A novel dilated convolutional neural network model for road scene segmentation

Yachao Zhang[1] and Yuxia Yuan[1,*]

[1]School of Electronic and Electrical Engineering, Zhengzhou University of Science and Technology
Email:316168430@qq.com;xdwangxd@163.com

Abstract

Road scene understanding is one of the important modules in the field of autonomous driving. It can provide more information about roads and play an important role in building high-precision maps and real-time planning. Among them, semantic segmentation can assign category information to each pixel of image, which is the most commonly used method in automatic driving scene understanding. However, most commonly used semantic segmentation algorithms cannot achieve a good balance between speed and precision. In this paper, a road scene segmentation model based on dilated convolutional neural network is constructed. The model consists of a front-end module and a context module. The front-end module is an improved structure of VGG-16 fused dilated convolution, and the context module is a cascade of dilated convolution layers with different expansion coefficients, which is trained by a two-stage training method. The network proposed in this paper can run in real time and ensure the accuracy to meet the requirements of practical applications, and has been verified and analyzed on Cityscapes data set.

Keywords: road scene segmentation, dilated convolutional neural network, scene understanding.

*Corresponding author. Email:xdwangxd@163.com

1. Introduction

Automatic driving is one of the hot and cutting-edge topics, mainly divided into visual perception, planning control, high-precision map construction and other modules. The visual perception module is mainly used for target detection and recognition, including detection of surrounding obstacles (pedestrian detection, vehicle detection, etc.), lane and line detection, traffic sign recognition, etc [1-3]. In automatic driving, only obstacle detection is not enough, and semantic understanding of obstacle targets on the road is also needed. Semantic level understanding can provide more information for planning and control, and it also plays an important role in building

semantic maps with higher accuracy. At present, semantic segmentation is a common method for road scene understanding. It predicts the category of each pixel of the image to achieve segmentation of the image, thus providing more abundant information, which is crucial for automatic driving [4,5].

2. Image semantic segmentation

In recent years, deep learning has developed rapidly. After AlexNet network was proposed [6], convolutional neural network (CNN) gradually occupies the mainstream position in various visual tasks. At the beginning, CNN was mainly used in image recognition tasks, which could

conduct end-to-end training on images and then make category prediction on images. In 2015, the reference [7] proposed fully convolutional neural network (FCN), which applies CNN to semantic segmentation task and replaces the full connection layer in CNN network with the convolutional layer. This design allows the network to accept images of any size.

After the FCN network was proposed, more and more researchers applied deep learning to semantic segmentation tasks [8-10]. At present, the research on semantic segmentation mainly focuses on two aspects: precision and speed. In order to improve the accuracy, the complex network structure is designed to increase the depth and width of the network. For example, DeeplabV3+ designed an ASPP module [11] to expand the receptive field of convolution with different cavity rates, aggregate multi-scale feature information, and then obtain the final segmentation result through up-sampling. PSPNet designs a Pyramid Pooling module, which uses spatial Pyramid Pooling to merge the context information of different regions in the feature graph. This kind of segmentation network has high precision but low operating efficiency. Some algorithms emphasize the segmentation speed and improve the running speed by reducing the network complexity and the number of network parameters. ENet achieves the purpose of improving segmentation efficiency by reducing the number of parameters through continuous fast down-sampling and asymmetric convolution. ESPNet decomposed standard convolution into point-wise convolution and spatial pyramid void convolution, which greatly reduced the number of network parameters. Although these networks improve the running speed, but the accuracy is greatly reduced, the segmentation accuracy is even lower than 60%, it is difficult to meet the needs of practical applications.

Automatic driving requires high algorithm accuracy and speed, and it is difficult to apply the algorithm with unbalanced speed and accuracy to automatic driving. Therefore, if semantic segmentation is to be applied to automatic driving, it is necessary to design a network with both speed and precision. To improve the speed of network operation, lightweight network (such as MobileNetV2 [12]) can be used as the backbone network to extract features. This kind of network operation speed is very fast, even can run on the mobile terminal, but the acquired feature information is not rich enough, and the down-sampling will cause the loss of edge and detail features, resulting in reduced segmentation accuracy. By analyzing these problems, it is found that the high-level feature maps in the network have better semantic information, while the low-level feature maps are rich in details. Better segmentation results can be obtained by integrating feature maps of different scales. Based on this, a new void CNN method is designed in this paper. When the input size is 512×1024, the segmentation accuracy is improved to 67.3%, and the operating speed is only 17ms.

