

# Hazy Images Segmentation Method Based on Improved DeeplabV3+

Dong Wenkuan<sup>a</sup>, Gong Shicai\*

School of Science, Zhejiang University of Science and Technology, Hangzhou, Zhejiang, 310000, China

<sup>a</sup>846064241@qq.com

\*Corresponding author

**Abstract:** To address the issue of decreased segmentation accuracy in foggy images and the difficulty in accurately segmenting heavy fog regions in traditional semantic segmentation models, we propose an improved segmentation model that combines image Levels adjustment and attention mechanisms. Firstly, the fog density of the image is estimated using an atmospheric scattering model, and the image is adjusted based on the estimated fog density to highlight the information that was originally obscured by the fog. A dual branch input (DB Input), is constructed for both heavy fog and light fog areas to enhance the feature learning of the model without destroying the initial information of the foggy image. An RCCA module, which is a spatial domain attention mechanism, is introduced at the end of the dual branch input to enhance the region attention ability of the module in different branches. Experiments are conducted on the datasets Foggy Cityscapes and Foggy Uavid. The results show that the improved model achieves 70.6% and 66.7% in the mIOU accuracy, respectively, which is a 5.2% and 5.4% improvement over the original model, indicating better segmentation results.

**Keywords:** fog density estimate; levels adaptive adjustment; attention mechanism; semantic segmentation; hazy images

## 1. Introduction

Semantic segmentation as an image pixel-level classification task has attracted extensive interest in academia and industries. It has shown great potential in many fields such as autonomous driving and drone monitoring<sup>[1, 2, 3]</sup>. Due to many reasons including air water vapor condensation and haze pollution, the image fog problem is widely present in visual scenes. With the intensive application of neural networks in the industry, it has gradually become a realistic requirement for various types of deep networks to operate properly in foggy scenes<sup>[4, 5]</sup>.

The proposal of FCN<sup>[6]</sup> enabled the segmentation accuracy of deep convolutional models to be greatly improved compared to traditional methods. Since then, deep models based on different improved ideas have emerged; Olaf Ronneberger<sup>[7]</sup> proposed the U-Net model with a symmetric encoder-decoder structure. U-Net passes low-level features into higher semantic information to recover the contours of the details; SegNet<sup>[8]</sup> optimizes the pooling process, which preserves the pooling information of the encoder and optimizes the process of the decoder; DeeplabV3+<sup>[9]</sup> improves the convolution structure by using a depth-separable convolution with the Atrous Spatial Pyramid Pooling (ASPP), which effectively extracts the multiscale information of the feature map.

For the picture defogging problem, researchers have conducted defogging in terms of physical models<sup>[10, 11]</sup>, nonphysical models<sup>[12, 13]</sup>, and deep learning models<sup>[14]</sup>. For example, Kaiming He invented the dark channel prior theory, which has become one of the current standard defogging algorithms due to its stability and efficiency<sup>[15]</sup>; Histogram equalization series of defogging methods achieve the effect of dehaze by adjusting the image gray value distribution for image enhancement<sup>[16, 17]</sup>; Cai et al. proposed DehazeNet<sup>[18]</sup> to obtain clear images based on the physical model of the atmosphere and use deep learning to estimate the haze parameters in the atmosphere.

Despite the development of both defogging and segmentation algorithms, the semantic segmentation task under fog graphs still needs further integration. Dai et al.<sup>[19]</sup> found that the segmentation of the light foggy images can guide the segmentation process of the heavy fog environments. Based on this idea, he proposed CMAda. However, the training of this model requires a

constraint on the fog concentration, and it needs to introduce more hyperparameters and higher computational effort; Lee<sup>[20]</sup> et al. consider fog as a special image style and learn the feature invariance of fog by constructing a depth model and a fog filtering channel. However, this model relies on three types of inputs, clear weather (CW), synthetic fog (SF), and real fog (RF), in the same scene, which has high requirements on the input dataset.

In order not to destroy the original information of the scene and not to impose additional constraints on the input data, an improved segmentation method combining image Levels adjustment based on fog concentration estimation and attention mechanism is proposed in this paper.

