SRNet: Striped Pyramid Pooling and Relational Transformer for Retinal Vessel Segmentation

Wei Yan*, Yun Jiang, Zequn Zhang, Yao Yan, Bingxi Liu

School of Computer Science and Engineering, Northwest Normal University, Lanzhou, Gansu Province, China

BMVC 2023

Abstract

- We propose a Relational Transformer Module for fusing image-level and patcheslevel information, which can combine the strengths of image-level and patches-level segmentation frameworks. The introduction of cross-transformer and self-attention effectively focuses on the local information of the vessels end and the long-distance context correlation of the vessels.
- According to the inherent characteristics of retinal vessels, we designed a novelStriped Pooling Pyramid Module, which can better capture the characteristic information of thin tree-like long blood vessels and further improve the segmentation accuracy.
- Based on the above innovations, we propose a retinal vessels segmentation network (SRNet). We conduct comprehensive experiments on DRIVE and CHASE datasets, all achieving state-of-the-art performance.

Visualizations

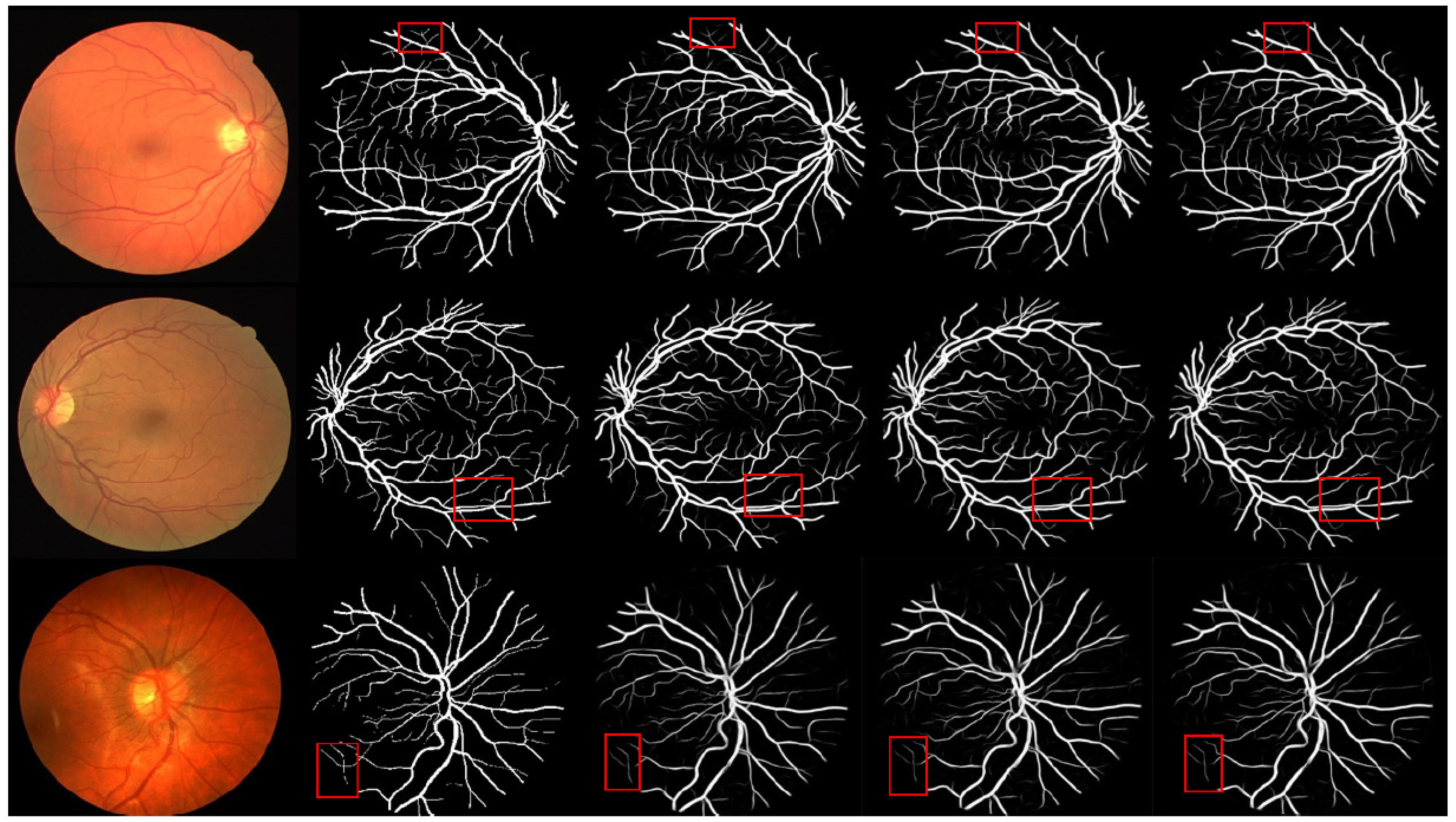

Figure 5: Visualization of segmentation results on DRIVE and CHASE datasets. The first and second lines are the DRIVE dataset, and the third line is the CHASE dataset. From Left to Right: Retina images, Ground truths, proposed **SRNet**, DA-Net[29], and CGANet [28]

Method

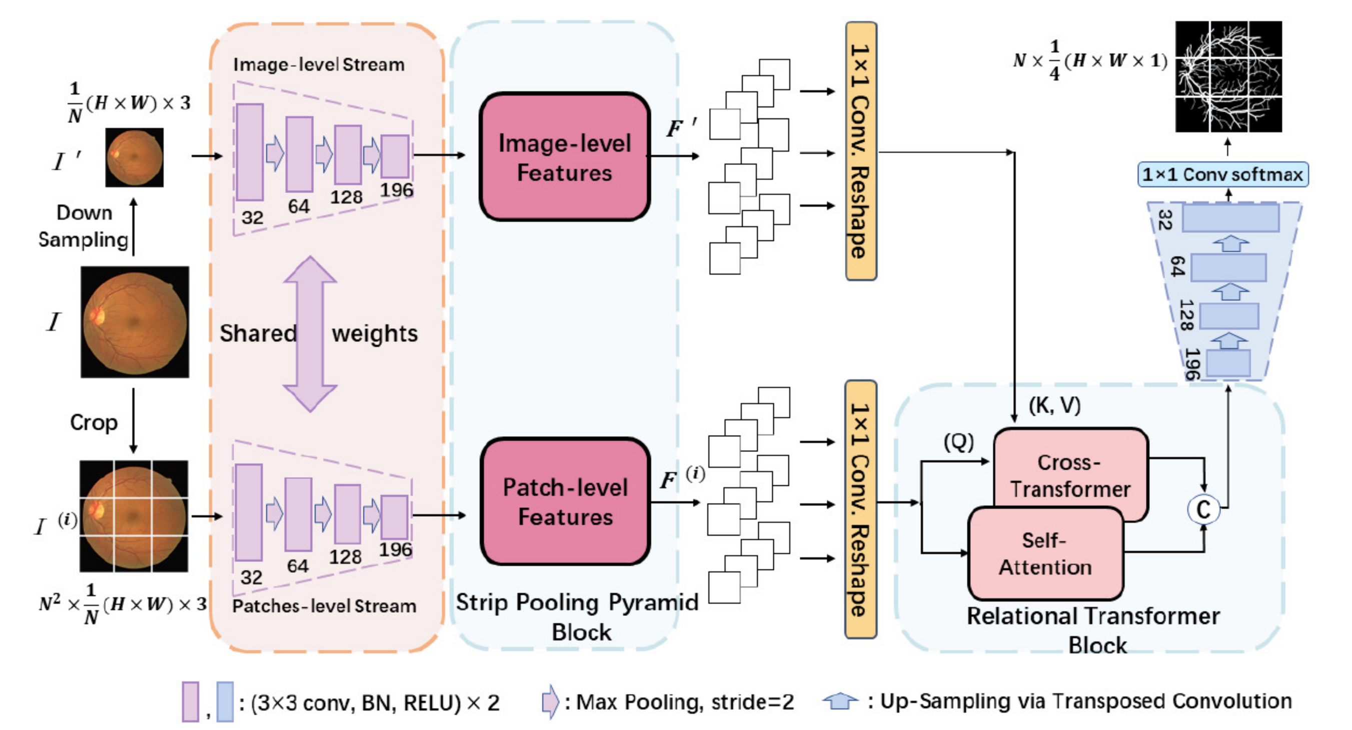

Figure 1: The overall architecture of our SRNet. The shared encoder consists of 4 layers of convolutional blocks, SPPM and RTM, and an up-sampling decoder.

