

## *Retraction*

# **Retracted: Automatic Detection of Small- and Medium-Sized Targets in High-Resolution Images Based on Computer Vision and Deep Learning Energy**

### **Wireless Communications and Mobile Computing**

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

Copyright © 2023 Wireless Communications and Mobile Computing. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] Y. Lv, Z. Yin, and Z. Yu, "Automatic Detection of Small- and Medium-Sized Targets in High-Resolution Images Based on Computer Vision and Deep Learning Energy," *Wireless Communications and Mobile Computing*, vol. 2022, Article ID 7983127, 8 pages, 2022.

## Research Article

# Automatic Detection of Small- and Medium-Sized Targets in High-Resolution Images Based on Computer Vision and Deep Learning Energy

Yi Lv <sup>1,2</sup>, ZhengBo Yin <sup>3</sup>, and Zhezhou Yu <sup>3</sup>

<sup>1</sup>College of Computer Science and Technology, Jilin University, Changchun, Jilin 130012, China

<sup>2</sup>College of Computer Science and Technology, Changchun Normal University, Changchun, Jilin 130032, China

<sup>3</sup>College of Innovation and Entrepreneurship, Changchun University of Chinese Medicine, Changchun, Jilin 130117, China

Correspondence should be addressed to Zhezhou Yu; [yuzz@ccsfu.edu.cn](mailto:yuzz@ccsfu.edu.cn)

Received 6 April 2022; Revised 15 May 2022; Accepted 24 May 2022; Published 7 June 2022

Academic Editor: Aruna K K

Copyright © 2022 Yi Lv et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to solve the problem that it is difficult for traditional manual feature algorithms to deal with complex image features quickly and automatically in high-resolution images, an automatic detection method for small objects in high-resolution images based on computer vision and deep learning energy is proposed. Starting from the deep learning target detection system, this paper studies many problems in the system according to the characteristics of target spatial deformation, more small targets, easy confusion, and rough regional proposal in high-score remote sensing images. The results showed that a combination of spatial deformation strength with spatial deformation resistance and multiphase coupling were all proposed to be comparable with spatial deformation. The process of tracking international products in the system has been extended to study local channels. Pay attention to the operation of the machine, make full use of rich semantic data of local characteristics and the world of spatial points, which will help to improve the accuracy of the purpose of the chaw, and identify ways to create a shared database. The proposed method is evaluated on three attribute, classification joint datasets; ACCNN on the AC-AID dataset, the performance of ACCNN is suboptimal on AC-AID and is very close to the best method for DCA fusion. The average error rate of DCA fusion is 5.67%; on the AC-UCM dataset, the error rate of ACCNN is 2.11% lower than the suboptimal method DCA fusion, and the error rate of DCA fusion is 6.16%. On the AC-Sydney dataset, the error rate of ACCNN is 1.86% lower than the suboptimal method DCA fusion, and the error rate of DCA fusion is 5.99%, which obtains quite competitive results, experimenting with high-resolution images.

## 1. Introduction

There are many kinds of aerospace remote sensing platforms, such as domestic Fengyun series, Beidou series, and Gaofen series satellites, foreign Landsat series, and sentinel series satellites, as well as various aviation platforms such as UAV, high-altitude balloon, and airship. The Gaofen-5 satellite payload can obtain hyperspectral force sensing images with 30 m spatial resolution and 330 spectral channels; Gaofen 11 satellite can obtain optical remote sensing images with submeter spatial resolution. The combination of multiple satellites in the high-score series improves the time resolution of remote sensing images. The improvement of image quality helps peo-

ple to understand the real-time dynamic information of surface coverage in a more comprehensive and clear way, which makes it possible to continuously monitor the surface by means of remote sensing. However, with the improvement of data quality and the reduction of acquisition difficulty, the amount of remote sensing image data is larger and larger, and the content is more complex. The traditional processing method needs manual intervention, which is difficult to achieve rapid and automatic information processing [1].

Massive and high-quality remote sensing images become easy to obtain with the development of technology. This phenomenon greatly increases the difficulty of remote sensing image data processing. Therefore, more efficient and automatic

algorithms are needed to process these data quickly, and the deep learning algorithm is used as the research basis. Traditional image processing algorithms are based on manual design features, such as scale invariant feature transform (SIFT) features, histogram of gradient (HOG) features, and bag of visual words (BOVW) features based on visual features. Although these algorithms can realize the processing of high-score remote sensing images to a certain extent, these algorithms have artificially designed features, have a certain bias in the selection of image features, and cannot extract the features most suitable for the current image, so the detection accuracy is poor. Moreover, with the implementation of major projects of China's high-resolution earth observation system and the launch of multiple high-resolution remote sensing satellites, the resolution of China's space-based forced sensing image has been accurate from 2.1 meters to 0.65 meters or even higher. The spatial details in the high-resolution forced sensing image are more and more sufficient, and the image content is more and more complex. The traditional image processing algorithm based on manual features is difficult to effectively deal with this complex image [2].

