Eng Int Syst (2019) 4: 185–192

Engineering
Intelligent Systems

Application of Deep Convolutional Neural Network in Computer Vision

Lepeng Chen* **and Chengjiang Wang†**

College of Electricity and New Energy, Three Gorges University, Yichang 443002, Hubei, China
College of Electricity and New Energy, Three Gorges University, Yichang 443002, Hubei, China

With the rapid development of computer technology, computer vision brings a lot of convenience to people's lives, study and work. Because computer vision is a newly developed field, there are still many problems in computer vision research. For example, it is not perfect for image segmentation, target detection, and image classification. This paper addresses computer vision, mainly to improve image segmentation and target detection in computer vision. In this paper, deep convolutional neural networks are used, which must first be constructed. Then, both the loss function of the network and the gradient magnitude are calculated. Finally, the sharing mode of the network and the target detection based on the bidirectional feature pyramid, are examined. The results show that the image segmentation and target detection using the deep convolutional neural network can make the gradient amplitude of the filler sub-branches in the basic network significantly larger than the gradient amplitude of the instance segmentation network. The sharing mode can bring about a 70% improvement in performance. Furthermore, the target detection method based on the bidirectional feature pyramid extracts more information.

Keywords: Computer Vision, Deep Convolutional Neural Network, Image Panorama Segmentation, Image Target Detection.

1. INTRODUCTION

Computer vision refers to the science of letting computers understand graphical and video information, extracting useful information from graphics and video and analyzing it. Computer vision is a fairly new and rapidly developing field of research and has become one of the important research areas of computer science. The field of computer vision is also the hottest topic among the artificial intelligence community. The technology in this field is considered to be compatible with the industry and has a significant and realistic impact on industry use cases. Some of the technologies in this field have surpassed the performance levels achievable by humans and exceeded the accuracy and reliability standards expected by most industries. The amazing advances in basic computer vision tasks, such as image classification, have enabled multiple technologies to becombined to create a new composite technology with applications that have never been explored in industrial environments to date.

The research on computer vision is very broad, and the research on image segmentation and image target detection in the field of computer vision has received extensive attention. In [1], in order to effectively segment medical images with rich noise, blurred boundaries and uneven intensity, a hybrid active contour model combining global information and local information is proposed. Then a new global energy functional is constructed and the adaptive weight is given. The results show that the method can effectively segment synthetic images and medical images, and has higher segmentation performance than other related methods. In [2], the author focuses strongly on the shortcomings of the initial user input in order to overcome the existing image segmentation techniques. A robust and efficient image segmentation technique based

*Email: ccxdxctgu@163.com
†Corresponding author (Email: ccxdxctgu@163.com)

on a mixed convex active contour and CHAN-VESE (CV) model is proposed. The experimental results show that the proposed algorithm outperforms the recently proposed indoor and outdoor image segmentation algorithms in processing time and segmentation accuracy. In [3], in order to overcome the serious complications of resection, the authors proposed a three-dimensional model of 30 female pelvises using ITK-SNAP and internal CAS segmentation pipeline extension technique for image segmentation. The experimental results show that the three-dimensional reconstruction of the female pelvis using ITK-SNAP before surgery can help to obtain a widely-used tibia;however, further comparative studies are needed to evaluate the effects before and after use. In [4], in order to solve the problem of target spatial location and time consuming in video surveillance system, the author proposes a target spatial localization algorithm based on RNN-LSTM deep learning. At the same time, OpenGL perspective imaging and photogrammetry consistency principle and 3D scene simulation imaging technology are used. Then according to the corresponding relationship between video image and simulation image, the target is located. The experimental results show that the target detection has high precision and has important reference value for the outdoor target location of video surveillance images. In [5], in order to realize the automatic and accurate identification of the shipwreck target in the side scan sonar (SSS) waterfall image, the author constructs a pipeline containing feature extraction, selection and shipwreck identification using the sample image. The ADABOOST model is constructed, and the shipwreck target is quickly detected by the nonlinear matching model. The shipwreck target recognition method in the SSS waterfall image is given. According to the wide combination of these different procedures, the shipwreck is accurately identified. The experimental results show that the correct recognition rate of the sample image is 97.44%, the false positive rate is 3.13%, and the missed detection rate is 0. These works indicate that image segmentation and target detection based on computer vision have been studied by many scholars, but since the field of computer vision is still in its infancy, many technologies for image segmentation and target detection are still not mature. Further and more comprehensive research is needed to improve the knowledge and understanding of image segmentation and target detection in computer vision.