## 2. Related Work

Deep learning models have a certain learning ability in thin fog areas, but this learning ability needs to be enhanced in areas with heavy fog. Figure 1 shows the segmentation results of the basic DeeplabV3+ model in clear and foggy images. As shown in the rectangle area, the model can still perform segmentation with the same performance even though the image contains fog. However, in heavily foggy areas such as the circular area, the model struggles to learn the features of the region due to the thick fog. Therefore, this paper combines image processing algorithms to process the thick fog areas and feed them into the deep learning model for learning.

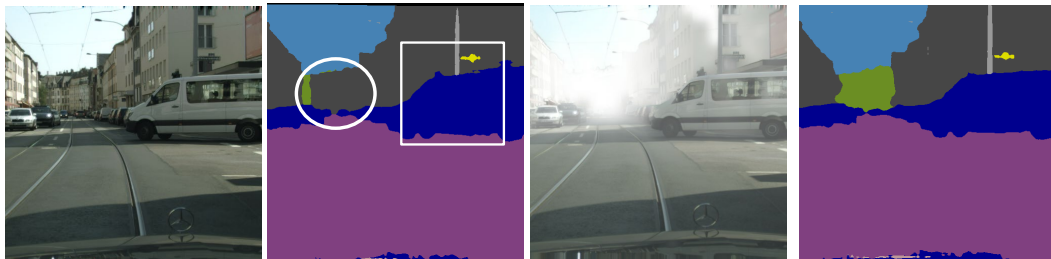


Figure 1: DeeplabV3+ segmentation effect in foggy images and clear images

### 2.1 Network Architecture

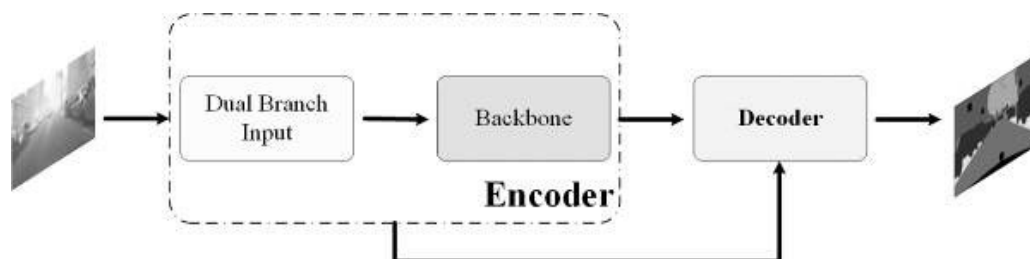


Figure 2: General structure of the model

This article is based on DeeplabV3+ and utilizes a structure consisting of the encoder and decoder. As shown in Figure 2, the encoder is composed of two parts: (1) a dual branch input module (DB Input) for foggy and hazy areas, which extracts the heavy fog features of the image and improved segmentation performance, and (2) a backbone module that employs the pre-trained Res101 structure to further learn and extract the fused information of the dual branch features.

The decoder module has two input sources: the complex semantics learned by the Backbone module and the detailed contour information merged by the DB Input module. The information from both sources is learned through residual modules.

### 2.2 Dual Branch Input

To enhance the transferability of the processing procedure across various datasets and segmentation models, an adaptive processing approach is proposed, where the algorithm autonomously adjusts the input image in different foggy conditions. Moreover, the input module constructed in this paper can be easily transferred to different segmentation models with only minor adjustments to the number of input module channels.

### 2.2.1 Estimation of Fog Concentration

Firstly, it is necessary to analyze the degree of haze in the input image. The atmospheric scattering model constructs a physical imaging theory that comprehensively considers the attenuation of incident light and atmospheric diffuse reflection based on the imaging principle, as shown below:

$$I(x) = J(x)t(x) + A(1 - t(x)) \quad (1)$$

$$t(x) = e^{-\beta d(x)}, t(x) \in [0, 1] \quad (2)$$

In the equation,  $I(x)$  and  $J(x)$  respectively represent the hazy image and clear image.  $A$  is the atmospheric light value that reflects the lighting conditions, and  $t(x)$  is the medium transmittance map, which is influenced by the attenuation coefficient  $\beta$  and the distance function  $d(x)$ . The larger the  $t$  value is, the less the image is disturbed. The difference in transmittance values between the hazy and haze-free images can reflect the degree of haze in the image. After extensive observation, He et al. found that in haze-free images, there always exists a channel with a grayscale value approaching 0. Estimation of the transmittance can be obtained based on this theory.