Deep learning is widely used in driverless, face recognition, machine translation, target recognition, emotion recognition, and so on. The reason why the deep learning algorithm is efficient is that it has the ability of self-learning from a large amount of data, avoids the interference of human factors on feature selection, and can get the expression closer to the best feature. For the complicated detail information in high-score remote sensing images, deep learning algorithm can obtain certain adaptability through a large number of sample learning, which is much better than the performance of traditional algorithms. Therefore, how to effectively introduce deep learning into the field of remote sensing image processing is an important topic worthy of research [3].

## 2. Literature Review

Toldinas believes that deep learning algorithms are generally supervised and need a large dataset with a large number of samples to train the whole model. The samples in the high-score remote sensing target detection dataset should have two labels, the real target location in the image and the corresponding real target category at each location [4]. Nasif research since the shooting angle of high-score remote sensing image is from high altitude, the angle of the target in the image is arbitrary, and the target in the image has rotation change. In addition, the size of the target is also affected by the height of the platform, and the target in the image has scaling changes. In addition, the tilt of the camera may cause image distortion, etc. There are a series of other spatial deformations in high-resolution remote sensing images. Therefore, in order to introduce the deep learning algorithm into the target detection field of high-score remote sensing images, it is necessary to build a relatively robust feature expression model for these spatial variation factors [5]. Mhango's work aiming at the problem of small target detection, which is one of the difficulties of target detection, the analysis shows that part of the difficulty of detection is that slight positioning deviation will have a fatal impact on small target detection. For small targets, it is assumed that

there is a positioning deviation of several pixels between the predicted boundary box and the real target box [6]. From the perspective of multiscale feature fusion, Ma improved FasterRCNN and proposed feature pyramid networks (FPN). Through the top-down integration of network deep features and shallow features layer by layer, the problem of lack of deep semantic features in small target detection is solved to a certain extent [7]. Liu added a fusion module TDM (top down modulation) on the side connection of the basic detection network. In the structure of TDM, the information fusion between network layers is no longer the same as FPN. First, the features of different scales are scaled to the same dimension, and then, the features are fused in the way of element superposition. Instead, the required features are screened out through convolution operation, and then, the features of different scales are combined by splicing, so as to avoid the negative impact of feature superposition on information [8]. Gong proposed DSSD (deconvolutional single-shot detector) network structure, in which D is deconvolution module. DSSD inputs the features generated by SSD (single shot detector) model into the deconvolution module and then outputs the modified feature map pyramid to form an hourglass structure composed of features of different scales. When fusing the features between adjacent layers, it is also necessary to scale the features of different scales to the same dimension and then fuse the features by multiplying the corresponding elements [9]. The above research work belongs to multiscale feature method. In addition, multiscale training method is also a widely used scheme to solve the problem of multiscale target detection. Recently, Liang proposed snip and sniper algorithm, which is an improvement on the traditional multiscale training method. Sniper is the accelerated version of snip. The research work of snip and sniper has effectively proved that even with relatively sufficient data, CNN (revolutionary neural network) is still difficult to use objects of all scales, but can only act on targets of a certain range of scales. Therefore, in the multiscale training, only the target area with the scale within the detection range is gradient retransmitted, which greatly reduces the computational overhead of the multiscale training method. Other methods focus on the impact of the design of a priori frame on detection performance [10]. Wang found in the research work of S3FD that the scale mismatch between the candidate frame and the small target and the scale mismatch between the receptive field and the small target make it difficult for the small target to be detected. Therefore, they proposed candidate box matching strategy to solve the above two scale mismatch problems [11]. Subsequently, Jiang proposed a strategy of densely generating candidate boxes in the research work of Facebook, that is, using candidate boxes with the same density to sample large/medium/small targets with different scales. This strategy of densely sampling small targets can effectively improve the recall rate of small object detection [12]. In pyramid boxing training, Shin and Yu expand microtraining models and improve the results of small-scale programs to maximize the diversity of small-scale instructional materials. The preframing model is usually focused on improving the results of a small target to visualize it correctly [13]. The research of Zang also shows the importance of accurate positioning. Some research teams have tried to apply superresolution technology to small

target detection and made pioneering work. Using the idea of superresolution reconstruction, Li and others proposed to transform the shallow features of small targets into corresponding high-resolution representations to improve the semantic feature representation of small targets [14]. Ahmed and others tried to apply the research of generative adversarial network (GAN) in superresolution reconstruction to small target detection and proposed the network structure of Sod-Mtgan [15]. Li has studied the first-order product search algorithm that directly creates product groups and parts integrated with a single finding. The whole process is easily compared to a two-stage target detection algorithm, and the representative algorithms include the Yolo series algorithm and the SSD series algorithm [16]. The operation of the StageG marker search algorithm is shown in Figure 1.