With the development of science and technology, the era of big data has followed, giving the deep convolutional neural network (CNN) a richer network structure. Compared with traditional machine learning, it has advantages in terms of feature expression and feature learning. Based on the deep learning algorithm deep convolutional neural network model, the computer vision field has achieved remarkable results in recognition ability. Deep convolutional neural networks have many applications in the field of computer vision, such as image classification, object detection, image segmentation, etc. In [6], the author designed a 13-layer convolutional neural network (CNN) for the use of hand-made functions in computer vision systems for fruit classification, using three data enhancement methods: image rotation, gamma Ma correction and noise injection. The experimental results show that the overall accuracy of the method is 94.94%, which is at least 5 percentage points higher than the advanced method. In [7], the author applies a deep convolutional neural network to the automatic detection of pulse, and proposes two deep neural network (DNN) structures that use short electrocardiogram segments (5s) to detect pulse, that is, they divide the pulse into pulseless electricity. Activity (PEA) or pulse produces rhythm (PR). Experimental results show that both architectures improve the performance of the most advanced methods developed to date, increasing it by more than 1.5 points. In [8], the author uses neural networks to detect the number of olive Ridley turtles. The results show that the model built using the neural network can detect 8% more turtles than bymanual count. This finding highlights the feasibility of combining the drone with the neural network to estimate the population levels of various marine animals. In [9], the author proposed a deep convolutional neural network (DCNN) model for the analysis of clinical ultrasound sonogram data. The experimental results show that the diagnostic accuracy of thyroid cancer is improved. In [10], in order to effectively use the deep learning technique to obtain the airfoil shape, the author proposes a method of implementing the airfoil inverse design using DCNN. The results show that the CNN method has a certain calculation time, high calculation efficiency and high prediction accuracy. The above literatures all show that deep convolutional neural networks can be widely used in various fields and solve some practical problems. Similarly, DCNNs can also be applied to related research in the field of computer vision, making huge contributions to the development of the computer field.

In order to study the field of computer vision, in this paper, we adopt the notion of deep convolutional neural networks to perform image panoramic segmentation and target detection on computer vision images. In order to achieve this operation, we first need to construct a network of deep convolutional neural networks, then calculate the loss function and gradient magnitude of the network and study the sharing mode of the network. After research, we can obtain image segmentation and target detection using deep convolutional neural network. The gradient amplitude of the filler sub-branch into the basic network can be significantly larger than the gradient amplitude of the instance segmentation network. The sharing model can bring 70% of the revenue. The performance improvement of the left and right, the target detection method based on the bidirectional feature pyramid extracts more information.

2. METHODS

2.1 Deep Convolutional Neural Network

Convolutional neural networks generally organize important data of two-dimensional input and gradually establish artificial multilayer neural networks. Throughout the whole network, each layer consists of two-dimensional planes, and in each plane, there are many relatively independent neurons, and adjacent neurons can be connected to each other, but if two nerves are at the same level, they cannot be connected. At present, the development of neural networks is booming, and has gradually become an important research element of speech analysis and image recognition. As learning from higher areas is progressive, it is also richer in application changes. In essence, the convolutional neural network is the first multi-layer neural network to be successfully developed. This model algorithm can make the input of multiple signals more convenient. With the gradual increase of learning depth, information learning is becoming a craze. Nowadays, convolutional neural networks can be perfectly applied in speech recognition, image recognition, speech processing, etc., and machine learning is developing in a stronger direction. The structure of the convolutional neural network usually includes modules such as an input layer, a Convolution Layer, an Activation layer, a Pooling Layer, a Full Connected Layer, and an output layer.

(1) Convolution layer

The convolutional layer is the most basic layer of CNNs, the purpose of which is to convolve the input image with the learnable filter and extract features. Each filter can produce a feature map. When CNNs is used for RGB images, the image is a 3D matrix, and each layer is the same. Each convolution layer of CNN is composed of a 3D filter of size D × H × W, and H × W is a spatial dimension, which is equivalent to the set of neurons, and d is the number of core (filter) feature channels.

