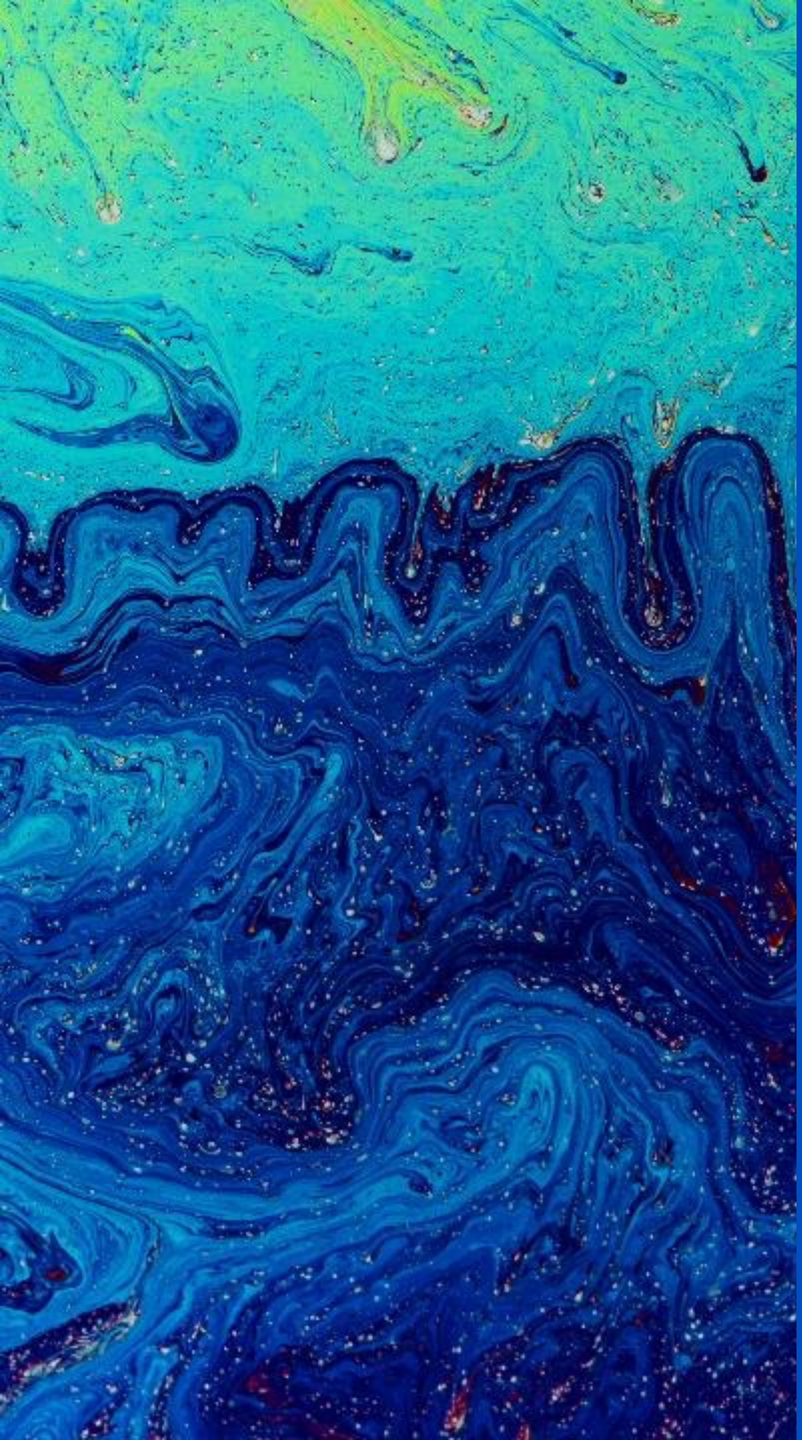
ICPR 2024 INDIA
27TH International Conference on Pattern Recognition
December 01-05, 2024, Kolkata, India

DD-Net: Dynamic Network Architecture for Optimized Curve Segmentation and Reduce Computational Redundancy