Based on the current research, based on a deep learning target detection system, this paper studies many problems in the system according to the characteristics of target spatial deformation, more small targets, easy confusion, and rough regional proposal in high-score remote sensing images. The results are as follows: convolution features combined with multilevel full connection features robust to spatial deformation are proposed to enhance the robustness of feature expression to spatial deformation. Using the method of synthesis, a joint dataset of three attributes and classifications is established for the training of multitask learning network.

### 3. Research Methods

**3.1. Hierarchical Robust Feature Extraction.** Classical RSI target detection methods are almost based on sliding window search for possible targets. In the sliding window search algorithm, the initial target proposal is generated by selecting from regions with different positions and scales. This kind of method uses brute force algorithm, which is time-consuming, laborious, and large amount of calculation. In recent years, in order to avoid exhaustive search, this paper uses the region of interest (ROI) generated by selective search algorithm.

In our HRCNN system, Alex net1491 is used as the basis for unpacking features. Then, three rotating layers with different spatial scales in the network were selected to create a map of the semantic energy of the hierarchical space. To avoid being affected by spatial deformation such as rotation and scaling, we developed rotational and scaling system (RSRE) in this framework. Finally, product research is done using class symbols and reworkers. It should be noted that the application HSS and RSRE models can be trained with the core network. Such a learning process makes the network viable for practical use [17].

The CNN model proposed is used to create a review team and site regression, respectively, to exclude results after the analysis and prediction sites. This document uses a vector support system (SVM) as the host for the class. Existing ROI localization can be improved by limiting box regression. This phrase uses the method of spinal therapy to regenerate.

**3.1.1. Introduction to SVM.**  $x_i$  represents the feature vector of the  $i$ -th sample (i.e., regional proposal), and  $y_i$  represents the corresponding category label. In the feature space, find

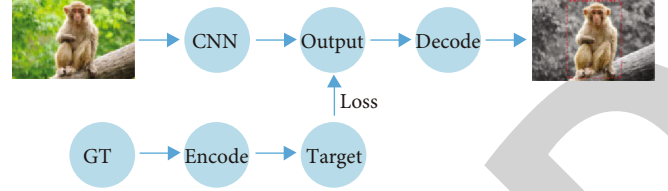


FIGURE 1: One stage target detection algorithm.

the critical samples to form the support vector, and construct the following linear discrimination plane to complete the task of category recognition, as shown in

$$f(\mathbf{w}, b) = \mathbf{w}^T x_i + b. \quad (1)$$

Matrix  $w$  and vector  $b$  are the parameters to be learned in the process of SVM training. When looking for support vectors, the optimization objective function is expressed as

$$\min_{w, b} \frac{1}{2} \|\mathbf{w}\|_2^2, \quad (2)$$

$$y_i (\mathbf{w}^T x_i + b) \geq 1, i = 1, 2, \dots, n.$$

However, in most cases, the input model cannot be segmented entirely using only the spatial linear function of the function. Therefore, the vector must be plotted in a nonlinear space. A kernel replacement function has been introduced, which provides better insights [18]. After production, the final Lagrange function is shown in

$$L(w, b, \Lambda) = \frac{1}{2} w^T w + \sum_{i=1}^l \alpha_i - \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j \kappa(x_i, x_j), \quad (3)$$

where  $\kappa(x_i, x_j)$  refers to the kernel function of the current convolution and  $\Lambda = (\alpha_1, \alpha_2, \dots, \alpha_l)$  is a nonnegative Lagrange multiplier. Finally, the concise expression of the optimization function is shown in

$$\min_w \max_{\alpha_i \geq 0} L(\mathbf{w}, b, \Lambda). \quad (4)$$

Existing ROI localization can be improved by limiting box regression. This phrase uses the method of spinal therapy to regenerate.  $X$  is the input function matrix,  $Y$  is the difference between the constraint box and the actual constraint box, and the return of the spine as the square error is shown at