3. Proposed road scene segmentation model

The full convolutional neural network (FCN) model enables the convolutional neural network to carry out intensive pixel prediction without the full connection layer and generate image segmentation graphs of any size, with faster operation speed than image block classification. FCN can be based on several structures (AlextNet, VGG-Net, GoogLeNet, SIFT-flow, VGG-16, etc.), VGG-16 is widely regarded as the most effective one. However, FCN is adapted from the traditional CNN. CNN was originally designed as an artificial neural network for image classification. Semantic segmentation is a dense prediction problem, which is structurally different from image classification. Under the same computational conditions, dilated convolutional networks (DCN) can provide a larger receptive field and are often used in real-time image segmentation. Based on this, this paper combines the advantages of full convolution and dilated convolution to build a semantic segmentation model of field road scene image based on dilated convolution neural network.

3.1. Dilated convolution

Dilated convolution [13,14] is a way of sampling data on a feature graph. Receptive fields can be increased without loss of resolution or coverage. Receptive field refers to the area size mapped by pixels in the feature map output by each layer of the network on the original image. The calculation formula of receptive field r_{i+1}^2 is as follows:

$$r_{i+1}^2 = [(r_i - 1) + (2l + 1)]^2 \quad (1)$$

Where r_i represents the length of the perceived field side of the i-th layer, and l represents the expansion coefficient of the dilated convolution.

The size of convolution kernel of dilated convolution is the same as that of ordinary convolution, that is, the number of parameters is unchanged in neural network, but it has a larger receptive field.

The dilated convolution in two dimensions can be defined as follows:

$$(F \circ k)(p) = \sum_{s+lt=p} F(s)k(t) \quad (2)$$

Where $\circ$ is the dilated convolution. p is its domain. F is the input image. s is its domain. k is the kernel. t is its domain. Compared with ordinary convolution, the

condition of dilated convolution changes from $s+t=p$ to $s+lt=p$, that is, each convolution kernel only operates with the element in the multiple position of l in the image F. When $l=1$, ∘ is the ordinary discrete convolution operation.

The convolutional neural network reduces the image size by pooling layer down-sampling while increasing the receptive field, and then uses up-sampling to change the image back to its original size. In this process, image information is lost, but the problem can be avoided to some extent by dilated convolution. Figure 1 shows the relationship between the void convolution and the receptive field, which increases exponentially.
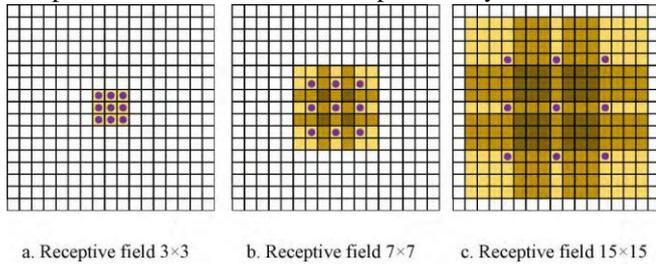

a. Receptive field 3×3   b. Receptive field 7×7   c. Receptive field 15×15

Figure 1. Expansion of receptive field due to dilated convolution

In Figure 1, the convolution kernels are all 3×3. Figure 1a adopts the dilated convolution of 11 (i.e., ordinary convolution) to operate on the original graph to obtain the feature graph of the first layer. The information represented by each element in the first layer is the information of element 3×3 in the original picture, that is, the receptive field is 3×3. Figure 1b adopts the cavity convolution of $l=2$ to operate the first layer to obtain the feature map of the second layer. Since the expansion coefficient is 2, the convolution kernel distribution is actually the position of the b in the figure. The receptive field of each element in the second layer is 7×7 relative to the original image. FIG. 1c adopts the cavity convolution of $l=4$ to operate on the second layer to obtain the feature map of the third layer. Similarly, the receptive field of each element in the third layer is 15×15.