Main content

01 Background

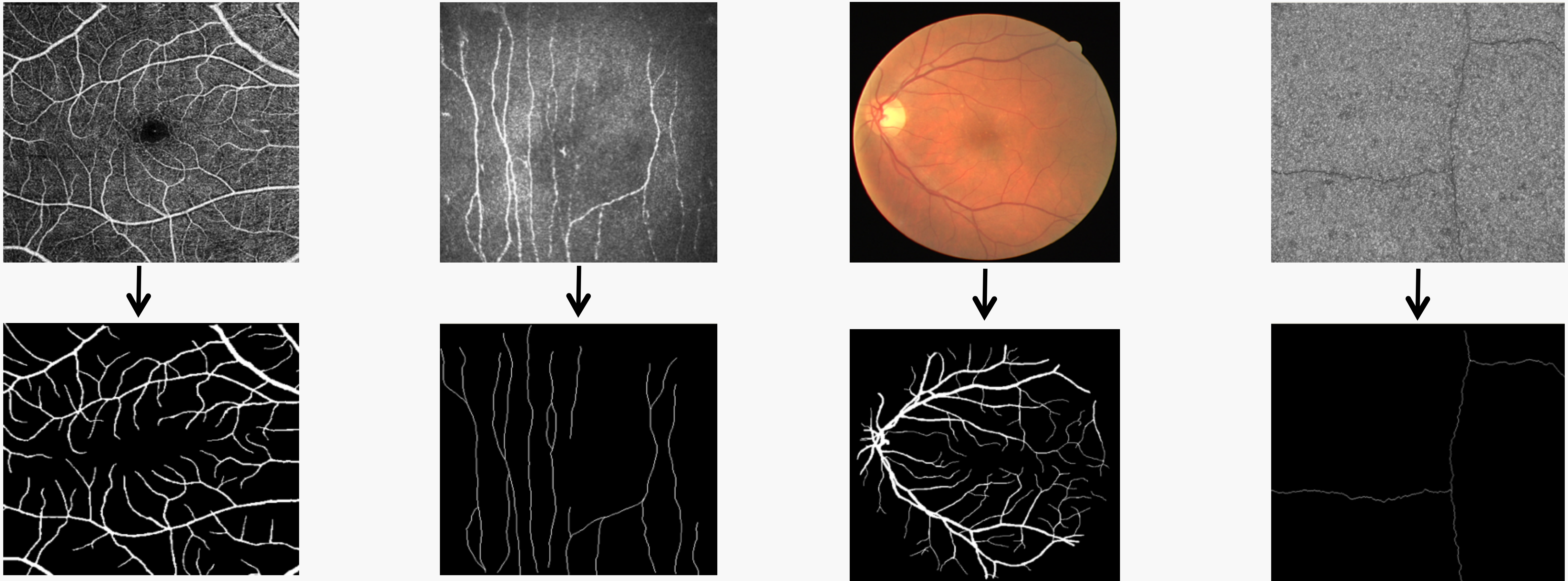
02 Proposed method

03 Experimental Results

04 Summarize

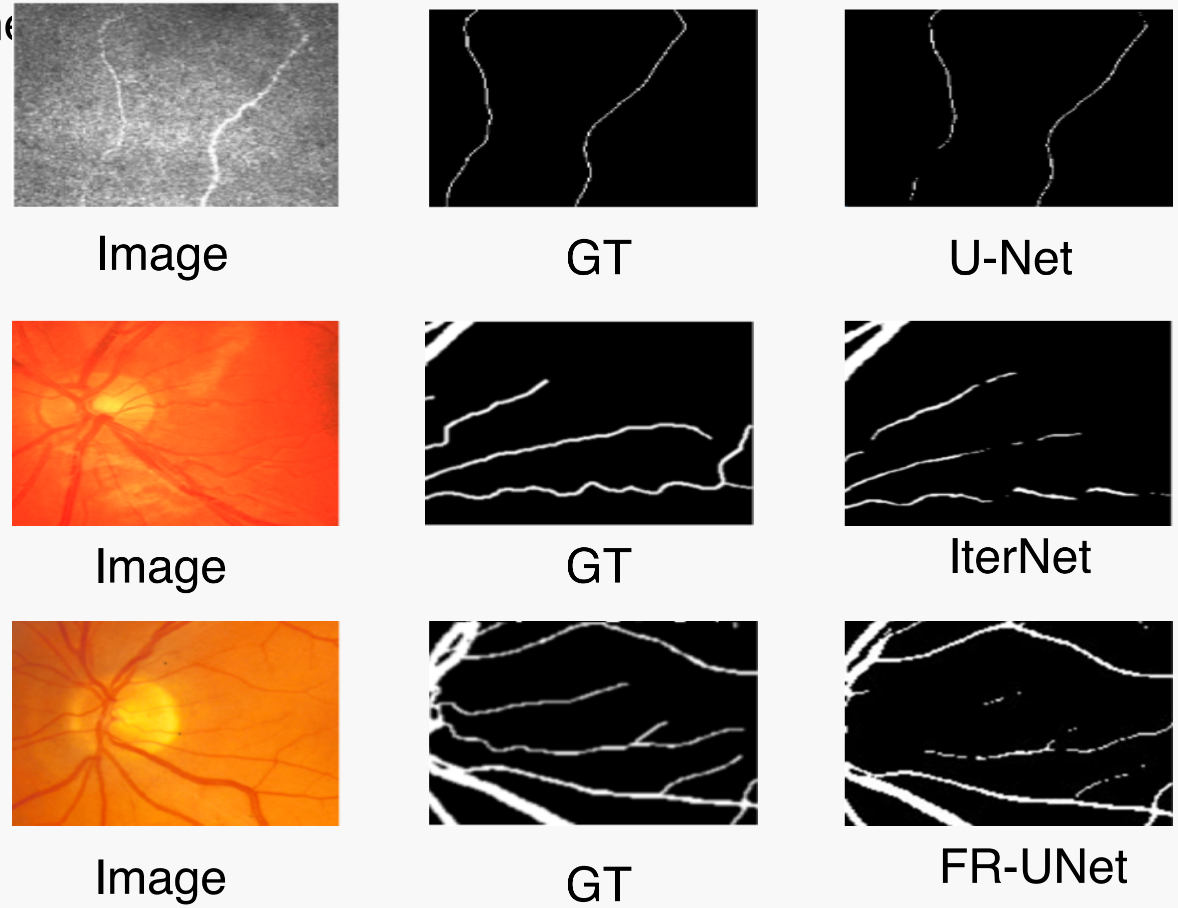
01 Background

Curve structure refers to a set of connected lines or curves with a certain width and morphological continuity in the image. Curve segmentation is to separate the curve from the background or other elements. In medical images, incorrect curve segmentation may affect the location and diagnosis of the lesion area. In the segmentation of road or wall cracks, incorrect