Figure 3 shows a comparison of transmittance between hazy and clear images. In the clear image, only the sky area has a transmittance approaching zero due to the distance function value being close to infinity. However, in the hazy scene, due to the presence of dense fog, the proportion of pixels with transmittance values approaching zero is increased. Therefore, selecting the proportion of low-transmittance pixels in the input image is used as an indicator of the hazy condition of the image. In this article, we select the proportion of pixels with transmittance values less than 0.1 as an estimate of the hazy condition.



Figure 3: Comparison of transmittance between hazy and clear images. (a) clear image (b) Transmittance of the clear image. (c) hazy image. (d) Transmittance of the hazy image

### 2.2.2 Levels Adaptive Adjustment

The essence of Levels adjustment is to perform gamma correction on the image, reducing the probability of high-gray pixels appearing while increasing the probability of low-gray pixels appearing during the adjustment process. At this time, reducing the loss of image information caused by heavy fog. The specific transformation process is as follows:

$$L_{out} = 255 \left( \frac{L_{in}}{255} \right)^{\frac{1}{m}} m \in [0, 10] \quad (3)$$

where  $L_{in}$  and  $L_{out}$  denote the input and output gray level. And  $m$  indicates the midtone, with the default value of 1. When the midtone is lowered, the highlight area information is enhanced to display, on the contrary, the shadow part display is enhanced. The update formula of  $m$  in Levels adjustment is:

$$m = 1 - \log_{256} \left( \frac{128}{L_m} \right) \quad (4)$$

where  $L_m$  is the grayscale value corresponding to  $m$ , and the initial value of  $L_m$  is 128.

The percentage of pixels below 0.1 in the transmittance map is selected and recorded as  $\lambda$ . According to the experiment, pixels with a transmission rate lower than 0.1 often correspond to the sky or heavy fog areas in the images. Larger  $\lambda$  values result in a higher percentage of low transmittance pixels, implying that smaller  $m$  values are needed to recover the masked information. The sum of gray

value frequencies of the image is calculated from high to low, and when the cumulative gray value frequency breaks through the threshold value determined by  $\lambda$ , the  $L_m$  value in Equation (4) is updated, and the adjusted image is obtained according to the requested  $m$  value.

Figure 4 shows the comparison of images before and after Algorithm 1 adjustment. The (a) and (c) parts in the figure are the overall images before and after processing by Algorithm 1, and the (b) and (d) parts are the corresponding local images. It can be seen from the figure that after adaptive color adjustment, the information in the foggy areas is restored, and the details and contours of objects hidden in the fog are displayed. By inputting the adjusted image into a deep learning model, it can help the model reduce the difficulty of obtaining feature information in foggy areas.

Algorithm 1: Levels adaptive adjustment

**Begin:**

**Input:**  $Img_{in}$

**Initialize:** gray value frequency of RGB channels:  $F_R = F_G = F_B = [0, 0, \dots, 0]_{1 \times 256}$ . RGB channels value:  $C_R, C_G, C_B, m=1$ . Input image size is  $h$  and  $w$ . Calculate the RGB mean value  $A$  of the first 0.1% pixel points of the image grayscale value, and calculate the transmittance  $T$ .