The step size of the filter is defined as the interval at which the filter moves in each spatial dimension. The fill parameter p corresponds to the number of pixels added to the outer edge of the input. Therefore, the step size can be considered as the input method of subsampling. Usually, the h=w=f Square filter is used. The output of this layer is calculated using equations (1), (2), and (3).

$$D_0 = N \tag{1}$$

$$H_0 = \frac{H_i - f + 2p}{S} + 1 \tag{2}$$

$$W_0 = \frac{w_i - f + 2p}{S} \tag{3}$$

(2) Pooling layer

The pooling layer is usually between convolutional layers;the pooling layer is decentralized and the input features can be subsampled. Pooling is achieved by sliding a filter over the input image. The input images (usually non-overlapping) are divided into different sub-regions, each sub-region being down sampled by a non-linear pooling function. The most common of these functions are the maximum and average pooling functions. The maximum pooling will return the maximum value from each sub-area, and the average pooling returns the average of the sub-areas. Pooling increases the robustness ofthe network by reducing the number of translation inconsistencies. In addition, the pooling removes unnecessary and redundant features, reduces the computational cost of the network, and improves the efficiency of the network.

(3) Activation layer

In general, the output of the convolutional layer is the input to the active layer. The active layer can consist of nonlinear functions such as SIGMOID, TANH, and RELU. RELU has a good effect in CNNs and is generally the preferred activation function. The RELU layer converts the negative input to 0, and does not change its value for the positive input, as shown in equation (4):

$$\mathrm{f(x)} = \begin{cases} x, x \geq 0 \\ 0 \quad \text{others} \end{cases} \tag{4}$$

where x is the input to the activation function RELU and f(x) is the active output.

(4) Fully connected layer

After the high-level features are extracted through the convolutional layer, the pooled layer and the RELU layer, the fully-connected layer is typically placed at the end of the network. The neurons in this layer are completely dependent on all activations of the previous layer. The most important role of the fully connected layer is that the neurons in that layer determine which features correspond to which categories. In short, a fully-connected layer can be thought of as a layer that provides a classifier. When neurons in the fully connected layer accept the activation of all input neurons, spatial information is lost, which is undesirable in image semantic segmentation where spatial information is very important.

(5) Classifier

The choice of classifier is determined by considering the current problem and the data used. In this paper, we use the SOFTMAX function to predict categories. For binary problems, the SOFTMAX function is reduced to logistic regression. The SOFTMAX function in equation (5) gives a probability that an input belongs to class c.

$$P_c = \frac{e^{(s_c)}}{\sum_{i=1}^{c} e^{s_i}} \tag{5}$$

where s is the network output of a particular class. For a single input, the sum of all probabilities between classes is always equal to 1, and the loss function is defined as the negative logarithm of the SOFTMAX function probability.

2.2 Image Panorama Segmentation

Image segmentation divides the image into techniques and processes that have homogenous regions with their respective characteristics, and extracts the target object of interest. It is a key step in image processing to conduct image analysis, and this is one of the basic problems in image processing and robot vision. It has become a hot topic in social research and is being used more and more widely in various fields such as remote sensing meteorological services, medical image analysis, military research, traffic image analysis, image compression, and image retrieval.

(1) Construction of panoramic segmentation network based on convolutional neural network

The schematic diagram of the panoramic segmentation network structure constructed in this paper first needs to give an input image. This paper uses the feature pyramid structure to provide feature maps for the filler semantic segmentation branch and the object instance segmentation branch. After the two sub-branches have generated their respective intermediate prediction results, the result is transmitted to the network fusion module and, finally, the prediction result of the panoramic segmentation is obtained.

(2) Loss function

$$\mathrm{L}(p_2, l_2) = \frac{1}{N_{cls}} \sum_{2} L_{cls}(p_2, p_2^*) + \lambda \frac{1}{N_{reg}} \sum_{2} p_2^* L_{reg}(t_2, t_2^*) \tag{6}$$

Where p_2 is the predicted probability that the second anchor is the target. If the anchor is positive, its label p_2^* is 1; if the anchor is negative, its label p_2^* is 0, t_2 is a vector representing the four parameterized coordinates of the predicted bounding box, and t_2^* is the coordinate vector of the true label bounding box corresponding to the positive anchor.