$$\|\mathbf{X}\theta - \mathbf{Y}\|_2^2, \quad (5)$$

where  $X$  is a  $N * L$  matrix and  $N$  and  $L$  are the number of samples and characteristic dimensions, respectively.  $Y$  is an  $N * 4$  matrix; each row of which is a coordinate regression vector, which is the coordinate difference between the real target box and the proposed target box. Each element of the vector corresponds to the abscissa of the center point of the target box, the ordinate of the center point, and the width and height

of the bounding box.  $\theta$  is a parameter matrix of parameter matrix  $L * 4$  of size, which is used to derive the regression matrix  $Y$  from the input feature  $X$  [19]. The solution method of  $\theta$  is

$$\theta = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}. \quad (6)$$

However, this problem is ill posed because  $X$  is not rank. Therefore, a regular term formula (7) is added:

$$\|\mathbf{X}\theta - \mathbf{Y}\|_2^2 + \alpha \|\theta\|^2. \quad (7)$$

The solver rule becomes

$$\theta(\alpha) = (\mathbf{X}^T \mathbf{X} + \alpha \mathbf{I})^{-1} \mathbf{X}^T \mathbf{Y}, \quad (8)$$

where  $\mathbf{I}$  is an element matrix with size  $L$ , which is a constant with small value, which is set to 0.001 here. So far, the above ill-posed problem has been transformed into a well-posed problem. The target detection method and process of deep learning features are shown in Figure 2.

**3.2. Target Positioning Based on Gated Axis Clustering Positioning Network.** It is often difficult to capture small objects in a wide range of search equipment with high resolution for remote controls. This is because the effect of the difference on the location of the smaller target is larger than that of the larger target. This is because the cross section of the smaller body is much smaller than the cross section of the larger body and much smaller than the IOU when the cross section of the approximate and correct spheres is reduced by the same rate. To solve this problem, we propose a new regional model to improve the accuracy in the area of small objects. The model has two sections. First, the process of determining global characteristics is planned in order to focus on the channel to study local characteristics. This process makes full use of rich semantic data of global characteristics and spatial meanings of local characteristics. The data selected in this way will be more useful in identifying small targets. Second, the axial agglomeration prediction (ACP) method is used to prepare rotation diagrams in different directions to avoid interference of different joints and to improve the exposure accuracy and then using the regression technique using models to illustrate the tools learned [20].

**3.3. Target Recognition Based on Attribute Cooperative Convolution Neural Network.** Classification of remote sensing images is one of the most important aspects of remote sensing image processing and is part of the goal of detecting remote sensing images. RSIs are known to be very difficult due to the diversity of their content, and it is difficult to distinguish between different events with similar descriptive concepts, such as deserts and the barren region. The Attribute Collaborative Convolutional Neural Network (ACCNN), which uses the material as a supplement, claims that it is difficult to classify negative structures. The Convolutional Neural Network feature first uses CNN and then distribution centers to determine the RSI status. Second, an attribute field is intended to

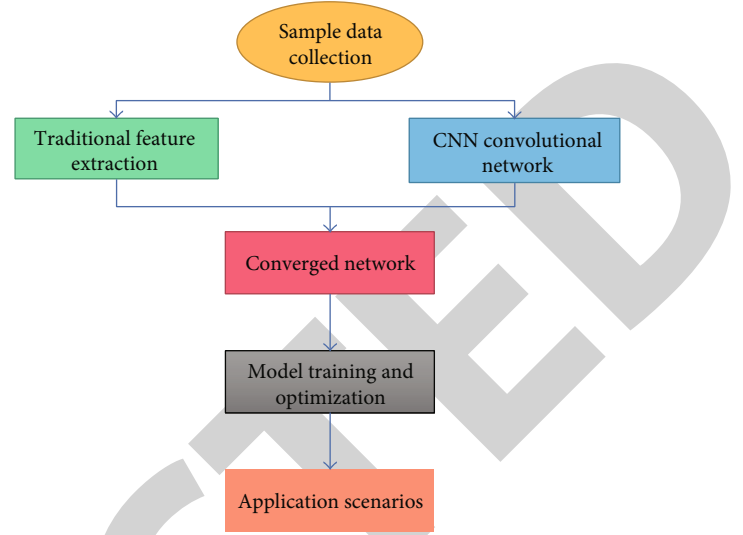


FIGURE 2: Target detection method and process of deep learning features.

predict image data. Through collaborative training, networks are able to be aware of current events. Because the input attribute branch and the subdirectory subdivide a layer to unpack the attribute, the subdivision branch will access the additional attribute data. Finally, the relationship between branches of distribution and behavior is learned through communication, which strengthens the exchange of information between the two positions. Our data sharing system (AC-AID, AC-UCM, and AC-Sydney) is designed to capture data behaviors [21].