**for**  $L = 1$  to 255 **do**

Calculate the number of pixels whose  $C_R, C_G,$  and  $C_B$  grayscale values are equal to  $L$ , noted as  $C_{RL}, C_{GL}, C_{BL}$ . Assign the  $C_{RL}, C_{GL}, C_{BL}$  values to the vector of  $F_R, F_G, F_B$  with index  $L$

**end for**

$F_{mean} = (F_R + F_G + F_B) / (3 \times H \times W)$

$\lambda = \text{sum}([T < 0.1])$

Threshold =  $\text{sum}(F_{mean}) \times \lambda$

**for**  $k = 255$  to 0 **do**

**if**  $\text{sum}(F_{mean}[k:255]) > \text{Threshold}$  **do**

Update the  $m$  value according to Equation (4)

**end if**

**end for**

Bring  $m$  into equation (3) to obtain the adjusted image  $Img_{out}$

**Output:**  $Img_{out}$

**End**



Figure 4: Impact of image Levels adjustment algorithm. (a) original foggy image (b) original heavy fog area image (c) image after Levels adjustment (d) heavy fog area after Levels adjustment

### 2.2.3 Input Structure

A dual branch input module for processing dense and thin fog, as shown in Figure 5, is constructed, and the images before and after adjustment by Algorithm 1 are passed to the dense and thin fog learning branches of the module. Firstly, high-view information extraction is performed by a  $7 \times 7$  convolution, and then feature dimensionality reduction is achieved by using a pooling layer. The dense fog branch extracts multi-scale integrated information through Avg Pooling and Max Pooling, and the subsequent blocks are residual modules with a channel number of 64 to guide the model to learn residual information.

A spatial domain attention mechanism, RCCA<sup>[21]</sup>, is introduced at the end of the two branches. By establishing the correlation relationship among all pixels, RCCA guides the model to focus on more critical information in complex input features, and improve the ability of the model to extract contextual information. Experiments have shown that the number of RCCA iterations  $R$  affects the ability to extract contextual information. In this paper, based on experiments,  $R=2$  is finally adopted.

In the original DeeplabV3+ model, the concat step whether it is in the upsampling process or ASPP module, the channel integration is directly performed through  $1 \times 1$  convolution, without more detailed processing. However, this article believes that features from different sources have different levels of importance. Therefore, after all the concat steps in the article, an SE Module<sup>[22]</sup> is introduced to help the

model learn the importance of different channels.

As shown in Figure 6, the SE module consists of three parts: Squeeze, Excitation, and Reweight. It takes a feature vector  $X$  with size  $H \times W$  and channel number  $C$  as input, compresses each 2D channel into a scalar that has a certain degree of the global receptive field, generates the weight for each feature channel through excitation operation, and applies 2 fully connected layers, one ReLU activation layer, and one sigmoid layer in this paper. Finally, the input layer is multiplied by the weight layer in a reweighting manner to recalibrate the weights of the input features.