segmentation may cause non-crack areas to be misjudged as cracks, or the real crack area to be missed.

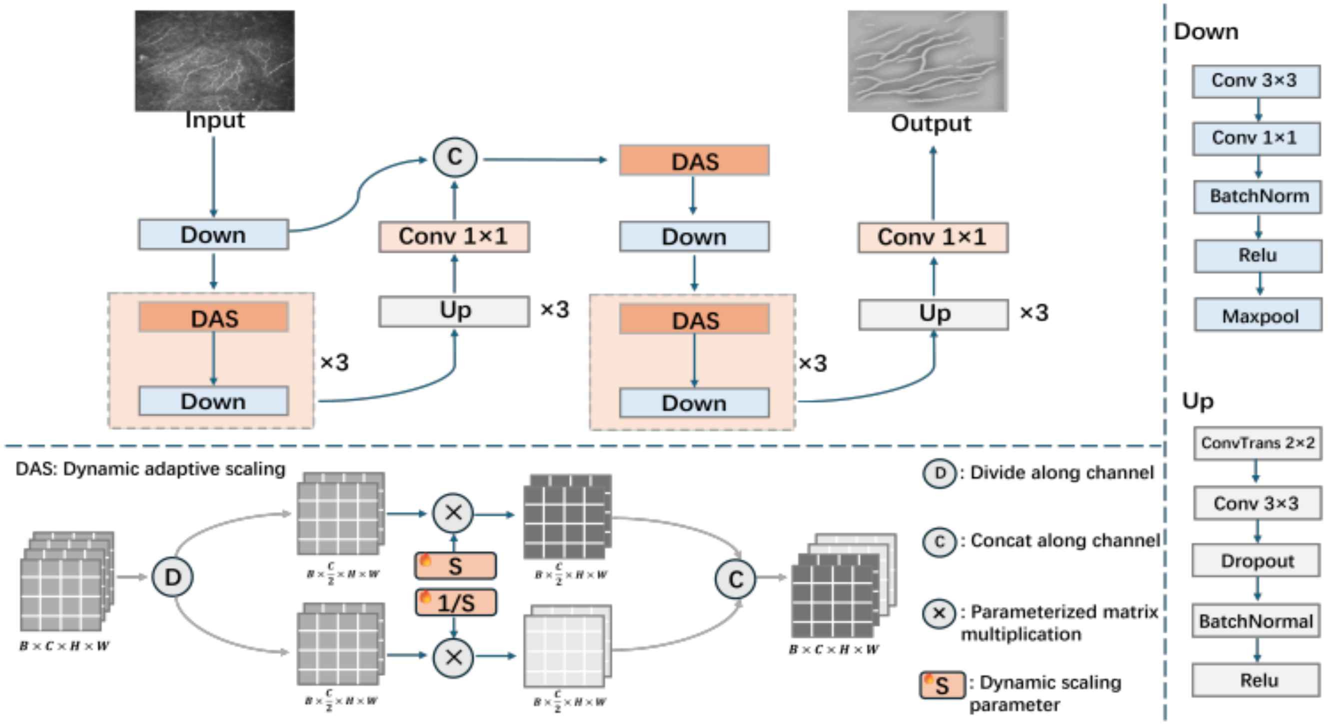


01 Background

At present, if the segmentation accuracy of deep learning-based segmentation technology is insufficient, the topological connectivity of the curve will be reduced, resulting in curve breakage. In addition, curves are often embedded in a background with large interference, and their texture or artifacts may have a high similarity with the target structure, which increases the difficulty of segme