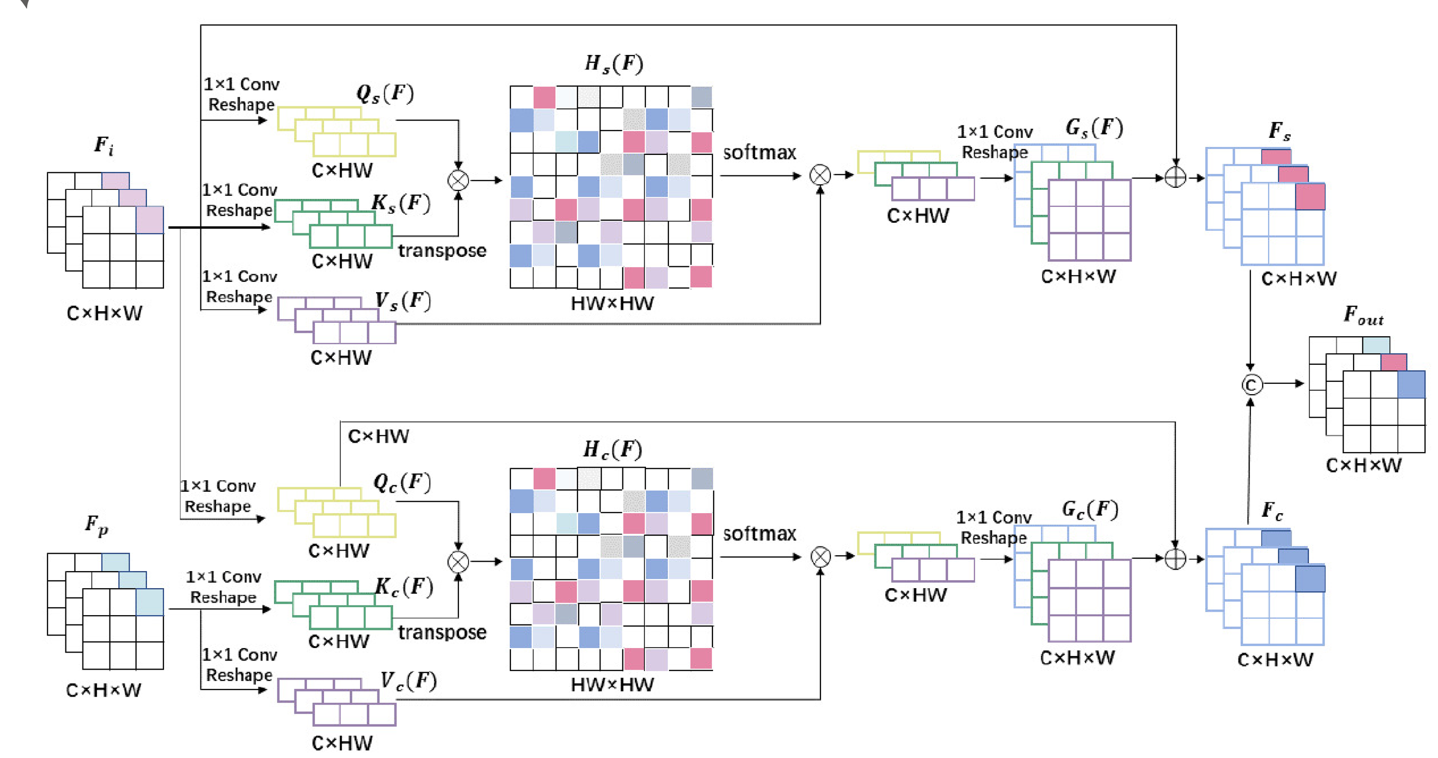

Figure 2: The details of the Relational Transformer Module(RTM).