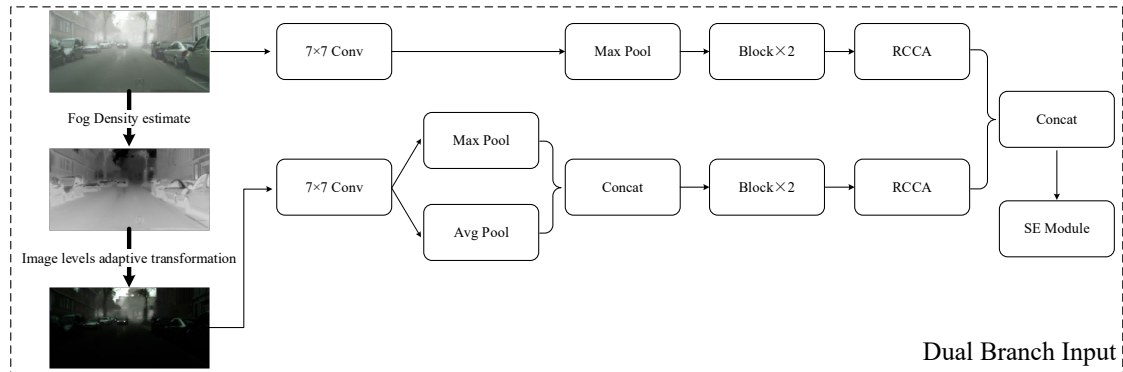


Figure 5: Dual Branch Input

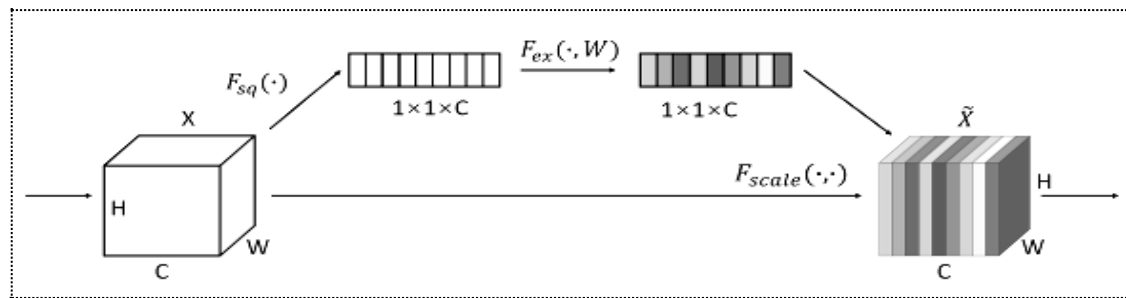


Figure 6: SE module

### 3. Training

#### 3.1 Datasets

On the basis of the Cityscapes dataset, Foggy Cityscapes<sup>[23,24]</sup> employs a simulated fog algorithm to synthesize high-quality foggy images. The dataset consists of two parts: weakly annotated and finely annotated images. For this experiment, we selected finely annotated images that contain three visibility versions, each with three constant attenuation coefficients. The attenuation coefficients are 0.005, 0.01, and 0.02, corresponding to visibility distances of 600, 300, and 150 meters, respectively.

In order to test the segmentation performance of the model under different conditions, this paper simulates different concentrations of haze images on the basis of the Uavid<sup>[25]</sup> dataset using the fog simulation algorithm in equation (1) and (2). Figure 7 shows some examples of the generated images.

Uavid is a remote sensing dataset of urban street scenes obtained by unmanned aerial vehicles at 4k resolution. Each image is used to generate three hazy images with different haze concentration parameters  $\beta=0.005$ ,  $\beta=0.01$ , and  $\beta=0.02$ .

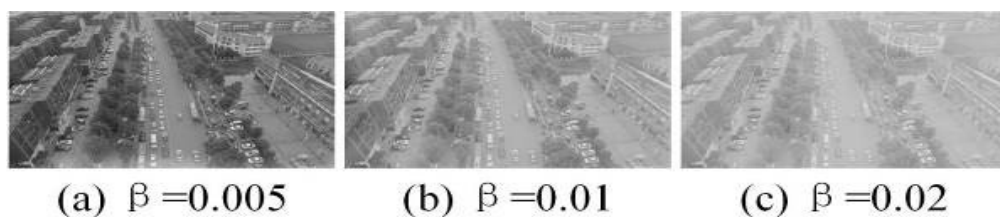


Figure 7: Foggy Uavid training set

### 3.2 Experiment

The experimental parameter settings in this article are shown in Table 1:

Table 1: Hyperparameters

Batch size	Output stride	Image size	Weight decay	Min learning rate	Max learning rate	Momentum	Iteration
12	8	769×769	1×10 <sup>-4</sup>	0.00001	0.1	0.9	50000

The experiment was conducted using the pytorch framework, with an NVIDIA A40 GPU, and a pre-trained Resnet101 was selected as the backbone. Image augmentation techniques such as random horizontal flipping and random Gaussian perturbation were used. Focal Loss was chosen as the loss function to address the class imbalance of the Foggy Cityscape and Foggy Uavid samples.

The evaluation metrics used were mean intersection over union (mIOU):

$$R_{mIOU} = \frac{1}{1+n} \sum_{i=0}^n \frac{p_{ij}}{\sum_{j=0}^n (p_{ij} + p_{ji}) + p_{ii}} \quad (5)$$

where  $n$  denotes the number of categories,  $p_{ij}$  denotes the number of pixels which the true category is  $i$  and the predicted category is  $j$ .