Compared with the three-layer 3×3 convolution kernel of traditional convolution, only 7×7 receptive fields can be obtained. The number of factors actually participating in the convolution of dilated convolution has not changed. The computation amount of dilated convolution remains unchanged, but the size of the convolution kernel increases, so that one eigenvalue in the feature graph corresponds to the original larger region, that is, a larger range of feeling can be obtained.

3.2. Context aggregation and front-end module based on dilated convolution

In recent years, in the study of convolutional neural network, Yanagase et al. [15] analyzed the expansion of filter but did not apply it. Sun et al. [16] simplified the structure of full convolutional neural network with cavities. Yu et al. [17] proposed the context module of dilated convolution, which systematically used dilated convolution to carry out multi-scale context aggregation, aiming to improve the performance of pixel prediction architecture by aggregating context information. The input and output of the module are C channel feature maps (C can represent the number of object classification in the image), and the input and output forms are the same. This module can therefore be plugged into an existing pixel prediction network. However, it does not have complete predictive network function, and requires a front-end network to provide feature maps as input, namely front-end module.

1) Context module

The context module has 8 layers. For the first 7 layers, 3×3 convolution kernels with different expansion coefficients are used for cavity convolution. The expansion coefficient 1 increases exponentially in each layer. So the convolution kernel of the small sensing region is used to obtain local features first, and then the convolution kernel of the large sensing region is used to divide features into more regions. Each convolution operation is followed by the element-by-element truncation operation Max to clip the expanded edges caused by the cavity convolution. The last layer performs 1×1×C convolution and produces the output of the module. Context module can be divided into two network forms, Basic and Large, according to different channels of convolution. Convolutional neural networks are usually initialized with randomly distributed samples [18,19]. However, experiments show that the standard random initialization scheme cannot improve the prediction accuracy of the context model, and the alternative initialization form with clear semantics is more effective.

The initialization scheme of basic network is as follows:

$$k^b(t,a) = 1_{[t=0]}1_{[a=b]} \qquad (3)$$

Where a is the index of the input feature graph, and b is the index of the output feature graph. The initialization scheme sets up all filters to pass input from each layer directly to the next layer. Experiments show that the back-propagation method can reliably obtain the context information of the network and improve the accuracy of the processed feature map.

Large networks differ from Basic networks in the use of more feature maps in deeper layers. Large networks also need to change the initialization scheme to solve the

problem of the difference in the number of feature maps of different layers. The method is as follows: set c_i and c_{i+1} as the number of feature maps of two continuous layers, divide C by c_i and c_{i+1} at the same time, which can be expressed as:

$$k^b(t,a) = \begin{cases} C/c_{i+1}, t = 0 \\ \varepsilon, others \end{cases} \quad (4)$$

Where $\varepsilon \sim N(0,\sigma^2)$ and $\sigma << C/c_{i+1}$.

2) Front-end module

The front-end module is also called front-end Prediction Module. Its function is to generate a feature map of a certain resolution to provide to the context module. The front-end module is derived from the research of Long et al. and Chen et al. Long et al. retained the last two pooling and striding layers in the traditional classification network. Chen et al. used expansion to replace striding layer and retain pooling layer. Xu et al. found that the prediction accuracy could be improved by removing the last two pooling layers of VGG-16 network to simplify the network. This paper adopts the front-end module modified by Xu et al [20].

3.3. Semantic segmentation model based on dilated convolution

According to the characteristics of the cavity convolution network mentioned above, this paper uses the structure of full convolutional neural network VGG-16 to fuse the cavity convolution to construct a front-end module with higher prediction accuracy, and uses the cascade of the cavity convolution layers with different expansion coefficients to carry out multi-scale context aggregation. The semantic segmentation model of road scene image thus constructed is shown in figure 2. The part before the final layer in the figure is the front-end module and the part after it is the context module. The front-end module takes a color image as input and generates C=11 feature maps as output. The context module can further predict the feature graph output by the front-end module.