02 Dynamic Structure Extraction Module

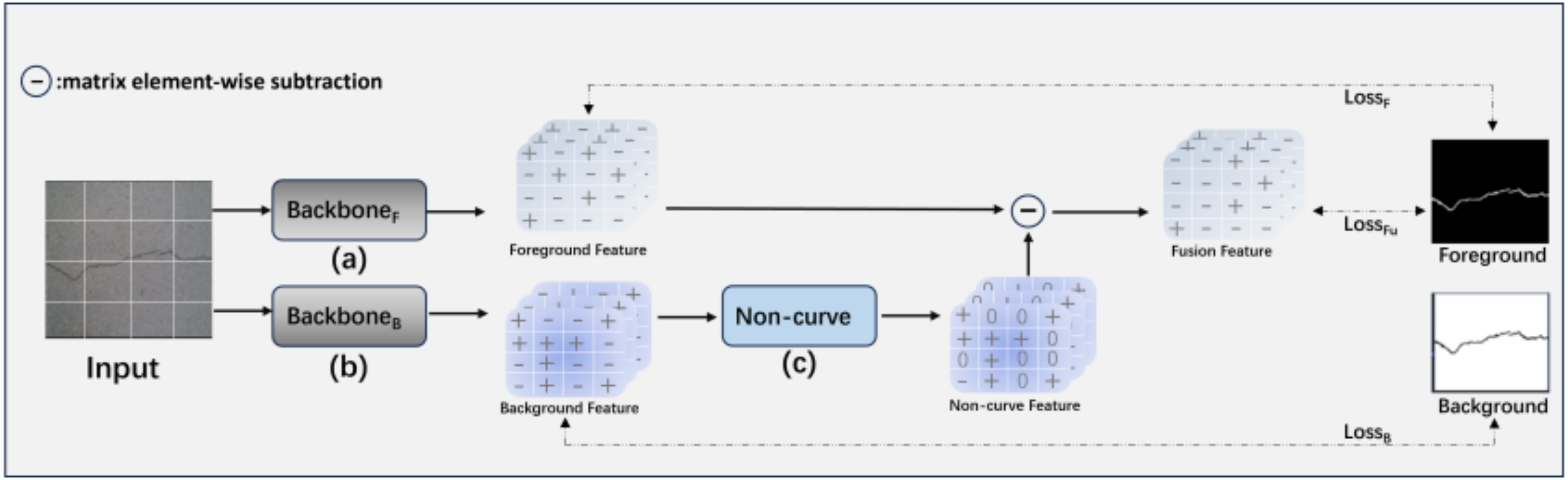


As depicted in the fig, the dynamic adaptive scaling process for each time can be described as follows:

$$x_{i+1} = \left(\frac{\mathbf{S}_i}{\mathbf{S}_{i+1}} \cdot \hat{x_i}\right) \cup \left(\frac{\mathbf{S}_{i+1}}{\mathbf{S}_i} \cdot \hat{x_j}\right)$$

where $\mathbf{S}$ is a learnable dynamic scaling parameter, and $\cup$ (.) is concat along the channel dimension. By introducing this dynamic structure, we can dynamically adjust the level of feature retention or discard in our

