EAI Endorsed Transactions
on Scalable Information Systems
Research Article EAI.EU

Region proposal network based on context information feature fusion for vehicle detection

Zengyong Xu[1,*]

[1]School of Automotive Studies, Henan College of Transportation, Zhengzhou 451100 China

Abstract

By using the traditional methods, the feature information extracted from vehicle target detection is insufficient, which leads to the low accuracy in identifying small target vehicles or blocked targets. Therefore, we propose a region proposal network (RPN) based on context information feature fusion for vehicle detection. RPN obtains feature vectors of fixed length as vehicle target features. Context information fusion network obtains the corresponding context information features on the feature maps of different layers. Finally, the two features are fused. In addition, in order to solve the problem of data imbalance, experiments on PASCAL VOC2007 and PASCAL VOC2012 data sets with difficult sample training show that the proposed method has significantly improved the mean average accuracy (mAP) compared with other methods.

Keywords: RPN, vehicle detection, context information fusion.

Received on 18 January 2022, accepted on 24 January 2022, published on 27 January 2022

doi: 10.4108/eai.27-1-2022.173161

*Corresponding author. Email: zxcvfdsa2224@foxmail.com

1. Introduction

Object detection refers to the recognition of the existence of known objects in a given image and the use of rectangular boxes to mark the location of objects. Convolutional Neural Network (CNN) is one of the classic object detection algorithms [1-3]. At present, target detection algorithms based on CNN are mainly divided into two categories [4]:

1) two-step method based on candidate box [5], which requires selective search algorithm to select;

2) Regression based one-step method [6], which uses CNN network to directly predict the category and position of the target. The method of obtaining candidate boxes is not strong in real time for target detection, but has high accuracy. The method based on regression does not need to obtain candidate boxes, and has strong real-time performance, but low detection accuracy.

In real life, the changes of image imaging conditions and environment significantly affect the appearance of objects, resulting in differences between objects. Even objects of the same category may affect the accuracy of target detection due to factors such as imaging time, location, weather conditions, camera, background, light and viewing distance. Context information can be understood as contextual information, which usually refers to any and all information that may affect the perception of the scene and the objects in it. Studies show that context information can improve the accuracy of target detection algorithm [7-10].

The current context information is mainly divided into two parts: 1) context-level context information around the target object, and 2) context information about the relationship between objects. Torralba [11] showed that

the recognition accuracy could be improved by using object shape and context information, and the context features were divided into three categories: semantic context (probability), spatial context (location) and scale context (size). Divvala et al. [12] evaluated more contextual information, such as local pixel context, geometric context, geographical environment, etc. Liu et al. [13] proposed inside-outside Net (ION) and external network. Externally, the sliced-loop neural network was used to integrate contextual information around the Region of Interest (ROI) [14,15], and internally, ROI features of different layers were extracted for fusion. Liu et al. [16] proposed Structure InferenceNet (SIN), which used the relationship between context information at the scene level and instances for target detection and recognition. Shrivastava et al. [17] proposed a Contextual Priming and Feedback network (CPF) to provide top-down Contextual information Feedback. Mei et al. [18] proposed Attention to Context CNN (ACCNN), which utilized local multi-scale CNN features and generates global features through Long Short-term Memory (LSTM) for target recognition. However, the extracted global information only improved the detection accuracy by 0.6%, and all the above algorithms extracted only the context information features of a certain layer, resulting in insufficient feature information. In addition, Wang et al. [19] proposed the combination of front-end network and multi-layer context module, but focused on the front-end network, and only applied to image segmentation.

In the training process of target detection, the target area in the whole image is smaller than the background area, that is, the positive sample is less than the negative sample. If the classifier is trained directly with these unbalanced data, the classifier may tend to classify all samples as negative samples. Shrivastava et al. [20] carried out secondary classification of samples that were difficult to correctly classify, and proposed Online Hard Example Mining (OHEM). All candidate frames were back-propagated, sorted according to Loss, and some candidate frames with large Loss were retrained as difficult samples to effectively improve the accuracy of target detection.