### 3.3 Contrast and Ablation Experiments

Table 2: Comparison of segmentation accuracy of different models

Algorithm	Foggy Cityscapes mIOU /%	Foggy Uavid mIOU/%
DeeplabV3+[9]	65.4	61.3
DDRNet[26]	65.5	60.4
MFANet[27]	66.8	61.4
FISS GAN[28]	66.4	61.7
Proposed algorithm	<b>70.6</b>	<b>66.7</b>

To compare the segmentation accuracy of different models on the Foggy Cityscapes and Foggy Uavid datasets, DDRNet, FISS GAN, MFANet, and the basic DeeplabV3+ were selected for comparative experiments. Under the same experimental conditions, each model was trained on the training set of the dataset and the results were output on the validation set. As shown in Table 2, the proposed model achieved mIOU accuracies of 70.6% and 66.7% in Foggy Cityscapes and Foggy Uavid, respectively, which are 5.2% and 5.4% higher than the original model. Compared with DDRNet, FISS GAN, and MFANet, the proposed model has better segmentation performance.

The model was trained under fog-free conditions with the same parameter settings to test the generalization performance of the model, and the comparison results are shown in the following table:

Table 3: Comparison of different models in Cityscapes and Uavid segmentation accuracy

Algorithm	Cityscapes mIOU /%	Uavid mIOU/%
DeeplabV3[9]	79.1	68.0
DDRNet[26]	79.5	63.0
MFANet[27]	73.9	63.8
FISS GAN[28]	75.8	64.2
Proposed algorithm	<b>79.9</b>	<b>68.2</b>

Table 3 shows the segmentation accuracy of the different models on the fog-free datasets Cityscapes and Uavid. Compared with the original model, the improved models both have some accuracy improvement, with increases of 0.8% and 0.2% respectively. It shows that the model with Dual Branch Input has a better improvement in the foggy scenes, and the introduction of dual-branch input does not reduce the performance of the model in clear haze-free scenes.

Table 4: Inference strategy on the Foggy Cityscapes and Foggy Uavid

FDE	LAA	SE Module	Foggy Cityscapes mIOU /%	Foggy Uavid mIOU /%
			65.4	61.3
	√		68.7	63.4
		√	66.1	61.9
√	√		70.2	65.5
√	√	√	70.6	66.7

To further verify the practical effects of the input module and the attention mechanism module, ablation experiments were conducted. The results are shown in Table 4, where FDE denotes Fog Density Estimate, LAA denotes Levels adaptive adjustment, and LA denotes fixing the Levels adjustment to a certain determined value. It means in Algorithm (1), a fixed value of  $\lambda$  is taken. In the experiment, the mean values obtained from the training sets of the datasets were approximately 0.1. As can be seen from the table, the introduction of Levels adjustment helps the model improve segmentation accuracy significantly, and the adjustment based on fog density makes the segmentation effect even more pronounced. The addition of the channel domain attention also makes the model have partial effect improvement.