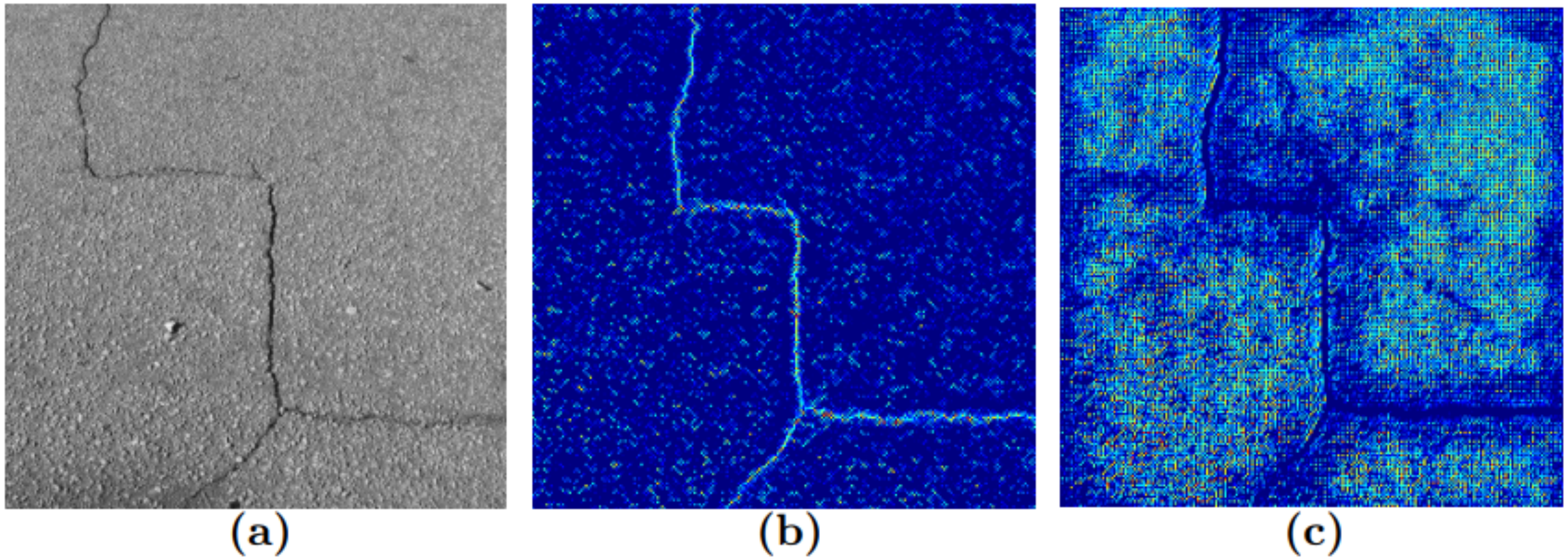
02 Non-Curve Suppression



Among them, the foreground is the prediction result of the curve, the background is the prediction result of the part outside the curve, and the non-curve background is the result after masking the prediction of the curve in the background prediction. For the output X_{fu} of each model cascade layer, it can be expressed as:

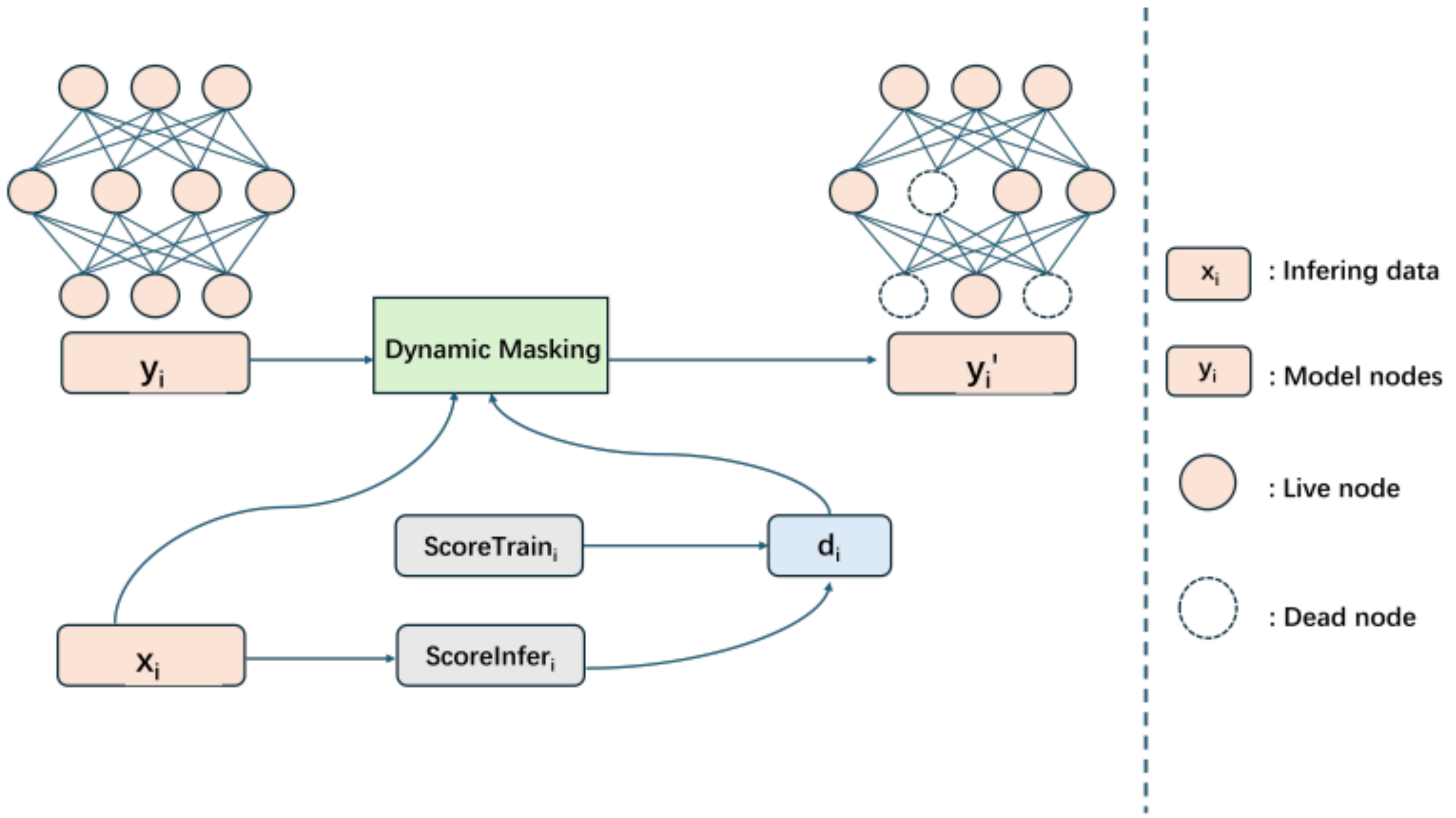
$$X_{fu} = X_f - \begin{cases} 0 & X_b < 0 \\ X_b & \text{other} \end{cases}$$