## 4. Result Analysis

According to the ms-coco evaluation protocol, this paper uses the average accuracy map to evaluate the effectiveness of the algorithm. The real and false indicators are the basis of computer vision. From TP, FP, and FN, the precision  $P$  and recall  $R$  can be calculated as

$$P = \frac{TP}{TP + FP}, \quad (9)$$

$$R = \frac{TP}{TP + FN}. \quad (10)$$

After the test results of a certain category are obtained, they are sorted according to the confidence, and the recall and precision obtained by taking each confidence as the division threshold are calculated. The corresponding maximum precision rate is obtained according to different recall rates, and then, the average accuracy AP of a certain class is obtained as

$$AP = \sum_{i=1}^n P_i (R_i - R_{i-1}). \quad (11)$$

Find the average value of all categories of AP to obtain the average accuracy map. At the same time, the map in this

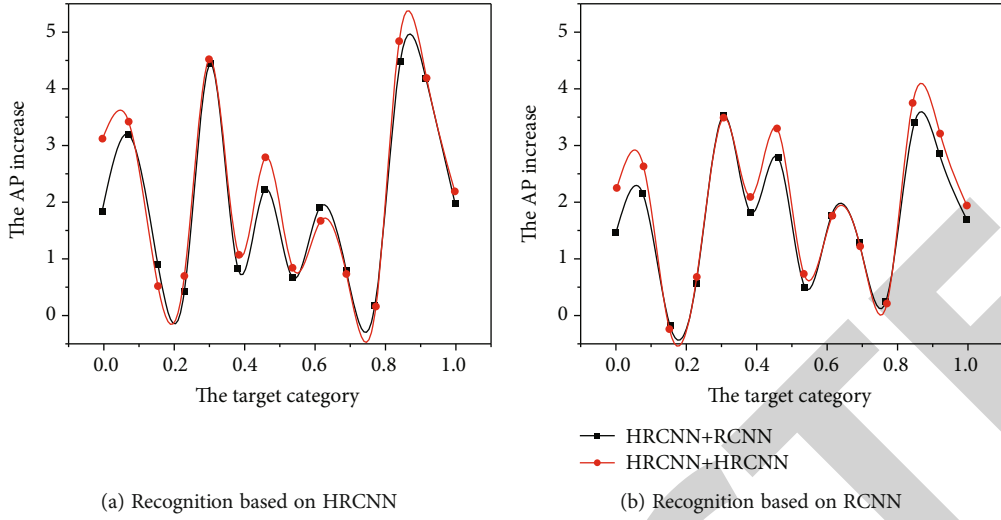


FIGURE 3: Comparison of regression strategies under the same identification strategy.

paper is the average value taken among multiple IOU thresholds, using 10 IOU thresholds with an interval of 0.05 from 0.50 to 0.95. The traditional method only calculates the index with a single IOU threshold of 0.50, and the prediction area is considered correct only when  $IOU > 0.50$ . Compared with traditional methods, the method of averaging multiple IOU thresholds can more comprehensively test the effectiveness of the detection algorithm. In addition, this paper also takes map and map 75 as a set of auxiliary reference data and the average maximum recall rate as another set of auxiliary reference data. Ar100 represents the maximum recall rate obtained by giving 100 detection results in each picture [22].

**4.1. HRCNN Regression Effect Detection.** When we look at different regression strategies when the properties used for recognition are similar, we see a difference: in general, the use of regression improves the AP, while the regression of the HRCNN function always improves the AP rather than the specific regression of the RCNN. For graphical sites, RCNN regression improved RCNN accuracy by 1.71% and HRCNN by 2.02%; HRCNN regression improved RCNN accuracy by 1.94% and HRCNN by 2.22%. To clearly explain this improvement, we compare different values of returns from the same angle, as shown in Figure 3. In the figure, the  $x$ -axis is the target and the  $y$ -axis is the AP value from AP, without getting results. It is clear that the side effects of HRCNN are stronger even if RCNN recognition or HRCNN recognition is used.