2. Proposed vehicle detection network

2.1. RPN and anchors

Faster R-CNN is a two-stage target detection method, in which an input image is sent to the convolutional layer for a convolution operation to obtain a feature image, and then the feature image is sent to a Region Proposal Network (RPN) to obtain a target candidate box. Finally, candidate boxes and feature maps are sent to pooling layer and full connection layer for classification.

We define three scales (8,16,32) and three aspect ratios (1:1,1:2,2:1) anchors in Faster R-CNN that represent nine possible sizes for each slide window in the original area of the RPN network when using the sliding window strategy for feature maps. Each point on the feature map generates 9 anchor points. In this paper, we define 6 kinds of anchors with different scales (1,2,4,8,16,32) and 3 kinds of height-width ratios (1:1,1:2,2:1), and obtain 18 kinds of anchors. The addition of small scale anchor points is conducive to the detection of small vehicle targets, as shown in figure 1.

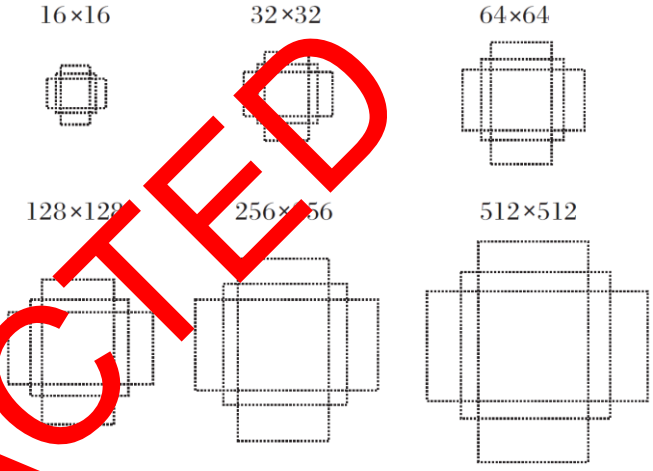

Figure 1. 18 anchor point sizes

2.2. Network structure

The proposed structure is shown in figure 2. Taking Faster R-CNN as the benchmark network, VGG16 was pretrained on ImageNet to initialize the network. For each input image, feature maps are generated through different convolutional layers and pooling layers. The feature graph was convolved with Conv5 to generate ROI features through RPN [21,22]. In conv4 and Conv5, the corresponding context information features of the same size and different scales were extracted for each ROI, and then the two connected context information features were input into the convolutional layer of 1×1 for normalization. Finally, the normalized context information features were fused with ROI features to generate fixed-length feature descriptors for each 512×7×7 Proposal Region. Two Fully Connected Layers (FC) process each descriptor and produces two outputs: a K-object class prediction, an adjustment to the bounding box of the proposed region.

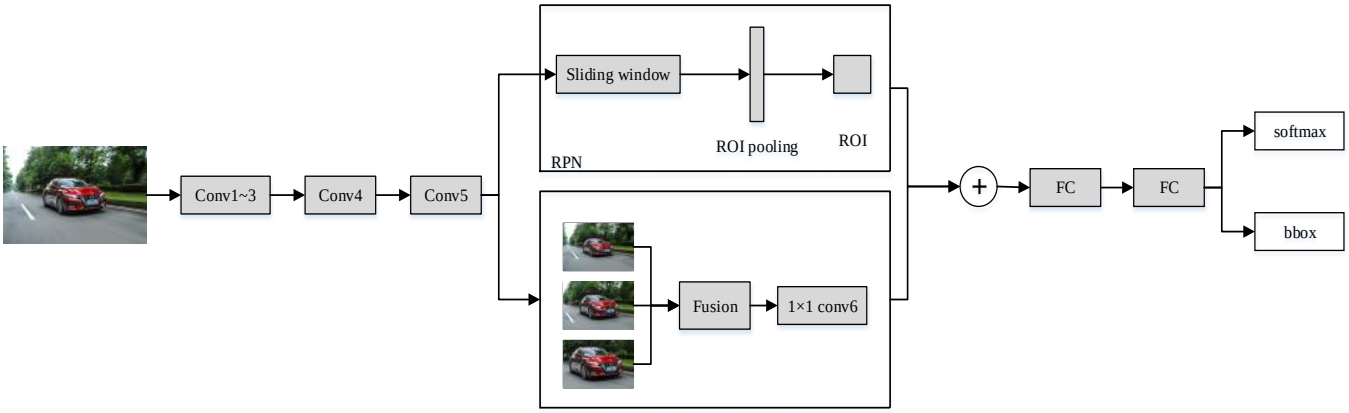

Figure 2. Proposed network.