02 Non-Curve Suppression



(a) represents the original image, while (b) and (c) depict the heat maps showcasing the model's attention to the foreground and background, respectively. Brighter colors indicate areas where the model pays more attention.

02 Dynamic Masking



$$ScoresTrain_i = \frac{1}{n}\sum_{j=1}^{n} \mathbf{x}_{ij} - \sigma(\mathbf{x}_{ij})$$

$$d_i = \left| \frac{1}{n}\sum_{j=1}^{n} \mathbf{x}_{ij} - \frac{1}{n}\sum_{j=1}^{n} \hat{\mathbf{x}_{ij}} + \sigma(\hat{\mathbf{x}_{ij}}) - \sigma(\mathbf{x}_{ij}) \right|$$

$$p_{ik} = \frac{\mathbf{x}_i \mathbf{y}_{ik}}{d_i^2}$$

Masking of useless parameter nodes based on data scores: Comparison between ScoreT rain acquired during model layer training and ScoreInfer obtained in the current inference process.

03 Experimental Results

Dataset	Method	F1↑	Pre.↑	Rec.↑	mIoU↑	Mcc↑	Qua.↑
CORN-1	U-Net [20]	65.73	74.85	59.20	49.28	65.72	49.11
	IterNet [9]	66.67	74.92	60.60	50.32	66.56	50.17
	CS-Net [16]	65.93	**75.80**	58.54	49.45	65.94	49.22
	FR-UNet [12]	66.83	75.08	60.72	50.46	66.71	50.34
	DD-Net	**68.93**	70.46	**68.12**	**52.93**	**68.50**	**52.51**
CRACKF.	U-Net [20]	67.29	69.24	70.24	52.13	67.67	27.66
	IterNet [9]	69.04	72.93	69.48	54.03	69.29	28.94
	CS-Net [16]	67.86	72.51	68.37	53.67	68.11	29.39
	FR-UNet [12]	69.23	66.34	**76.19**	53.81	69.48	27.33
	DD-Net	**72.63**	**74.72**	72.63	**57.76**	**72.39**	**31.44**
CRACKT.	U-Net [20]	65.68	71.69	61.39	49.32	65.97	48.86
	IterNet [9]	71.16	**80.22**	64.61	55.86	71.65	55.65
	CS-Net [16]	66.60	80.11	57.44	50.56	67.53	50.47
	FR-UNet [12]	72.45	78.86	67.53	57.36	72.70	57.03
	DD-Net	**74.99**	79.14	**71.61**	**60.49**	**75.08**	**60.48**

$$MCC = \frac{tp \times tn - tp \times fn}{\sqrt{(tp + fp) \times (tp + fn) \times (tn + fp) \times (tn + fn)}} \tag{7}$$