For the classification loss function, L_{cls} is the logarithmic loss function of the object class (including the background), and the formula is:

$$L_{cls}(p_i, p_i^*) = -\log\left[p_i^* p_i + (1 - p_i^*)(1 - p_i)\right] \tag{7}$$

For the regression loss function, the formula is:

$$L_{reg}(t_i, t_i^*) = smooth_{L_1}(t_i - t_i^*) \tag{8}$$

The definition of the $smooth_{L_1}$ function is:

$$smooth_{L_i}(\mathrm{x}) = \begin{cases} 0.5x^2 \\ \pm x - 0.5 \end{cases} \tag{9}$$

p_i^* means that only the anchor with a positive prediction will have a regression loss. In other cases, the gradient of the loss function will not be returned.

2.3 Image Target Detection

Target detection is widely used for image retrieval, video surveillance, military reconnaissance, and in other fields. It is intended to automatically identify the classification information and location information of the target from complex scenes. In view of the poor portability of traditional modeling methods, recent deep learning has become the main means of research on target detection. In order to improve the accuracy of target detection based on deep learning, this paper studies the target classification and target location as the main breakthrough points.

(1) Network construction

Mask R-CNN is the best detection accuracy in the current target detection framework, so this paper studies and improves on this framework. It adds a mask prediction branch based on Faster R-CNN, and integrates FPN into the Residual Network (Res Net) and improves Ro I Pooling as Ro I Align Pooling in the prediction frame extraction process. The use of bilinear interpolation replaces the simple rounding down of the original method. Specific steps are as follows:

A. Input the image to the shared convolutional layer extraction feature, and generate a multi-scale feature map using FPN.

B. Perform side connection on the generated feature map, and double-upsample the feature map of each stage and add the tensor to the adjacent lower layer.

C. Use RPN to generate candidate regions for different size feature maps, and input them into the feature map to obtain the prediction frame.

D. Conduct classification and regression positioning of prediction frames.

(2) Target detection method based on bidirectional feature pyramid

In deep convolutional neural networks, FPN combined with various backbone networks has become the most effective and commonly-used method for feature extraction. In the process of feature extraction, feature maps of different sizes are generated by different layers, and then the feature maps are adjacent. The feature map of the layer is connected side by side from top to bottom, and finally the different size feature maps generated by the aliasing effect are convolved as an output. The advantage of this method is that different levels of information can be merged together for processing. The disadvantage is that the positioning information is lacking in the strong semantic feature map, especially the performance of the upper layer is more obvious, which ultimately affects the accuracy of detection. Hence, this study is intended to address the shortcomings of FPN, propose a target detection method based on a bidirectional feature pyramid, add a bottom-up path, and fuse the feature map as the input of RPN. The robustness of the algorithm is enhanced. The positioning information in the semantic feature map further improves the accuracy of the detection.

The classification loss function in the network training loss can calculate the sparse SOFTMAX cross entropy between the probability and the label, and measure the probability in the discrete classification task in which the categories are mutually exclusive. Loss is calculated using positive and negative samples.

$$\mathrm{softmax}(x)_1 = \frac{\exp(x_1)}{\sum_j \exp(x_j)} \tag{10}$$

$$H'_y(y) = -\sum_{1} y'_1 log y_1 \tag{11}$$

The specific steps of the loss function are shown in the formula above. First, the SOFTMAX classification is calculated by using equation 10; where x_1 is the score of the first category in the forecast, and the obtained result SOFTMAX is the vector of each category probability, and the highest value is the classification result. The output vector of equation 10 and the actual label of the sample are then used to calculate the cross entropy using equation 11; where y'_1 refer to the first value in the actual label, and y_1 is the value of the first element in the output vector of equation 10. Since the output is a vector, the result of the cross entropy needs to sum the vectors, and the loss is the average of the vectors.

Table 1 Gradient descent algorithm steps.

Step 1	Input: objective function f(x), gradient function g(x) = ∇ f(x)
Step 2	Let k = 0, take the initial value $x^0 \in R^n$
Step 3	Calculate the k iteration value f(x^k)
Step 4	Calculate the gradient $g_k = g(xk)$, stop iterating when $\|g_k\| < \varepsilon$, let $x^* = x^k$; otherwise let the search direction $p_k = -\nabla f(x^k)$, find the step size λ_k. The equation f($x^k + \lambda_k p_k$) = $\min_{\lambda \geq 0} f(x^k + \lambda p_k$ is established.
Step 5	Let $x^{k+1} \leftarrow x^k + \lambda_k p_k$ and calculate $f(x^{k+1})$, when $\| f(x^{k+1}) - f(x^k)\| < \varepsilon$ or $\|x^{k+1} - x^k\| < \varepsilon$ stops iteration, so $x^* = x^{k+1}$
Step 6	Otherwise, let k = k + 1 repeat Step 4 and Step 5
Step 7	Output: the minimum value of f(x)x^* and the optimal solution f(x^*)