2.3. Context information extraction

Context information features play an important role in target detection, but Faster R-CNN only extracts ROI without considering context information features. In addition, in the VGG network, a 2×2 Maxpooling is performed for each convolutional layer from conv1 to conv4. After four times of maximum pooling, the feature maps are down-sampled to 1/16 of the original. For example, a 32×32 area was reduced to 2×2 through "conv5", and a 16×16 block became a pixel. Therefore, when 7×7 is sampled in ROI Pooling, a lot of information is lost, resulting in poor performance of the target detection algorithm in small object detection.

In this paper, the features of context information are extracted at conv4 and conv5 layers. With the increase of convolution and pooling layers, the resolution of feature maps decreases gradually. Conv4 feature maps were 28×28 in size. The Conv5 feature map is 14×14 in size. Compared with Conv5, Conv4 has higher resolution but less semantic information. Therefore, three small size context information features of 1.5x, 2x and 3x were extracted from conv4. In conv5, context information features of 1.5x, 2x and 3x were extracted from conv5. The method of extracting the integer multiples of the original candidate box is easy to calculate the extracted original candidate box.

$$B_1 = [x, y, w, h] \tag{1}$$

Where (x, y) is the center coordinate of the candidate frame, w and h are the width and height of the candidate frame respectively. Extracting context information features is n-fold larger than the original candidate box. Candidate boxes for context information characteristics is:

$$B_2 = [x, y, nw, nh] \tag{2}$$

In conv4, n=1.5,2,3. In conv5, n=1.5,2,4. Figure 3 shows the extracted context information features of Conv5.

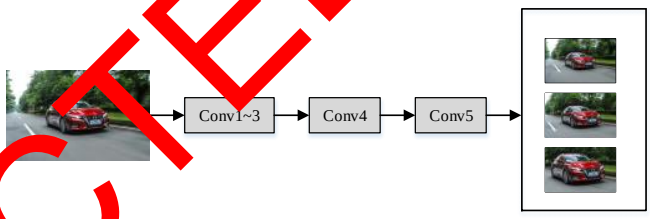

Figure 3. Contextual information of conv5

If the coordinates of the upper left and lower right corner of the original candidate box B_1 are (x_1, y_1) and (x_2, y_2) respectively, then,

$$x = \frac{(x_1 + x_2)}{2},\ y = \frac{(y_1 + y_2)}{2} \tag{3}$$

$$w = x_2 - x_1,\ h = y_2 - y_1 \tag{4}$$

For the candidate box B_2 of the improved context information feature, the coordinates of the upper left corner and lower right corner are (x_1', y_1') and (x_2', y_2') respectively, then,

$$x_1' = x - \frac{w \cdot roi\_sclae}{2} \tag{5}$$

$$x_2' = x + \frac{w \cdot roi\_sclae}{2} \tag{6}$$

$$y_1' = y - \frac{h \cdot roi\_sclae}{2} \tag{7}$$

$$y_2' = y + \frac{h \cdot roi\_sclae}{2} \tag{8}$$

Where roi_scale refers to the size of context information features. Round function is used to round the coordinates and output the positions on the corresponding feature graph.

2.4. Normalization operation

After extracting the context information features of different layers, the context information features are fused by concat layer. The features extracted from different convolutional layers of neural network have different scales. If the features are simply combined, the learning rate will be unstable.

Normalization is required to match the context information characteristics with the order of magnitude of ROI characteristics. This paper mainly studies L2 normalization layer and 1×1 convolution layer normalization. In CNN, L2 normalization layer is in the convolution of $N \times H \times W \times C_1$ to $N \times H \times W \times C_2$ to perform normal form calculation for different channels of samples in Batch:

$$y = \frac{x}{\sqrt{\sum_{i=1}^{k} x_i^2}} = \frac{x}{\sqrt{x^T x}} \tag{9}$$

Where $x = [x_1, x_2, \cdots, x_k]$ is the input vector. $y = [y_1, y_2, \cdots, y_k]$ is the output of the forward pass.

