptable ☐Needs revision	
ABSTRACT	☑Clear and adequate ☐Confusing ☐Should be condensed ☐Does not present key results ☐Should be rewritten	
LANGUAGE	☑Grammatically correct ☐Needs revision ☐Unacceptable	
ILLUSTRATIONS	☑Good ☐Extra figures required ☐Quality inadequate ☐Needs revision	
TABLES	☑Good ☐Extra tables required ☐Should be rearranged ☐Needs revision	
ABBREVIATIONS, FORMULAS, UNITS	☑Confirm to acceptable standards ☐Need revision ☐Should be explained ☐A notation list is necessary	
REFERENCES	☑Appropriate ☐Need revision ☐Incorrect ☐Insufficient	
FINAL RECOMMENDATION	☑Acceptance ☐Revision ☐Rejection	

Comments:

Topic of this paper is important and related to this conference. The author's work is novel and suggestive, and the author makes investigation in the proposed method. There is no major modification opinion, only please pay more attention to English grammar and format your paper according to the template provided by conference homepage. I hope see you to present your paper work in CGIV 2022.

100020
北京市朝阳区朝外大街乙12号办公楼9层0-1008 北京众达德权知识产权代理有限公司
张晓冬(010-58790480)

发文日：

2022年05月31日

申请号或专利号：202220496401.0　　发文序号：2022052602007900

申请人或专利权人：周明哲

发明创造名称：一种自动化温控系统

授予实用新型专利权通知书

1．根据专利法第四十条及实施细则第五十四条的规定，上述实用新型申请经初步审查，没有发现驳回理由，现作出授予实用新型专利权的通知。

申请人收到本通知书后，还应当按照办理登记手续通知书的规定办理登记手续。

申请人办理登记手续后，国家知识产权局作出授予实用新型专利权的决定，颁发相应的专利证书，同时予以登记和公告。

期满未办理登记手续的，视为放弃取得专利权的权利。

法律、行政法规规定相应技术的实施应当办理批准、登记等手续的，应依照其规定办理。

2．授予专利权的实用新型申请是以
2022年3月7日提交的说明书；
2022年3月7日提交的说明书附图；
2022年5月19日提交的权利要求书；
2022年3月7日提交的说明书摘要；
2022年3月7日提交的摘要附图为基础的。

3．审查员依职权修改内容为：
注：在本通知书发出后收到的申请人主动修改的申请文件，不予考虑。

审 查 员：梁莉　　审查部门：实用新型审查部

联系电话：010-62089208

证书号第16780476号

实用新型专利证书

实用新型名称：一种自动化温控系统

发　明　人：周明哲

专　利　号：ZL 2022 2 0496401.0

专利申请日：2022年03月07日

专利权人：周明哲

地　　　址：030600 山西省晋中市榆次区张庆乡张庆村7-92户

授权公告日：2022年06月21日　　　　授权公告号：CN 216790363 U

国家知识产权局依照中华人民共和国专利法经过初步审查，决定授予专利权，颁发实用新型专利证书并在专利登记簿上予以登记。专利权自授权公告之日起生效。专利权期限为十年，自申请日起算。

专利证书记载专利权登记时的法律状况。专利权的转移、质押、无效、终止、恢复和专利权人的姓名或名称、国籍、地址变更等事项记载在专利登记簿上。

局长
申长雨

申长雨

2022年06月21日

其他事项参见续页

证书号第16780476号

专利权人应当依照专利法及其实施细则规定缴纳年费。本专利的年费应当在每年03月07日前缴纳。未按照规定缴纳年费的，专利权自应当缴纳年费期满之日起终止。

申请日时本专利记载的申请人、发明人信息如下:

申请人:

周明哲

发明人:

周明哲