3. EXPERIMENT

(1) Evaluation of image panorama segmentation algorithm

The panoramic segmentation task requires a class prediction for each pixel of the full image and an instance prediction for different instances. Since category prediction and instance prediction in panoramic segmentation are closely related to semantic segmentation and instance segmentation, semantic segmentation and instance segmentation can be merged into panoramic segmentation by means of simple merge operations. Current research considers it as filler class semantic segmentation and object class. The instance splits two independent subtasks and obtains the result of the panoramic segmentation by simple merge. The traditional method will predict the semantic division of the filler and the segmentation of the object instance through two independent networks, and then simply merge the prediction results of the two networks to obtain the result of the panoramic segmentation. The algorithm designs a new network structure. The two sub-tasks of the panoramic segmentation can share the low-level image features, and at the same time, the task prediction can be performed through the respective classification regression layers, and the effects of mutual interference are obtained.

(2) Evaluation of image target detection algorithm

The image object detection algorithm, based on a deep convolutional neural network, uses a back propagation algorithm. It is difficult to complete the training of neural networks with simple learning rules, so effective learning algorithms are indispensable. In 1988, RUMELHART et al. proposed a back-propagation algorithm, also known as an error back propagation algorithm, which makes the training of neural network simple and feasible. At the same time, it is also the core algorithm for parameter updating of convolutional neural networks. Back propagation algorithm is the most commonly used method of neural network training. By calculating the bias of the error on the parameters, the error is reduced along the gradient direction, and iteratively, the error is minimized. The core method is to use chained rules to pass the error layer by layer, thus updating the weight of each layer. In convolutional neural networks, back propagation is different from that used for general neural networks. First, in the forward propagation, the input is compressed, and the inverse derivation error requires up-sampling; second, the volume product operation is convoluted by tensor and summed by several matrix convolutions. The output of direct matrix multiplication is different from the general network, so the recursive calculation method of error is different.

Another algorithm based on deep convolutional neural network image target detection is the gradient descent algorithm. This algorithm includes many different methods such as batch gradient descent algorithm, Stochastic Gradient Descent (SGD) algorithm and compromise gradient descent algorithm. The SGD algorithm achieves optimal results in the shortest path by continuously judging and selecting the optimal path under the current target. Given the regression equation, the optimal variable required for the least squares method is to minimize the square of the deviation between the calculated value and the actual value. The SGD algorithm needs to solve the optimal data under the current position by continuously obtaining the partial derivative. The gradient descent algorithm is also an iterative algorithm that updates the value of x by successive iterations of the initial value x_0, minimizing the objective function until convergence. Since the direction in which the function value drops the fastest is the negative gradient direction, the update of x is performed in the negative gradient direction during the iterative process. Suppose x^* is the minimum point of f(x), f(x^*) is the optimal solution, and the execution steps of the gradient descent algorithm are shown in Table 1.

4. RESULTS AND DISCUSSIONS

4.1 Gradient Magnitude Comparison

In order to explore the gradient differences of different self-branches, the difference in the magnitude of the gradients introduced by different sub-branches in the same network layer is plotted. As shown in Figure 1, the blue curve indicates that the gradient formed by the filler sub-branch supervision signal is back-transferred to the gradient value of the last layer of res5, and the yellow curve indicates that the gradient formed by the example sub-branch supervision signal is back-transferred to the last layer of res5. Wherein, the abscissa of the curve represents the iteration order number, and the ordinate represents the average gradient size of the layer, and the calculation method is the root mean square calculation of the gradient input of the output feature map of the layer. The same initialization method and hyper parameter configuration are