In order to simplify calculation and improve prediction accuracy, the front-end module is improved on the basis of VGG-16. The specific construction method is as follows: Pooling4 and Pooling5 layers in VGG-16 are removed, and three convolution layers in Convs are changed into dilated convolution with expansion coefficient of 2. The convolution of FC6 layer is changed to the dilated convolution with an expansion coefficient of 4 to keep the feeling field unchanged. In addition, the padding operation of VGG-16 intermediate feature map is used to conduct down-sampling with pooling layer, which is applicable to traditional classification network. However, noise may be introduced in the operation, which is unnecessary and unreasonable in pixel prediction. Therefore, the padding operation is deleted.

The context module constructed is the cascade of different expansion coefficient cavity convolution layers. The specific structural parameters of each layer are shown in Table 1. There are 8 layers including the final output layer. The first 6 layers are empty convolution with expansion coefficients of 1, 1, 2, 4, 8 and 16 respectively. As the resolution of the original image becomes 64×64 pixels after sampling under the front layer of the front-end module, the exponential expansion of the receptive field after the sixth layer is stopped in the context module design, and the receptive field of the seventh and eighth layers is 67×67. In order to facilitate comparison, two network forms, Basic and Large, are designed according to the different number of channels in the output characteristic graph.

These modifications enable parameter initialization using traditional VGG-16 networks and produce higher resolution output. The empty convolutional neural network is the semantic segmentation model of road scene image.

Table 1. Architecture parameters of context module network

Layer	Convolution	Dilation	Truncation	Receptive field	Output channels for feature maps	
					Basic network	Large network
1	3×3	1	Yes	3×3	C	2C
2	3×3	1	Yes	5×5	C	2C
3	3×3	2	Yes	9×9	C	4C
4	3×3	4	Yes	17×17	C	8C
5	3×3	8	Yes	33×33	C	16C

6	3×3	16	Yes	65×65	C	32C
7	3×3	1	Yes	67×67	C	32C
8	3×3	1	No	67×67	C	C

3.4. Model training

Convolutional Architecture for fast feature embedding based on deep learning framework [27] builds semantic segmentation model of road scene image based on empty convolution. CAFFE uses deploy.prototxt for the definition of the cavity convolution algorithm. solver.prototxt Sets training parameters. solve.py performs network training. infer.py calls the model to generate the results of semantic segmentation. In this paper, the test hardware environment is Intel Core i7-6700HQ@2.60GHz quad-core eight-thread processor, 16GB memory, 6GB video memory Nvidia GeForce GTX 1060 graphics card.

With the increasing number of layers, the accuracy of the DCN recognition model is also improving, but it also brings the problem that the model is prone to fall into the local minimum [21-24]. Therefore, in the actual deep network model training, some scholars generally adopt the parameters of the previous better convergence model to initialize the initial parameters of the new model. The existing convolutional neural network models such as SSD (Single Shot Multibox Detector) and DeepID adopt the strategy of pre-training.

In this paper, while pre-training initialization parameters are adopted, two-stage training [25] is adopted for the model. The specific steps are as follows:

1) Using the VGG-16 model parameters trained on ImageNet to initialize the DCNN network to be trained.

2) Manually select 500 simple images with obvious features and few object categories to train the model separately. After many tests, it is determined that the learning rate of weight parameters is 10^{-4}, mini-batch size is 14, momentum is 0.9, weight decay is 0.000 5. Stochastic gradient descent method is used for training, and model parameters are saved after model convergence. Because the image is simple, the model converges quickly.

3) Retrain the model saved in the previous step on all training sets, initialize the model with the parameters obtained in the previous step, reduce the learning rate to 10^{-5}, and update all network weights and parameters through training.

4. Experiments and analysis

Data set. Network training and experimentation are conducted on Cityscapes data sets. Cityscapes is a road scene segmentation dataset commonly used in the field of

semantic segmentation, including 5000 fine labeled images and 19998 coarse labeled images. Among them, 5000 fine annotation images are divided into 2975 training set images, 500 validation set images and 1525 test set images. In this paper, only fine annotated image data is used, not coarse annotated data. There are 30 original annotation categories in the data set. Similar to other papers, this paper selects 19 common categories (such as cars, pedestrians and buildings) for training and testing. The average intersection ratio (mIoU) is used as the measurement of segmentation accuracy [26-30], and the processing time and frame rate of each image are used as the measurement of segmentation speed.