After Recurrent Neural Network (RNN) expanding, it is actually a full-connection layer and has different activation and structure compared with CNN. Therefore, the effect of L2 norm on CNN is far inferior to that of RNN. However, the convolution kernel of 1×1 can reduce or raise dimension to normalize the dimensions of different features, which is simpler than L2 norm. In addition, 1×1 convolution layer is required to ensure that the length of feature before entering the full connection layer is consistent with the original method. Therefore, 1×1 convolution layer conv6 is used in this paper to normalize the extracted context information features.

2.5. Fusion strategy

It adds a new layer (Eltwise layer) to the original network structure. This layer has multiple inputs, one output, and three operations (product, max, sum). *product* is the multiplication of corresponding elements. This operation will make the system relatively unstable and vulnerable to the influence of the weaker party, resulting in difficult convergence of the network. *max* means taking the maximum value of the corresponding element, which is equivalent to the set model in the network to some extent. When neither branch can detect the object, sum loses its mutual support advantage over sum [23]. *sum* means adding the corresponding input elements, and this layer defines the coeff parameter, which is used to adjust the weight. In this paper, element integration is chosen by element summing.

2.6. Negative sample classification

In order to solve the problem of classifier performance degradation caused by too many negative samples in the process of target detection, the difficult samples are automatically selected to join the training in the training process, which makes the training faster and more effective. Feature maps and all ROIs were propagated forward, followed by classification regression, and the loss of each ROI was calculated, including classification (cls) and positioning (dec). Using non-maximum suppression (NMS) selects the ROI of the first B (B= 128, jointly determined by the number of images contained in each batch during training N(N=1 in this paper) and the minimum batch-size of training (batch size =128 in this paper) with the largest loss according to ROI and corresponding loss iteration, and then inputs the network (for forward and backward) to learn and carry out gradient propagation. In network training, the overlapping rate of predicted candidate frame and marked candidate frames (Intersection-over) is adopted without human intervention Union (IOU) as the evaluation index. The ROI of forward propagation is the candidate area where the overlap rate of prediction candidate box and marker candidate box IOU is greater than 0.5; the ROI of backward propagation is the candidate area where the overlap rate of prediction candidate box and marker candidate box IOU is in the interval $[g, 0.5)$; g is set as 0.0 in the paper.

3. Experiment and result analysis

3.1. Experimental configuration

The experimental environment was as follows: UBUNTU16.04 system, 16 GB memory, Intel I7-9700K, 3.60 GHz CPU, GeForce RTX2070 8 GB video memory, deep learning platform CAFFE combined with CUDA8.0 and CUDNN V5.1. First, the Vggnet-16 network is pre-trained on the ImageNet data set for initialization, with an initial learning rate of 0.001 and momentum of 0.9, and Stochastic Gradient Descent (SGD) is used as loss propagation. Compared with Faster R-CNN, the threshold range parameters of negative samples are set differently. When the threshold value is [0.0,0.5], it is regarded as negative samples to reduce the difference between the number of negative samples and positive samples. After changing the threshold range, the training model was more balanced. However, when testing, the threshold range of negative samples is the same as the parameter in Faster R-CNN. Both the method and baseline in this paper are implemented on the Caffe framework, and the end-to-end training is adopted without the need to train multiple models, which greatly shortens the training time.

3.2. Experimental data sets and comparative methods

The data set used in this paper is PASCAL VOC, and VOC data set contains 20 classes. The PASCAL VOC2007 dataset contains 9963 images and 24,640 objects, of which the training set and validation set contain 5011 images and the test set contain 4952 images. There are 11540 images in the training set and validation set on the PASCAL VOC2012 dataset, with a total of 27450 objects [24,25].

Table 1 shows the statistics of the PASCAL VOC2007 dataset, which is divided into two main subsets: Trainval and Test. Trainval data is further divided into training (Train) and validation (Val) sets. By counting the number of images for each subset and class separately, it can be seen that there is a particularly large amount of data for the category person, which in previous datasets was basically a synonym for "pedestrian". However, on VOC data set, there are images of people engaged in various activities, such as horse riding and car riding, which make the detection of Person more complicated and increase the difficulty of target detection.