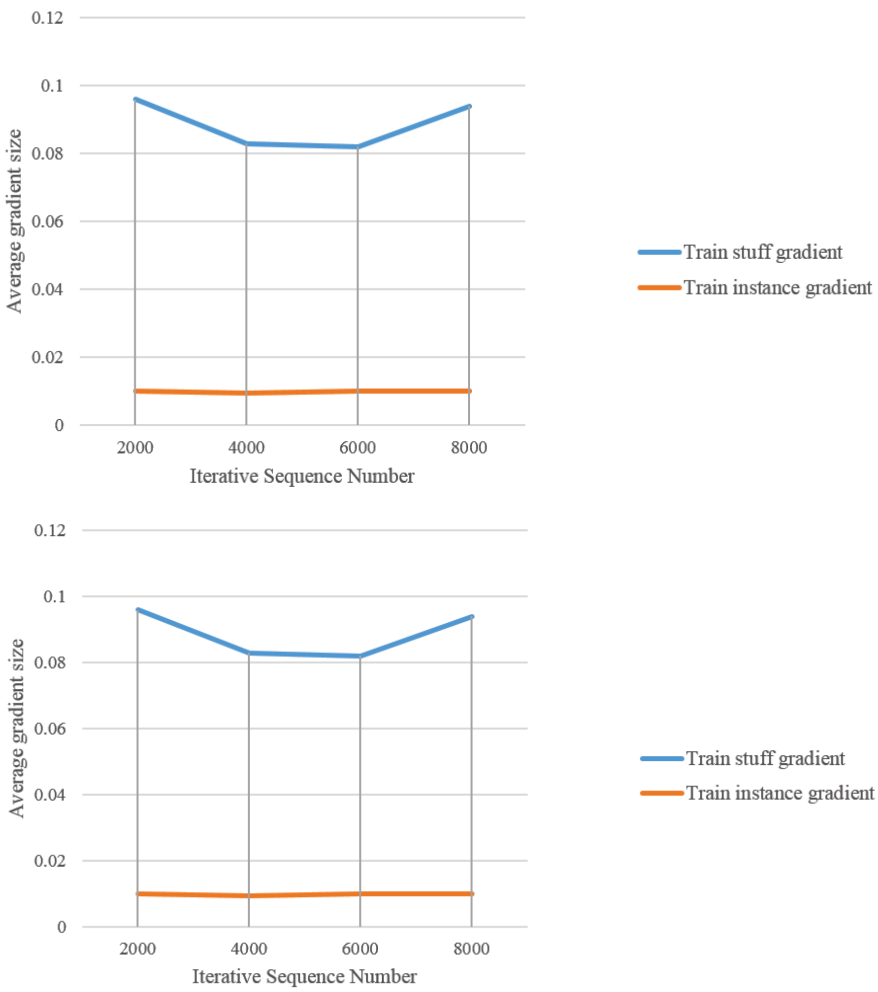

Figure 1 Comparison of the gradient magnitudes of different self-branched incoming basic networks.

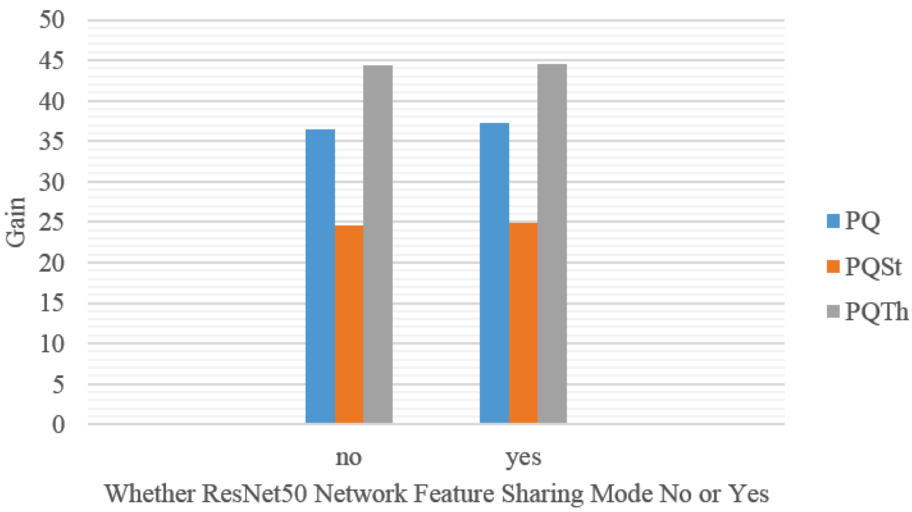

Figure 2 Experiment of ResNet50 network feature sharing mode gain.

used when performing separate training for each branch. It can be clearly seen in Figure 1 below that the gradient amplitude of the filler sub-branch into the basic network is significantly larger than the gradient amplitude of the instance segmentation network; hence, if the gradient is not balanced, the instance segmentation network will be significantly interrupted. This means that the underlying network cannot learn the common basic features that are suitable for instance segmentation.