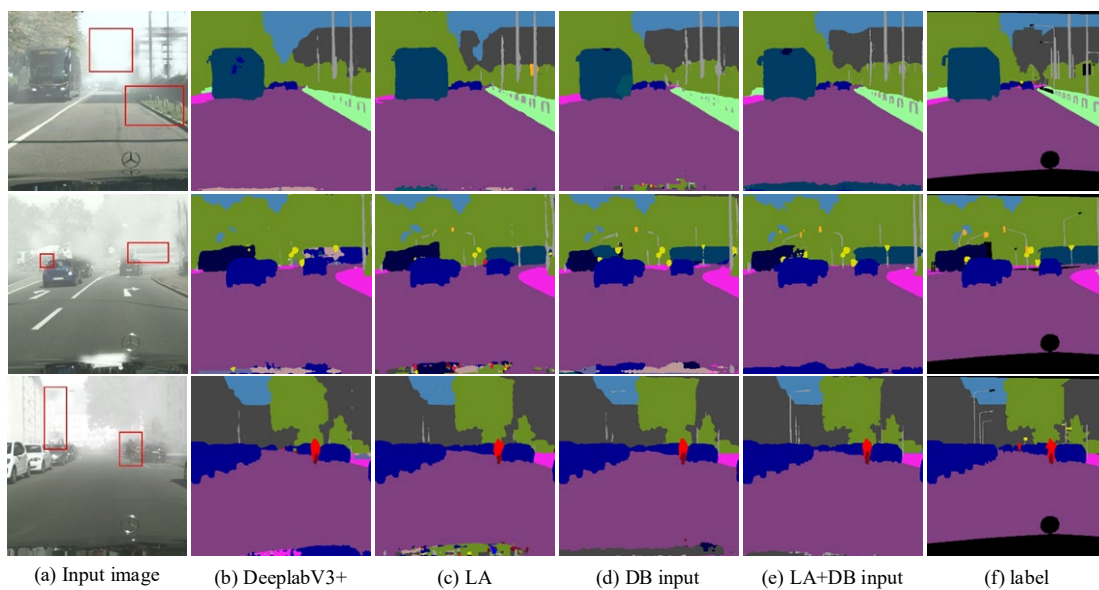


Figure 8: Comparison of segmentation results for DeeplabV3+ with different improvement methods

Figure 8 shows the segmentation results of different improvement methods. As can be seen from the figure, the introduction of the DB Input allows the model to have finer segmentation in regions containing heavy fog such as the sky and the distant scenes. For example, in the first row of the image in Figure 8, the original model identifies the distant building as the sky due to dense fog, while the improved model accurately identified the building contours. Introducing the SE Attention helps the model perform better in local details such as people and traffic signs. The improved model constructed by combining the above two has a higher segmentation accuracy and obtains more accurate prediction results in the highlight areas and detailed contours of the image.

#### 4. Summary

This paper proposes an improved DeeplabV3+ segmentation model that combines image Levels adaptive adjustment and attention mechanisms to address the problem of decreased segmentation accuracy in foggy scenes and insufficient segmentation details in dense fog areas. Firstly, the fog concentration of the image is estimated using the dark channel prior theory, and the image Levels is adjusted according to the estimated value to restore rich detail information in the heavily foggy areas. The Dual Branch Input module, which incorporates the spatial domain attention mechanism RCCA, is constructed to guide the model to enhance its attention and learning abilities in thin and dense obscure regions. The proposed model achieves more accurate segmentation and improves mIOU accuracy by 5.2% and 5.4% respectively on the Foggy Cityscapes and Foggy Uavid datasets, demonstrating the effectiveness of the proposed method. In future work, the evaluation of the global fog concentration of

the image could be combined with lightweight neural networks, or lightweight deep learning networks could be used to recover fog details in foggy images, allowing the model to better extract image feature information in subsequent segmentation tasks.