Data enhancement. A similar approach to BiSeNet is adopted in data enhancement. In the process of training, images are normalized by means and variances, and then the images are randomly flipped horizontally and transformed by random scales (including 0.75, 1.0. 1.5, 1.75, 2.0). Finally, in the training process, the fixed size of the image is randomly cut out for training.

4.1. Comparison with different modules

First, the front-end module is tested. Only the front-end module constructed in this paper (hereinafter referred to as front-end) is used for training and testing to verify whether the modification of VGG-16 is effective.

Then test the combination of the context module and the front-end module, and insert the constructed context module of Basic and Large into the front-end module (hereinafter referred to as front-end+Basic and front-end+Large respectively). Set the learning rate to 10^{-5}, the number of iterations to 4000, and initialize the context module. Since the receptive field of the context network is 67×67, a buffer of width 33 is used to populate the input feature graph. Two-stage training method is used to train network models. In the testing process of each network model, the image reading operation is realized by calling Python third-party library functions. The results are shown in table 2.

Table 2. Module validation

Module	mIoU/%	Time/ms
front-end	88.2	18

front-end+Basic	92.3	20
front-end+Large	93.1	19

As can be seen from the results in Table 2, compared with the method of directly integrating feature maps at different levels, the segmentation accuracy of the front-end+Large method in this paper is improved to 93.1% by considering different characteristics of feature maps at both high and low levels, indicating that the proposed method can significantly improve the segmentation accuracy. Figure 2 is the result.
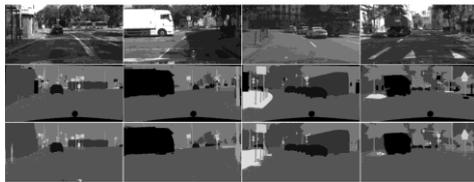

Figure 2. Segmentation result(from top to bottom, Input image, marked true value, predicted value)

4.2. Comparison with other networkds

SegNet and ENet are several commonly used semantic segmentation networks [17]. This paper compares the segmentation accuracy and speed of these networks. Table 3 shows the segmentation accuracy, running time and frame rate of different networks in Cityscapes data set 2048×1024 input resolution. It can be seen from the experimental results that the accuracy of the proposed method is higher and the real-time performance of the algorithm can be guaranteed at the same input size.

Table 3. Segmentation precision of several networks

Method	mIoU/%	Time/ms	Frame/s
SegNet	56.2	641	2
ENet	58.4	49	19
Proposed	72.5	30	32

The semantic segmentation results of the front-end+Large network model for the shaded road test set and the unshaded road test set are shown in Table 4, and the segmentation effect is shown in Figure 3.

Table 4. Test results of shadowed and unshadowed roads

Evaluation	Shadowed road	Unshaded road
MIoU/%	73.2	73.4
Time/s	21	19

Original image   Segmentation result   Original image   Segmentation image

Shadowed road   Unshadowed image

Figure 4. Front-end+Large network's image segmentation results for shadowed and unshadowed roads

5. Conclusion

In this paper, based on the different characteristics of high level and low level feature graphs, a network module is designed to effectively fuse different level feature graphs, which improves the problem that the fusion effect between feature graphs is not ideal in the traditional semantic segmentation problem. On the basis of this module, an efficient semantic segmentation network with both precision and speed is constructed. Different from many previous methods that directly improve the accuracy by increasing the complexity of the network or improve the speed by sacrificing the accuracy of the network, the cavity convolutional neural network achieves a good balance between the speed and accuracy. On the data set Cityscapes, it can achieve a segmentation accuracy of 72.5% on the basis of running time of 17 ms. Experimental results show that the proposed method can maintain both real-time performance and robustness, and can be further applied to practical scenarios such as intelligent driving. In the future, we will continue to optimize the segmentation results on this basis, considering the extraction of details from low-level features into high-level feature images.