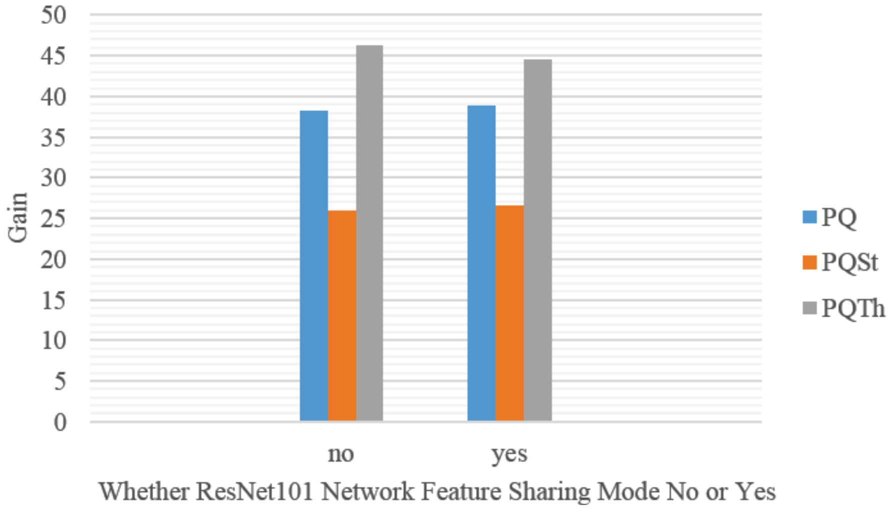

Figure 3 ResNet101 network feature sharing mode gain exploration experiment.

a1 b1

Figure 4 Mask R-CNN (left) and this method (right) comparison chart.

4.2 Shared Mode can Bring Gain to Network Prediction

For the exploration of feature sharing mode, the following experiments are carried out. First, the experiments determined whether the feature-sharing mode can benefit the convolutional neural network and, secondly, they identified the form of shared network structure that is most effective for the panoramic segmentation task. First, an analysis was conducted to determine whether the shared-feature model has a positive effect on the convolutional neural network under different model sizes, as shown in Figure 2 and Figure 3. In the experiment, the ResNet50 small network and the ResNet101 large network were used as the basic network respectively, and the network was built according to the sharing mode. The experimental results show that this feature-sharing mode does not affect the feature learning of the two self-branches, but the total panoramic segmentation index PQ is optimized. Through this sharing mode, the two sub-networks can learn more general basic features through training, so that the prediction of both sub-networks can be optimized and improved. The test results show that the sharing mode can bring about a performance improvement of about 0.7 for different network model sizes.

The experimental results show that the sharing mode can bring about 70% performance improvement for different network model sizes. In addition, the experimental results show that the more basic features are shared, which is the greatest prediction gain for the entire network.

4.3 Extracting Information Using Target Detection Methods

In the subjective evaluation experiment, images were compared using Mask R-CNN and the detection results of this method. Of these, a1 on the left is the bounding box detection result of Mask R-CNN, and b1 on the right side is the result of the method used in this paper. It can be intuitively seen that this method extracts more information than does the original FPN structure. As shown in Figure 4, a1 and b1, the car on the left edge of the image and the bag in the middle of the figure are detected bythe Mask R-CNN method. But they didn't detect either. All the methods in this paper are detected, and there is no missed detection compared with mask R-CNN. Moreover, according to the second set of data and the first and third sets of data analysis, the target size of the Mask R-CNN missed detection is large and small, and the location also includes the edge and middle of the image. The reason for this is that the feature extraction is not rich enough;the problem is not just that the small target is difficult to detect. The experimental results show that the feature information fully utilized in this method can improve the accuracy of detection.

5. CONCLUSIONS

Computer vision is a new field, and panoramic image segmentation is a new and important task in the field of computer vision. It has broad application prospects in various fields such as scene understanding, driverless vehicles, new retail, and intelligent robots. In this paper, the purpose of panoramic image segmentation based on convolutional neural network is studied. The design of image panoramic segmentation network based on convolutional neural network is constructed. The feature-sharing mode and loss function are proposed in order to achieve panoramic segmentation in a network. Training and prediction, the proposed image panoramic image segmentation method based on a convolutional neural network, has excellent prediction results under different basic models. Image target detection is also an important task in the field of computer vision. In this paper, a target detection method is proposed based on a bidirectional feature pyramid. The experimental results show that the proposed method can improve the detection accuracy. Of course, in the follow-up work, the following shortcomings need to be addressed:

(1) In this paper, a detailed analysis and exploration of panoramic segmentation and related problems was carried out. An end-to-end panoramic segmentation network was and proposed and designed. Although some progress has been made, its universal applicability needs to be improved.