Table 1. Statistics of PASCAL VOC2007 dataset

	aero	bike	bird	boat	bus	car	cat	dog	tv
Trainval	238	243	330	181	186	713	337	421	256
Test	204	239	282	172	174	721	322	418	229

In addition, the images in the PASCAL VOC dataset were not taken specifically for object recognition purposes. An image usually contains multiple categories of objects, which may lead to objects extending to the outside of the image or shielding each other between objects in an image, increasing the difficulty of target detection. Therefore, it is not enough to rely only on the features of the object itself for target detection.

In this paper, mean average precision (mAP) is used as the evaluation index of detection accuracy. The comparison methods are as follows: Fast R-CNN, Fast R-CNN+context, Faster R-CNN, Non-Multi Layer Context (Non-MLC), AC-CNN, Single Shot Multibox Detector (SSD300), ION, SIN, HyperNetwork (HyperNet), CPF, proposed network.

3.3. Normalization method and fusion strategy

The results of testing different normalization approaches on the PASCAL VOC2007 dataset are as follows. The mAP of L2 normal layer is 74.5%, and the mAP of 1×1 convolution layer is 77.4%. Using L2 normalized layer can not reduce the loss, but reduce the network performance, which is not different from the effect of not adding 12 normalized layer. Using 1×1 convolutional layer mAP can improve 2.9%. Therefore, 1×1 convolution layer is selected for normalization operation [26-28].

The mAP of the three different fusion strategies is as follows: sum is 77.4%, prod has no data, and max is 76.5%. Thus, the element-by-element summation strategy has the best performance. Therefore, the context information features and the object's own features are selected to be integrated element by element.

3.4. Experimental result

Public VGGNET-16 is selected as the initialization model. Although ResNet version of Faster R-CNN can improve the target detection accuracy, ResNet is slower than VGGNET-16 in speed, requiring more training times. In addition, the emphasis of this paper is to verify the characteristics of context information to improve the accuracy of target detection. In summary, VGGNet-16 model is selected.

Training is performed on the PASCAL VOC2007 dataset and the presented method is evaluated on the PASCAL VOC2007 dataset. Table 2 compares the test results of the five methods. c adds context information on the basis of Fast R-CNN without adopting other optimization strategies, and 67.3% mAP is obtained, which is 1.4% higher than Fast R-CNN. Because we detect vehicles in this paper, we only display the results of vehicle.

Table 2. Test results of 5 methods on PASCAL VOC2007 dataset

Method	car	bus	Big vehicle	mAP
Fast R-CNN	78.2	77.3	79.6	65.9
Fast R-	72.8	67.1	69.9	67.3

CNN+context				
Faster R-CNN	79.2	75.1	78.0	68.5
Non-MLC	80.9	78.8	73.6	69.3
Proposed	82.4	78.7	83.9	71.1

In the proposed method, anchor points of six scales (1,2,4,8,16,32) and three aspect ratios (1:1,1:2,2:1) are used to obtain 18 anchor points of different sizes to strengthen the detection of small targets. The number of iterations of the whole network training is 80 000. The initial learning rate of the first 50,000 iterations is 0.001, and the learning rate of the last 30,000 iterations is reduced to 0.0001. Set the momentum parameter to 0.9, the weight attenuation parameter to 0.0005, and the effective batch size to 2. If 18 anchor points of different sizes are adopted in Faster R-CNN without adding context information features, 69.3% mAP can be obtained. The proposed method obtained 71.1% mAP, which was 2.6% higher than Faster R-CNN. It can be seen that the feature of context information is helpful to target detection and plays a great role in improving accuracy. The Faster R-CNN takes 0.139s to test an image, while the method in this paper only takes 0.117s. Therefore, the method in this paper is more effective both in time and accuracy.