References

[1] Liu T, Yin S. An improved particle swarm optimization algorithm used for BP neural network and multimedia course-ware evaluation[J]. Multimedia Tools & Applications, 76(9):11961-11974, 2017.

[2] 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.

[3] 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.

[4] Wang S, Sun G, Zheng B, et al. A Crop Image Segmentation and Extraction Algorithm Based on Mask RCNN[J]. Entropy, 2021, 23.

[5] Huang C, Wang J. Research on ARKFCM Algorithm Based on Membership Constraint and Bias Field Correction in Neonatal HIE Image Segmentation Method[J]. Mathematical Problems in Engineering, 2021, 2021(99):1-11.

[6] Han X, Zhong Y, Cao L, et al. Pre-Trained AlexNet Architecture with Pyramid Pooling and Supervision for High Spatial Resolution Remote Sensing Image Scene Classification[J]. Remote Sensing, 2017, 9(8):848.

[7] Filipe M, Rafael M, Alberto B. Retinal Vessel Segmentation based on Fully Convolutional Neural Networks[J]. Expert Systems with Applications, 2018, 112:S0957417418303816-.

[8] Hl A, Sw A, Yc A, et al. Semantic segmentation of 3D indoor LiDAR point clouds through feature pyramid architecture search[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2021, 177:279-290.

[9] 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

[10] Yang Sun, Yin Shoulin, Teng Lin, Liu Jie. Study on wavelet transform adjustment method with enhancement of color image[J]. Journal of Information Hiding and Multimedia Signal Processing, v 9, p 606-614, May 2018.

[11] Jin Z, Sw B, Jq C, et al. A fast X-shaped foreground segmentation network with CompactASPP[J]. Engineering Applications of Artificial Intelligence, 97.

[12] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. -C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 4510-4520, doi: 10.1109/CVPR.2018.00474.

[13] Cheng H, Ma S, Lee H, et al. Millimeter Wave Path Loss Modeling for 5G Communications Using Deep Learning with Dilated Convolution and Attention[J]. IEEE Access, 2021, PP(99):1-1.

[14] Zeng J, Zhang D, Li Z, et al. Semi-Supervised Training of Transformer and Causal Dilated Convolution Network with Applications to Topic Classification[J]. Applied Sciences, 2021, 11(12):5712.

[15] Yanagase Y, Suzuki S, Kokubun Y, et al. Box-like filter response and expansion of FSR by a vertically triple coupled microring resonator filter[J]. Lightwave Technology Journal of, 2002, 20(8):1525-1529.

[16] Sun S Y, Li Z, Li J, et al. A memristor-based convolutional neural network with full parallelization architecture[J]. IEICE Electronics Express, 2019.

[17] 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.

[18] 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. DOI: 10.1177/1550147719852036

[19] Shoulin Yin, Lei Meng and Jie Liu. A New Apple Segmentation and Recognition Method Based on Modified Fuzzy C-means and Hough Transform[J]. Journal of Applied Science and Engineering. Vol. 22, No. 2, pp. 349-354, 2019.

[20] Fu X, Fang B, Zhou M, et al. Active Contour Driven by Adaptively Weighted Signed Pressure Force Combined with Legendre Polynomial for Image Segmentation[J]. Information Sciences, 2021.

[21] J. Wang, H. Li, S. Yin and Y. Sun, "Research on Improved Pedestrian Detection Algorithm Based on Convolutional Neural Network," 2019 International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), 2019, pp. 254-258, doi: 10.1109/iThings/GreenCom/CPSCom/SmartData.2019.00063.

[22] Shoulin Yin, Jing Bi. Medical Image Annotation Based on Deep Transfer Learning[J]. Journal of Applied Science and Engineering. Vol. 22, No. 2, pp. 385-390, 2019.

[23] 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.

[24] Jing Yu, Hang Li, Shoulin Yin. Dynamic Gesture Recognition Based on Deep Learning in Human-to-Computer Interfaces [J]. Journal of Applied Science and Engineering, vol. 23, no. 1, pp.31-38, 2020.

[25] 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

[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] 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

[29] 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.

[30] 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).