**4.2. Comparison of Positioning Model Accuracy in HRCNN.** The comparative experimental detection results of GACL net in HRCNN and NWPU VHR-I0 show that our model has the best mapping on both datasets. In addition, based on fast RCNN backbone network, GACL net improves the accuracy of HRRSD and NWPU VHR-I0 by 1.5% and 2.5%, respectively. Based on the fast RCNN backbone network, GACL net adopts the regional proposal network (RPN) as the

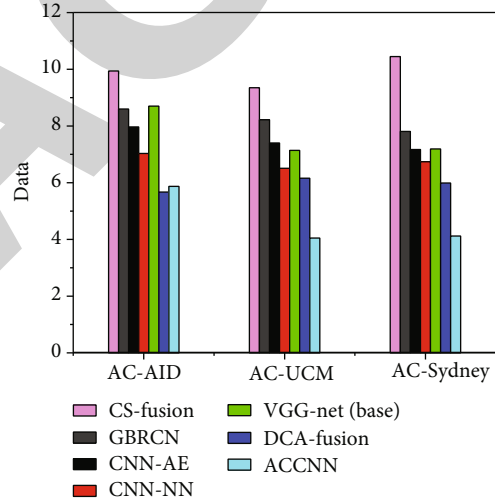


FIGURE 4: Classification error rate of different methods on three datasets.

regional proposal, which has increased by 0.6% and 0.4%, respectively, on HRRSD and NWPU VHR-I0 [23].

**4.3. Comparison of Error Rate of Attribute Cooperative Convolutional Neural Network (ACCNN) in Dataset.** The classification error rate of ACCNN in three datasets is shown in Figure 4, and the attribute error rate on each dataset is shown in Figures 5–7.

The results of the AC-AID test include the results of the behavioral analysis and the distribution results. The results of the behavior attribute are shown in Figure 5, where “Attribute ID” represents the different objects. As shown in Figure 5, the error for this product ranges from 0.65% to 33.05%. The average error is 5.06%. The results showed that the area of performing arts worked well. Compared with the main VG-net, the average error rate of VG-net decreased by 2.87%. The performance of ACCNN in AC-AID is not good

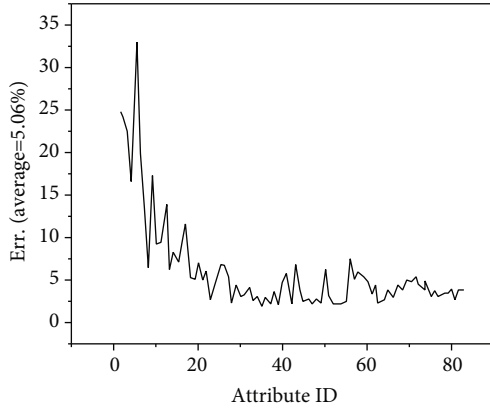


FIGURE 5: Attribute error rate of ACCNN on AC-AID dataset.

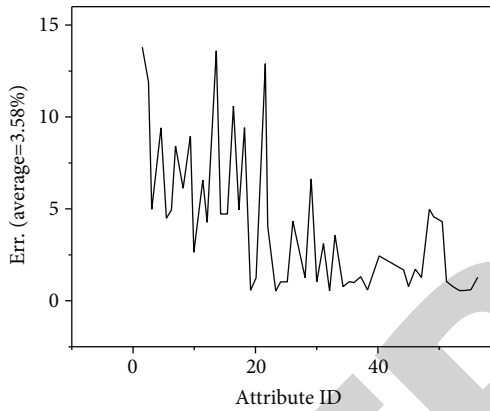


FIGURE 6: Attribute error rate of ACCNN on AC-UCM dataset.

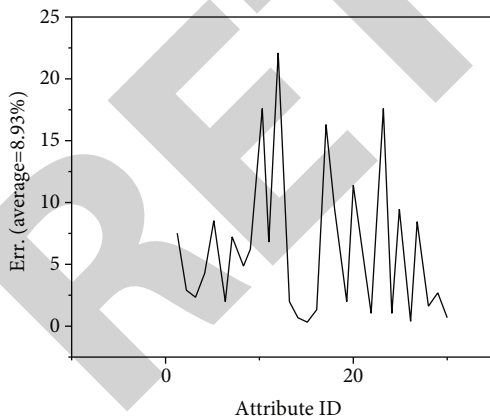


FIGURE 7: Attribute error rate of ACCNN on AC-Sydney dataset.

and is close to the best method of DCA melting. The average DCA smelting error rate is 5.67%. Thus, this method demonstrates the advantages of AC-AID.

AC-UCM-based testing has included classification and analysis. The results of the behavioral hypothesis are shown in Figure 6. The error for the various groups varied from 0.00% to 13.81%, and the mean error for these characteris-

tics was 3.58%. The results showed that the area of performing arts worked well. Compared with 7.14% error of base, the error of this algorithm is reduced by 3.09% and reduced by 4.05%. This result is the result of the optimal distribution of AC-UCM. In addition, the ACCNN melting error is 2.11% lower than the negative of DCA melting, and the DCA melting error is 6.16%. Therefore, the method mentioned in this document gets the SOTA accuracy in the AC-AID database [24].