(2) In terms of the detection accuracy for large targets, the improvement rate is small. The next step is to eliminate any redundant information on the basis of the two-way feature pyramid to achieve better detection results. The threshold is established manually. There are some disadvantages. The next step is to find a way to adaptively select the size of the threshold. In the experiment, it was found that for some images, the target near the edge of the image has a higher response on the feature map, although this is not evident in the final test result. The next step is to eliminate this boundary effect as much as possible.

REFERENCES

1. Fang, L., QIU, T., Zhao, H., & LV, F. (2019) "A hybrid active contour model based on global and local information for medical image segmentation". *Multidimensional Systems and Signal Processing*, 30(2), pp.689–703.
2. DERICHE, M., Amin, A., & Qureshi, M. (2019) "Color image segmentation by combining the convex active contour and the Chan VESE model". *Pattern Analysis and Applications*, 22(2), pp.343–357.
3. ALBANESI, G., GIANNINI, A., Carbone, M., Russo, E., MANNELLA, P., Ferrari, V., & SIMONCINI, T. (2019) "Computed-tomography image segmentation and 3D-reconstruction of the female pelvis for the preoperative planning of SACROCOLPOPEXY: preliminary data". *International UROGYNECOLOGYJournal*, 30(5), pp.725–731.
4. Liu, F., Chen, Z., & Wang, J. (2019) "Video image target monitoring based on RNN-LSTM". *Multimedia Tools and Applications*, 78(4), pp.4527–4544.
5. Zhu, B., Wang, X., Chu, Z., Yang, Y., & Shi, J. (2019) "Active learning for recognition of shipwreck target in side-scan sonar image". *Remote Sensing*, 11(3), pp.243.
6. Zhang, Y. D., Dong, Z., Chen, X., JIA, W., Du, S., Muhammad, K., & Wang, S. H. (2019) "Image based fruit category classification by 13-layer deep convolutional neural network and data augmentation". *Multimedia Tools and Applications*, 78(3), pp.3613–3632.
7. ELOLA, A., ARAMENDI, E., IRUSTA, U., PICON, A., Alonso, E., Owens, P., & IDRIS, A. (2019) "Deep Neural Networks for ECG-Based Pulse Detection during Out-of-Hospital Cardiac Arrest". *Entropy*, 21(3), pp.305.
8. Gray, P. C., Fleishman, A. B., Klein, D. J., MCKOWN, M. W., BEZY, V. S., LOHMANN, K. J., & Johnston, D. W. (2019) "A convolutional neural network for detecting sea turtles in drone imagery". *Methods in Ecology and Evolution*, 10(3), pp.345–355.
9. Li, X., Zhang, S., Zhang, Q., Wei, X., Pan, Y., Zhao, J. & Yang, F. (2019) "Diagnosis of thyroid cancer using deep convolutional neural network models applied to SONOGRAPHIC images: a retrospective, MULTICOHORT, diagnostic study". *The Lancet Oncology*, 20(2), pp.193–201.
10. SEKAR, V., Zhang, M., SHU, C., & KHOO, B. C. (2019) "Inverse Design of Airfoil Using a Deep Convolutional Neural Network". *AIAA Journal*, 57(3), pp.993–1003.

Lepeng Chen was born in 1999, and is an undergraduate at the School of Electrical and New Energy. Now there is one patent for invention, and many of them have been put into practice. He has participated in the Enzhipu Smart Car Competition. Lepeng'scurrent research focus is the convolutional neural network algorithm.
E-mail: ccxdxctgu@163.com

Chengjiang Wang Man, Baokang, Hubei, was born in 1971. He received his master's degree from Wuhan University. Now he is a professor of electrical engineering, master's tutor, electrical engineering department, and academic backbone of China's Three Gorges University. He is engaged mainly in the research and development of condition monitoring and fault diagnosis of power equipment, discharge and insulation phenomena, power system relay protection, power informatization, etc.
E-mail: ccxdxctgu@163.com