## References

- [1] Grigorescu S, Trasnea B, Cocias T, et al. A survey of deep learning techniques for autonomous driving [J]. *Journal of Field Robotics*, 2020, 37(3): 362-386.
- [2] Feng D, Haase S C, Rosenbaum L, et al. Deep multi-modal object detection and semantic segmentation for autonomous driving: Datasets, methods, and challenges [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2020, 22(3): 1341-1360.
- [3] Fu K, Zhu X Y, Lv Q X, et al. UAV Command Intent Identification Technology Based on Deep Learning [J]. *Ordnance Industry Automation*, 2022, 41(10): 41-44+59.
- [4] Dong S, Wang P, Abbas K. A survey on deep learning and its applications[J]. *Computer Science Review*, 2021, 40: 100379.
- [5] Lee T, Mckeever S, Courtney J. Flying free: a research overview of deep learning in drone navigation autonomy [J]. *Drones*, 2021, 5(2): 52.
- [6] Long J, Evan S, Trevor D. Fully convolutional networks for semantic segmentation[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2015: 3431-3440.
- [7] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation [C]//2015 International Conference on Medical Image Computing and Computer-assisted Intervention. Heidelberg: Springer, 2015: 234-241
- [8] Vijay B, Alex K, Roberto C. SegNet: A deep convolutional encoder-decoder architecture for image segmentation [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(12): 2481-249512.
- [9] Chen L C, Zhu Y, Papandreou G, et al. Encoder-Decoder with atrous separable convolution for semantic image segmentation[C]//2018 European Conference on Computer Vision. Heidelberg: Springer, 2018: 801-818.
- [10] Agrawal, Subhash C, Anand S J. A comprehensive review on analysis and implementation of recent image dehazing methods [J]. *Archives of Computational Methods in Engineering*, 2022, 29(7): 4799-4850.
- [11] Zhou J C, Liu D S, Xie X, et al. Underwater image restoration by red channel compensation and underwater median dark channel prior [J]. *Applied Optics*, 2022, 61(10): 2915-2922.
- [12] Hassan H, Bashir A K, Ahmad M, et al. Real-time image dehazing by super pixels segmentation and guidance filter[J]. *Journal of Real-Time Image Processing*, 2021, 18: 1555-1575.
- [13] Mondal K, Rabidas R, Dasgupta R. Single image haze removal using contrast limited adaptive histogram equalization based multiscale fusion technique [J]. *Multimedia Tools and Applications*, 2022, 19: 1-26.
- [14] Manu C M. MSDNet: a novel multi-stage progressive image dehazing network[C]//2021 Proceedings of the Twelfth Indian Conference on Computer Vision, Indian: Graphics and Image Processing, 2021: 1-9
- [15] He K M, Zhang X, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2016:770-778.
- [16] Dhal K G, Das A, Ray S, et al. Histogram equalization variants as optimization problems: a review[J]. *Archives of Computational Methods in Engineering*, 2021, 28: 1471-1496.
- [17] Rao B S. Dynamic histogram equalization for contrast enhancement for digital images[J]. *Applied Soft Computing*, 2020, (89): 106114.
- [18] Cai B, Xu X, Jia K, et al. DehazeNet: An end-to-end system for single Image haze removal[J]. *IEEE Transactions on Image Processing*, 2016, 25(11): 5187-5198.
- [19] Dai D, Sakaridis C, Hecker S, et al. Curriculum model adaptation with synthetic and real data for semantic foggy scene understanding[J]. *International Journal of Computer Vision*, 2020, 128(5): 1182-1204.
- [20] Lee S, Son T, Ewak S. Fifo: Learning fog-invariant features for foggy scene segmentation. [C]//2022 Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Louisiana: IEEE Press, 2022: 18911-18921.
- [21] Huang Z L, Wang X G, Huang L C, et al. Ccnet: Criss-cross attention for semantic segmentation [C]// 2019 IEEE/CVF International Conference on Computer Vision. New York: IEEE Press, 2019: 603-612.
- [22] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//2018 IEEE Conference on Computer



*Vision and Pattern Recognition. New York: IEEE Press, 2018: 7732-7741.*

[23] Cordts M, Omran M, Ramos S, et al. *The cityscapes dataset for semantic urban scene understanding[C]//2016 Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE Press, 2016: 3312-3323.*

[24] Sakaridis C, Dai D, Van G L. *Semantic foggy scene understanding with synthetic data[J]. International Journal of Computer Vision, 2018, 126(9): 973-992.*

[25] Lyu Y, Vosselman G, Xia G S, et al. *UAVid: A semantic segmentation dataset for UAV imagery[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 165: 108-119.*

[26] Hong Y, Pan H, Sun W, et al. *Deep dual-resolution networks for real-time and accurate semantic segmentation of road scenes [J]. arXiv preprint arXiv:2101.06085, 2021.*

[27] Ren D D, Li J B, Zhao J Y. *MFANet: multi-task fusion attention network for semantic segmentation of haze image [J]. Journal of natural science of Heilongjiang university, 2021, 38(5):608-616.*

[28] Liu K, Ye Z, Guo H, et al. *FISS GAN: A generative adversarial network for foggy image semantic segmentation [J]. IEEE/CAA Journal of Automatica Sinica, 2021, 8(8): 1428-1439.*