The performance results of AC-Sydney are shown in Figure 7. Equipment errors increased from 0.00% to 22.22%, with an average error of 8.93%. The results show that the academic sectors have worked well. The error of this algorithm decreased from 3.06% to 4.13% compared to the standard algorithm error 7.19. This benefit is the best benefit for AC-Sydney. In addition, the ACCNN melting error is 1.86% less than the DCA melting error and the DCA melting error is 5.99%. Therefore, there is full SOTA certification of the AC-Sydney database [25]. The test results show that the algorithm works better than the comparison algorithm. The reason is that behavioral branching and relationship fragmentation have been introduced and weight-sharing mechanisms have been developed [26].

## 5. Conclusion

High-score remote sensing image has the characteristics of large imaging range, high spatial resolution, and complex image content. The goal of target detection task based on high-score remote sensing image is to locate several categories of targets of interest from a large number of high-score remote sensing images and analyze and predict the target categories at each location. Through the investigation of the existing algorithms of the research task, it is found that the traditional manual feature method cannot deal with the complex and changeable high-score remote sensing images well: the deep learning neural network algorithm rising in recent years can adaptively learn the hidden features in the data, free from the interference of human factors, and can better express the changeable target features in high-score remote sensing images. The following results were achieved:

- (1) In the research of finding feature expression methods that can deal with spatial deformation such as rotation and scaling, aiming at the phenomenon that the traditional deep learning algorithm adopts a single global feature, this paper proposes to combine multilevel convolution features to enhance spatial information to help achieve more accurate target location, that is, target location in the detection system. Aiming at the spatial deformation problems such as the scaling and rotation of targets in the image, a convolution feature combined with multilevel full connection features robust to spatial deformation is proposed to enhance the robustness of feature expression to spatial deformation
- (2) In the research of finding a target location method that can deal with the interproblem of small target detection, aiming at the problem that the detection

accuracy of small target is greatly affected by the location accuracy, a method of projecting the convolution feature map to the horizontal and vertical spatial directions is proposed to avoid the interference of coordinate prediction in the horizontal and vertical orthogonal directions from receiving additional dimension information, so as to effectively improve the location accuracy

- (3) In the research of finding target recognition methods that can deal with the problems of complex and confusing content of high-score remote sensing images, aiming at the problems of complex images and easy confusion between nonsimilar targets and backgrounds, this paper proposes to use attribute learning task to provide additional auxiliary information for classification (recognition) task and fully share convolution layer between two task branches to enhance the discrimination ability of classifier. In addition, a combined dataset of three attributes and classifications is established by using the method of synthesis for the training of multitask learning network. The proposed method is evaluated on three attribute and classification joint datasets, and quite competitive results are obtained