The PASCAL VOC2007 training set and PASCAL VOC2012 training set were jointly trained, and 8 methods were evaluated on PASCAL VOC2007 data set. The test results are shown in Table 3. As the training set increases, the number of iterations of training needs to be increased. The learning rate of the first 60000 iterations is set at 0.001, and the learning rate of the last 40000 iterations is set at 0.000 1. The proposed method achieves 77.4% mAP on PASCAL VOC2007 data set, which is 4.2% better than Faster R-CNN without context information, and also better than other methods with context information.

Table 3. Test results of 8 methods on PASCAL VOC2007 dataset

Method	car	bus	Big vehicle	mAP
AC-CNN	81.5	80.1	81.9	72.0
Faster R-CNN	83.1	84.7	81.9	73.2
SSD300	84.2	83.0	84.5	74.3
ION	85.1	85.4	85.3	85.1
SIN	86.9	88.6	77.1	76.0
HyperNet	87.4	83.1	71.4	76.3
CPF	86.5	85.1	78.2	76.4
Proposed	87.2	85.1	70.9	77.4

Among the 20 categories in the dataset, the proposed method can improve the results more effectively for those categories that are easily obscured or considered as background, such as sofas, people, tables and chairs. In addition, the detection accuracy of vehicles with specific backgrounds, buses and other categories has also been greatly improved, and the addition of context information provides assistance for the detection of these categories.

4. Conclusion

This paper proposes a vehicle target detection algorithm based on RPN. Context information plays an important auxiliary role in target detection. In order to make effective use of context information, the combined context information and the characteristics of the object itself. Comparative experiments on different data sets show the effectiveness of the proposed method. Both contrast method did not join the context information and some of the ways to add context information, testing results of this method has obvious promotion, especially from the earlier context information levels of can enrich the characteristic information of the target detection to avoid some target loss of information, and improve small target or easily obscured target detection accuracy. Future efforts will be made to solve the problem of adaptive selection of context information to further improve the accuracy of target detection.

Acknowledgements.

The author greatly appreciates the anonymous review by the reviewers.

References

[1] Shoulin Yin, Ye Zhang, Shahid Karim. Large Scale Remote Sensing Image Segmentation Based on Fuzzy Region Competition and Gaussian Mixture Model[J]. IEEE Access. volume 6, pp: 26069 - 26080, 2018.

[2] Asif Ali Laghari, Hui He, Shahid Karim, Himat Ali Shah, Nabin Kumar Karn, "Quality of Experience Assessment of Video Quality in Social Clouds", Wireless Communications and Mobile Computing, vol. 2017, Article ID 8313942, 10 pages, 2017. https://doi.org/10.1155/2017/8313942

[3] Fang B, Fang L. Concise feature pyramid region proposal network for multi-scale object detection[J]. The Journal of Supercomputing, 2018:1-11.

[4] Ju M, Luo J, Liu G, et al. ISTDet: An efficient end-to-end neural network for infrared small target detection[J]. Infrared Physics & Technology, 2021, 114(7):103659.

[5] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137-1149, 1 June 2017, doi: 10.1109/TPAMI.2016.2577031.

[6] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.

[7] Yin Shoulin, Liu Jie, Teng Lin. A new krill herd algorithm based on SVM method for road feature extraction[J]. Journal of Information Hiding and Multimedia Signal Processing, v 9, n 4, p 997-1005, July 2018.

[8] Chen C, Wang G, Peng C, et al. Exploring Rich and Efficient Spatial Temporal Interactions for Real-Time Video Salient Object Detection[J]. IEEE Transactions on Image Processing, 2021, PP(99):1-1.

[9] Laghari A A, Jumani A K, Laghari R A. Review and State of Art of Fog Computing[J]. Archives of Computational Methods in Engineering, 2021(5).

[10] Yin Shoulin, Liu Jie, Li Hang. A Self-Supervised Learning Method for Shadow Detection in Remote Sensing Imagery[J]. 3D Research, vol. 9, no. 4, December 1, 2018. https://doi.org/10.1007/s13319-018-0204-9

[11] Torralba A. Contextual Priming for Object Detection[J]. International Journal of Computer Vision, 2003, 53(2):169-191.

[12] S. K. Divvala, D. Hoiem, J. H. Hays, A. A. Efros and M. Hebert, "An empirical study of context in object detection," 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 1271-1278, doi: 10.1109/CVPR.2009.5206532.