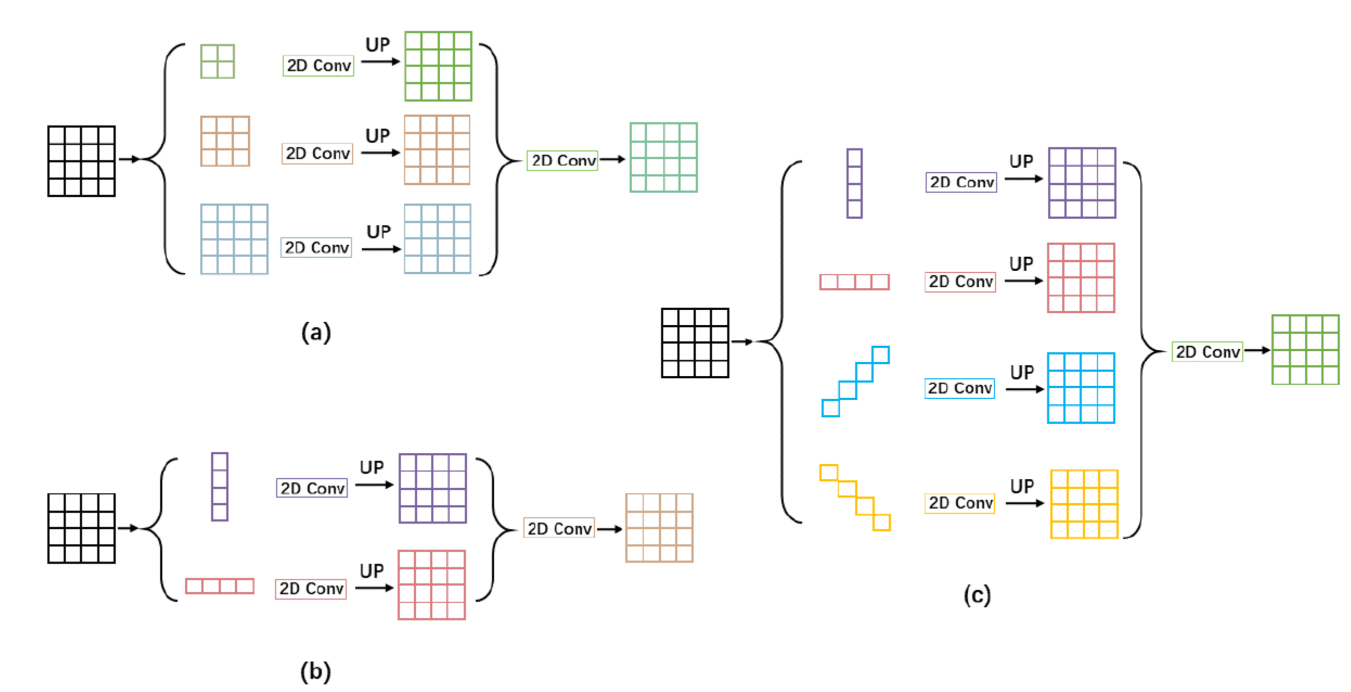

Figure 3: (a) Regular square pooling pyramid. (b) Regular strip pooling pyramid. (c) Our proposed pooling pyramid with a diagonal pooling kernel.