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

- [1] O. M. Lawal, "Yolomuskmelon: quest for fruit detection speed and accuracy using deep," *Learning Access*, vol. 9, pp. 15221–15227, 2021.
- [2] B. Ramalingam, A. A. Hayat, M. R. Elara, B. F. Gómez, and S. Subramanian, "Deep learning based pavement inspection using self-reconfigurable robot," *Sensors*, vol. 21, no. 8, p. 2595, 2021.
- [3] Z. Jiang, J. Zhou, and H. Huang, "Relationship between manifold smoothness and adversarial vulnerability in deep learning with local errors," *Chinese Physics B*, vol. 30, no. 4, article 048702, 2021.
- [4] J. Toldinas, A. Venkauskas, R. Damaevius, A. Grigaliūnas, and E. Baranauskas, "A novel approach for network intrusion detection using multistage deep learning image recognition," *Electronics*, vol. 10, no. 15, p. 1854, 2021.
- [5] A. Nasif, Z. A. Othman, and N. S. Sani, "The deep learning solutions on lossless compression methods for alleviating data load on iot nodes in smart cities," *Sensors*, vol. 21, no. 12, p. 4223, 2021.
- [6] J. K. Mhango, E. W. Harris, R. Green, and J. M. Monaghan, "Mapping potato plant density variation using aerial imagery and deep learning techniques for precision agriculture," *Remote Sensing*, vol. 13, no. 14, p. 2705, 2021.
- [7] L. Ma, Y. Li, J. Li, J. M. Junior, and M. A. Chapman, "Boundarynet: extraction and completion of road boundaries with deep learning using mobile laser scanning point clouds and satellite imagery," *IEEE Transactions on Intelligent Transportation Systems*, vol. PP (99), pp. 1–17, 2021.
- [8] S. Z. Liu, R. S. Sinha, and S. H. Hwang, "Clustering-based noise elimination scheme for data pre-processing for deep learning classifier in fingerprint indoor positioning system," *Sensors*, vol. 21, no. 13, p. 4349, 2021.
- [9] B. Gong, R. Cai, Z. Cai, Y. Ding, and M. Peng, "Selection of acoustic modeling unit for Tibetan speech recognition based on deep learning," *MATEC Web of Conferences*, vol. 336, no. 6, p. 06014, 2021.
- [10] F. Liang, C. Li, and X. Fu, "Evaluation of the effectiveness of artificial intelligence chest ct lung nodule detection based on deep learning," *Journal of Healthcare Engineering*, vol. 2021, 10 pages, 2021.
- [11] J. Wang and P. Srikantha, "Stealthy black-box attacks on deep learning non-intrusive load monitoring models," *IEEE Transactions on Smart Grid*, vol. 12, no. 4, pp. 3479–3492, 2021.
- [12] L. Jiang, X. Wang, W. Li, L. Wang, and L. Jia, "Hybrid multi-task multi-information fusion deep learning for household short-term load forecasting," *IEEE Transactions on Smart Grid*, vol. 12, no. 6, pp. 5362–5372, 2021.
- [13] Y. H. Shin and D. C. Lee, "True orthoimage generation using airborne lidar data with generative adversarial network-based deep learning model," *Journal of Sensors*, vol. 2021, 25 pages, 2021.
- [14] S. Zang, L. Mu, L. Xian, and W. Zhang, "Semi-supervised deep learning for lunar crater detection using ce-2 dom," *Remote Sensing*, vol. 13, no. 14, p. 2819, 2021.
- [15] S. Ahmed, U. Kamal, and M. K. Hasan, "Dfr-tds: a deep learning based framework for robust traffic sign detection under challenging weather conditions," *IEEE Transactions on Intelligent Transportation Systems*, vol. PP (99), pp. 1–13, 2021.
- [16] J. Li, J. Su, C. Xia, M. Ma, and Y. Tian, "Salient object detection with purificatory mechanism and structural similarity loss," *IEEE Transactions on Image Processing*, vol. 30, pp. 6855–6868, 2021.
- [17] L. I. Yundong, D. O. Han, L. I. Hongguang, X. Zhang, B. Zhang, and X. I. Zhifeng, "Multi-block ssd based on small object detection for uav railway scene surveillance," *Chinese Journal of Aeronautics*, vol. 33, no. 171, pp. 179–187, 2020.
- [18] B. Gu, R. Ge, Y. Chen, L. Luo, and G. Coatrieux, "Automatic and robust object detection in x-ray baggage inspection using deep convolutional neural networks," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 10, pp. 10248–10257, 2021.
- [19] S. Yuan, K. Wang, Y. Shan, and J. Yang, "Multi-scale object detection method based on multi-branch parallel dilated convolution," *Jisuanji Fuzhu Sheji Yu Tuxingxue Xuebao/Journal of Computer-Aided Design and Computer Graphics*, vol. 33, no. 6, pp. 864–872, 2021.
- [20] Z. Huang, H. Chen, B. Liu, and Z. Wang, "Semantic-guided attention refinement network for salient object detection in optical remote sensing images," *Remote Sensing*, vol. 13, no. 11, p. 2163, 2021.
- [21] N. Huang, Y. Yang, D. Zhang, Q. Zhang, and J. Han, "Employing bilinear fusion and saliency prior information for rgb-d salient object detection," *IEEE Transactions on Multimedia*, vol. (99), p. 1, 2021.



- [22] J. S. Chou and C. H. Liu, "Automated sensing system for real-time recognition of trucks in river dredging areas using computer vision and convolutional deep learning," *Sensors*, vol. 21, no. 2, p. 555, 2021.
- [23] A. Cummaudo, R. Vasa, J. Grundy, and M. Abdelrazek, "Requirements of api documentation: a case study into computer vision services," *IEEE Transactions on Software Engineering*, vol. PP (99), pp. 1–1, 2020.
- [24] Z. Chen, Z. Yan, H. Jiang, Z. Que, and Z. Xu, "Detecting coal pulverizing system anomaly using a gated recurrent unit and clustering," *Sensors*, vol. 20, no. 11, p. 3271, 2020.
- [25] M. Pan, A. Liu, Y. Yu, P. Wang, and H. Zhu, "Radar hrrp target recognition model based on a stacked cnn-bi-rnn with attention mechanism," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 99, pp. 1–14, 2021.
- [26] M. Yang, W. Xu, Z. Sun, H. Wu, and L. Li, "Mid-wave infrared polarization imaging system for detecting moving scene," *Optics Letters*, vol. 45, no. 20, pp. 5884–5887, 2020.