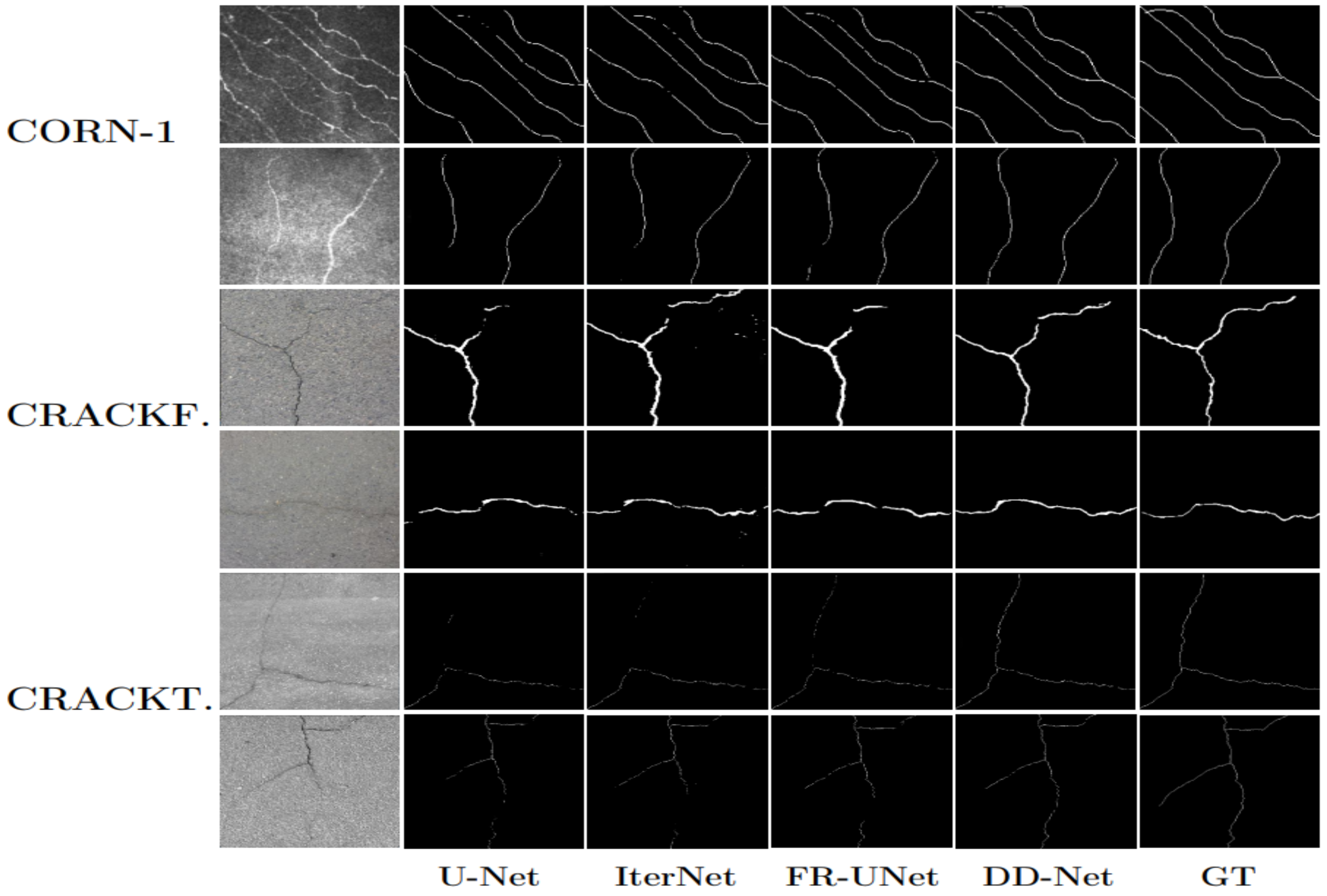
Consider one-pixel width segmentation, let the confusion matrix between the prediction skeleton and the ground truth skeleton of the thin curve are: TP_{sk}, FP_{sk}, TN_{sk}, FN_{sk}. The *Quality* is defined as:

$$Quality = \frac{TP_{sk}}{FP_{sk} + TP_{sk} + FN_{sk}} \tag{8}$$

To evaluate the effectiveness of our method on curve segmentation, we conduct qualitative and quantitative experiments on three different datasets, including CRACKTREE200 , CORN-1 , and CRACKFOREST.

To augment the data, we utilize horizontal flipping and random cropping techniques. All the networks are optimized by Adam with an initial learning rate of 0.001. During training, all training samples are cropped to a size of 256 × 256, and our network is trained with a batch size of 4.