[13] J. Liu, X. Gao, N. Bao, J. Tang and G. Wu, "Deep convolutional neural networks for pedestrian detection with skip pooling," 2017 International Joint Conference on Neural Networks (IJCNN), 2017, pp. 2056-2063, doi: 10.1109/IJCNN.2017.7966103.

[14] Shahid Karim, Ye Zhang, Shoulin Yin, Muhammad Rizwan Asif. An Efficient Region Proposal Method for Optical Remote Sensing Imagery[C]. IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, pp: 2455-2458, July 2018.

[15] Shoulin Yin, Ye Zhang and Shahid Karim. Region search based on hybrid convolutional neural network in optical remote sensing images[J]. International Journal of Distributed Sensor Networks, Vol. 15, No. 5, 2019. SCI&EI(JA) DOI: 10.1177/1550147719852036

[16] Y. Liu, R. Wang, S. Shan and X. Chen, "Structure Inference Net: Object Detection Using Scene-Level Context and Instance-Level Relationships," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 6985-6994, doi: 10.1109/CVPR.2018.00730.

[17] Shrivastava A, Gupta A. Contextual Priming and Feedback for Faster R-CNN[C]// European Conference on Computer Vision. Springer International Publishing, 2016.

[18] H. Mei et al., "Exploring Dense Context for Salient Object Detection," in IEEE Transactions on Circuits and Systems for Video Technology, doi: 10.1109/TCSVT.2021.3069848.

[19] Wang Y, Hu S, Wang G, et al. Multi-scale dilated convolution of convolutional neural network for crowd counting[J]. Multimedia Tools and Applications, 2020, 79(1-2):1057-1073.

[20] A. Shrivastava, A. Gupta and R. Girshick, "Training Region-Based Object Detectors with Online Hard Example Mining," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 761-769, doi: 10.1109/CVPR.2016.89..

[21] Lin Teng, Hang Li, Shoulin Yin, Shahid Karim &Yang Sun. An active contour model based on hybrid energy and fisher criterion for image segmentation[J]. International Journal of Image and Data Fusion. Vol.11, No. 1, pp. 97-112. 2020.

[22] Shoulin Yin, Hang Li, Lin Teng, Man Jiang & Shahid Karim. An optimised multi-scale fusion method for airport detection in large-scale optical remote sensing images [J]. International Journal of Image and Data Fusion, vol. 11, no. 2, pp. 201-214, 2020. DOI: 10.1080/19479832.2020.1727573

[23] Y. Zhu, C. Zhao, J. Wang, X. Zhao, Y. Wu and H. Lu, "CoupleNet: Coupling Global Structure with Local Parts for Object Detection," 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 4146-4154, doi: 10.1109/ICCV.2017.444.

[24] Xiaowei Wang, Shoulin Yin, Ke Sun, Hang Li, Jie Liu and Shahid Karim. GKFC-CNN: Modified Gaussian Kernel Fuzzy C-means and Convolutional Neural Network for Apple Segmentation and Recognition [J]. Journal of Applied Science and Engineering, vol. 23, no. 3, pp. 555-561, 2020.

[25] Shoulin Yin, Hang Li, Desheng Liu and Shahid Karim. Active Contour Modal Based on Density-oriented BIRCH Clustering Method for Medical Image Segmentation [J]. Multimedia Tools and Applications. Vol. 79, pp. 31049-31068, 2020.

[26] Khan AA, Shaikh ZA, Laghari AA, Bourouis S, Wagan AA, Ali GAAA. Blockchain-Aware Distributed Dynamic Monitoring: A Smart Contract for Fog-Based Drone Management in Land Surface Changes. Atmosphere. 2021; 12(11):1525. https://doi.org/10.3390/atmos12111525

[27] Laghari A A, He H, Shafiq M, et al. Application of Quality of Experience in Networked Services: Review, Trend & Perspectives[J]. Systemic Practice and Action Research, 2019, 32(5):501-519.

[28] Laghari A A, Laghari K, Memon K A, et al. Quality of Experience (QoE) Assessment of Games on workstations and Mobile[J]. Entertainment Computing, 2020, 34:100362.