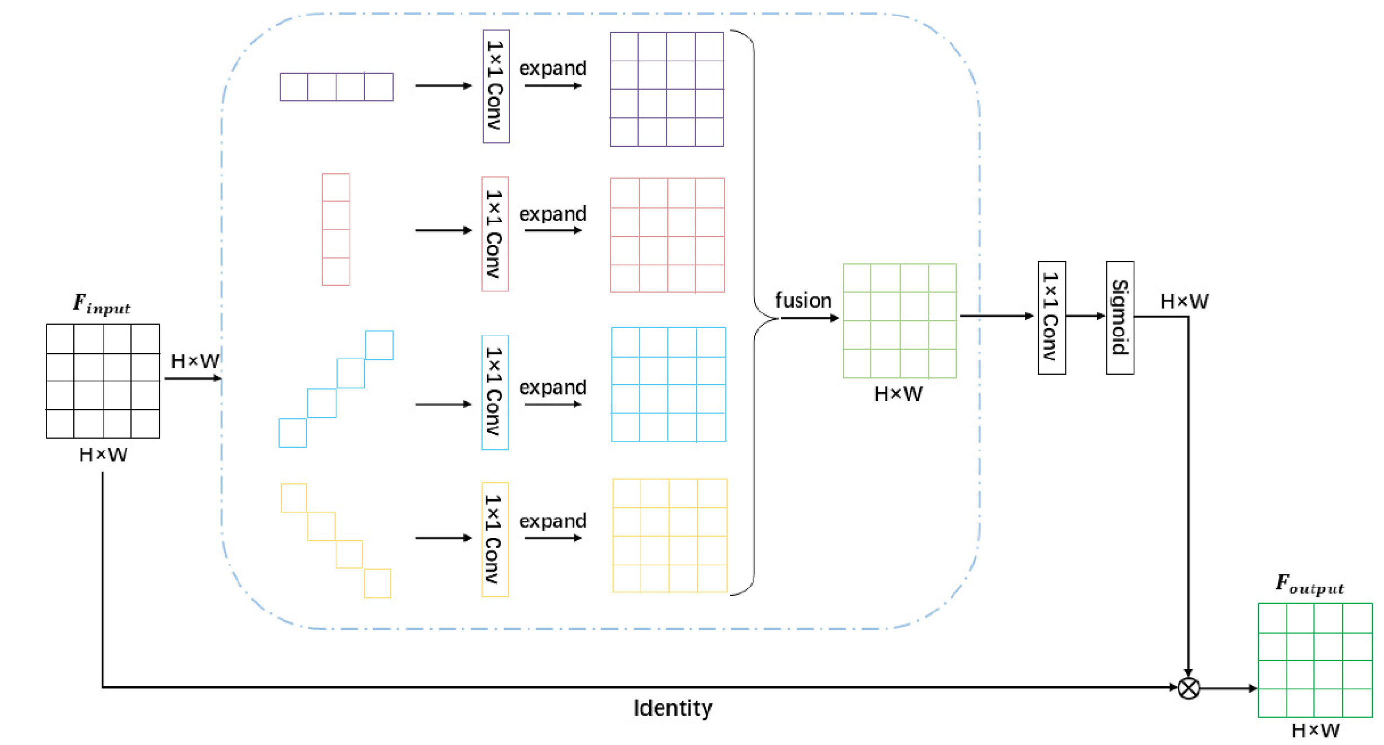

Figure 4: Schematic illustration of the Striped Pooling module(SPPM).

Experimental Results

Table 2: Comparison with other state-of-the-art methods on the DRIVE dataset. The best result is marked in **bold** and the second best result is underlined.

Method	year	Input Type	Acc	Sen	AUC
JL-UNet[34]	2018	Patches	95.56	77.92	97.84
MS-NFN[33]	2018	Patches	95.67	78.44	98.07
CE-Net[8]	2019	Image	95.45	83.09	97.79
CTF-Net[30]	2020	Patches	95.67	78.49	97.88
CGA-Net[28]	2021	Image	96.47	83.05	98.65
SCS-Net[32]	2021	Image	96.97	82.89	98.37
DA-Net[29]	2022	Joint	97.07	85.57	99.03
SRNet(our)	2023	Joint	**97.09**	**85.68**	**99.13**

Table 3: Comparison with other state-of-the-art methods on the CAHSE dataset. The best result is marked in **bold** and the second best result is underlined.

Method	year	Input Type	Acc	Sen	AUC
JL-UNet[34]	2018	Patches	96.10	76.33	97.81
MS-NFN[33]	2018	Patches	96.37	75.38	98.25
CE-Net[8]	2019	Image	96.89	81.52	98.30
CTF-Net[30]	2020	Patches	96.48	79.48	98.47
CGA-Net[28]	2021	Image	97.06	86.78	98.12
SCS-Net[32]	2021	Image	97.44	83.65	98.67
DA-Net[29]	2022	Joint	97.66	87.04	99.08
SRNet(our)	2023	Joint	**97.82**	**87.06**	**99.17**