03 Experimental Results



03 Ablation Study

Dataset	Method	F1↑	mIoU↑	CR↓
CORN-1	DD-Net w/o DM	68.92	52.92	100%
	DD-Net	**68.93**	**52.93**	**87.35%**
CRACKF.	DD-Net w/o DM	72.59	57.71	100%
	DD-Net	**72.63**	**57.76**	**87.35%**
CRACKT.	DD-Net w/o DM	74.98	60.48	100%
	DD-Net	**74.99**	**60.49**	**87.4%**

Dataset	Method	F1↑	mIoU↑	Mcc↑
CORN-1	IterNet	66.67	50.32	66.56
	IterNet w/ DS	**67.35**	**51.09**	**67.03**
CRACKF.	IterNet	69.04	54.03	69.29
	IterNet w/ DS	**69.39**	**54.30**	**69.67**
CRACKT.	IterNet	71.16	55.86	71.65
	IterNet w/ DS	**71.42**	**56.34**	**72.45**

$$CR = \frac{\sum_{i=1}^{n} |x_i|}{N \sum_{j=1}^{m} |y_j|}$$

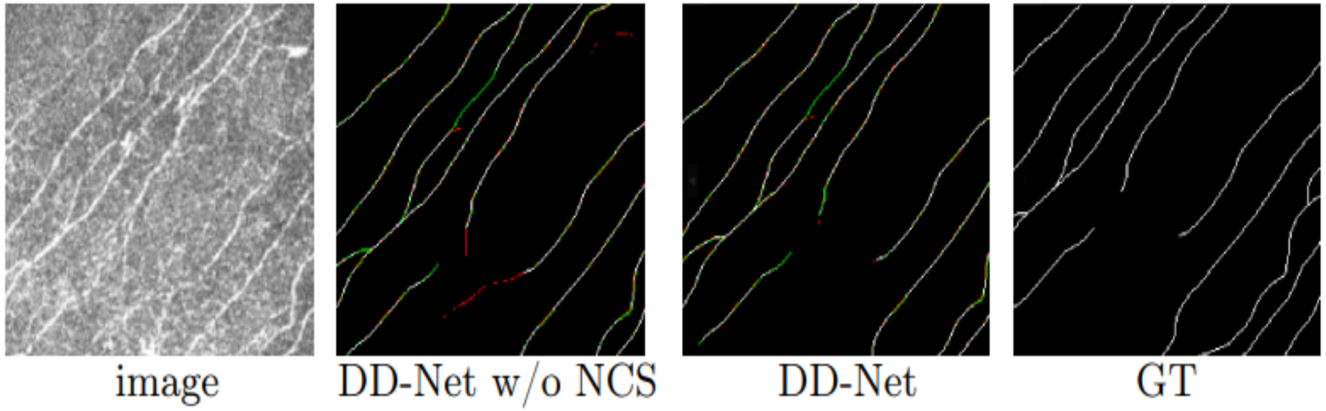
Dataset	Method	F1↑	mIoU↑	Mcc↑
CORN-1	U-Net	65.73	49.28	65.72
	U-Net w/ DS	**67.14**	**50.81**	**66.93**
CRACKF.	U-Net	67.29	52.13	67.67
	U-Net w/ DS	**68.94**	**53.86**	**69.23**
CRACKT.	U-Net	65.68	49.32	65.97
	U-Net w/ DS	**67.38**	**51.51**	**68.39**

It can be observed that the introduction of DS improves the performance of U-Net on various metrics across different datasets.

03 Ablation Study

Dataset	Method	F1↑	mIoU↑	Mcc↑
CORN-1	U-Net	65.73	49.28	65.72
	U-Net w/ NCS	**66.94**	**50.57**	**66.82**
CRACKF.	U-Net	67.29	52.13	67.67
	U-Net w/ NCS	**68.67**	**53.78**	**69.22**
CRACKT.	U-Net	65.68	49.32	65.97
	U-Net w/ NCS	**71.45**	**56.26**	**72.22**

Dataset	Method	F1↑	mIoU↑	Mcc↑
CORN-1	DD-Net w/o NCS	67.76	51.60	67.34
	DD-Net	**68.93**	**52.93**	**68.50**
CRACKF.	DD-Net w/o NCS	71.83	56.93	71.83
	DD-Net	**72.63**	**57.76**	**72.39**
CRACKT.	DD-Net w/o NCS	73.96	59.25	74.05
	DD-Net	**74.99**	**60.49**	**75.08**



Non-curve suppression serves as a valuable aid in suppressing erroneous foreground information.

The performance enhancement achieved by NCS sets the ability to generalize on different data. Certainly, we are interested in understanding the contribution of NCS to our DD-Net model. By comparing the performance of DD-Net with and without NCS, we can quantify the specific improvement brought about by this technique.

04 Summarize

• We propose a dynamic structure based on trainable dynamic scaling parameters to scale global feature, selectively retaining interesting parts of the model and proportionally discarding irrelevant information.

• We propose a dynamic inference strategy designed to adaptively adjust computations based on varying spatial locations within image data. This method can adaptively reduce the amount of redundant calculations only through inference.

• We introduce a non-curve suppression mechanism to enhance the preservation of the global feature integrity of the data by supplementing background feature information.

THANK YOU！

cyx529630@gmail.com