

Partial Class Activation Attention for Semantic Segmentation

CVPR, 2022

语义分割的类激活注意

Motivation

解决语义分割中存在的类内不一致问题：

语义分割中，同一类别物体的像素特征会由于纹理、光照、位置的不同而有显著差异。因此，我们往往需要用全局的一致性来消除这种局部的特殊性。

先前解决这个问题的典型办法有金字塔池化、扩张卷积、自注意力。

自注意力的模型：首先计算像素关系，然后通过基于关系图的加权聚合来增强特征。

思路：使用Partial CAM来计算像素到类的关系，然后提出PCAA，将PCAM收集到的局部类表示聚合到一个全局类中心，来用作特征增强的基础。

Framework

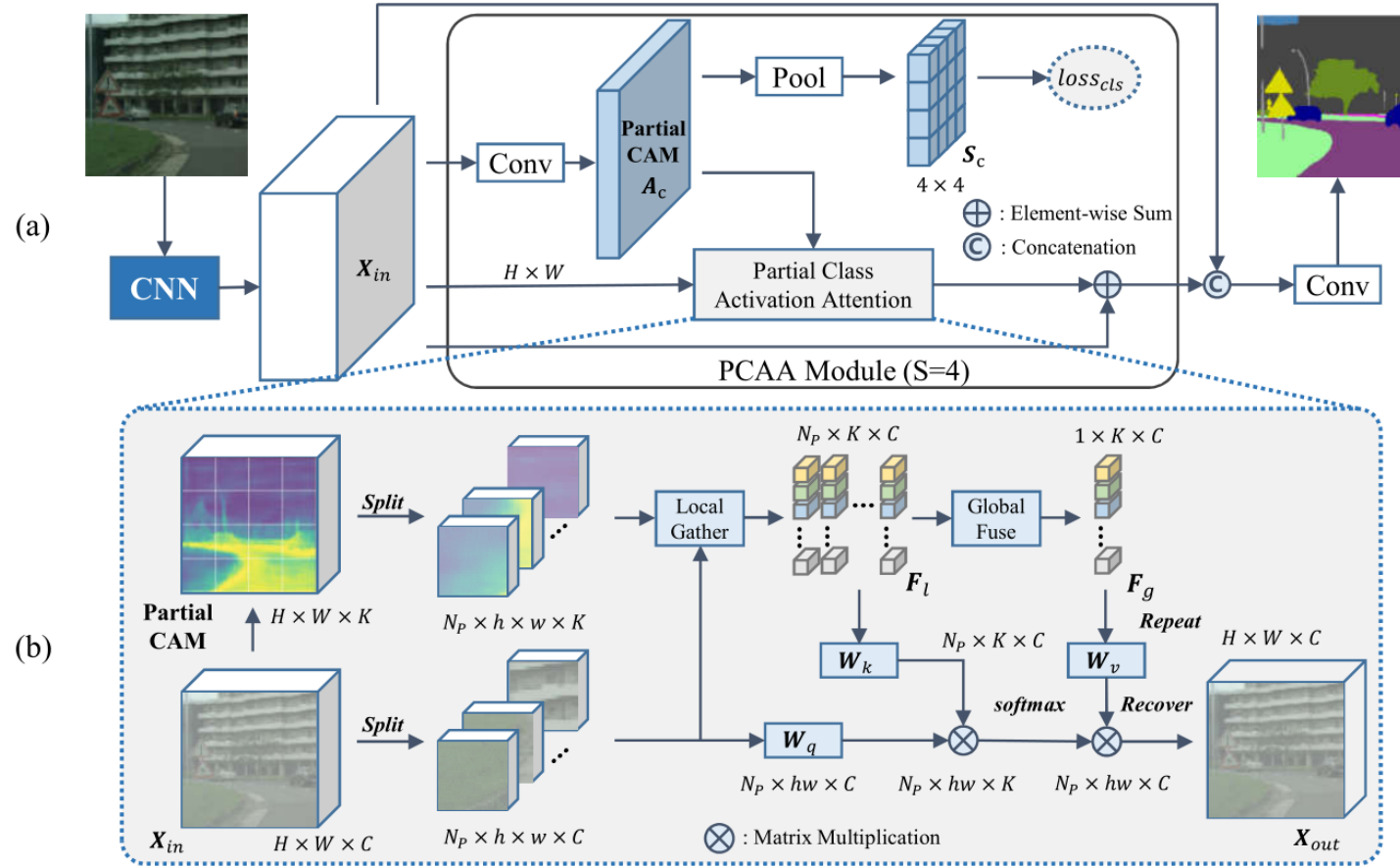


Figure 2. Detailed architecture of the proposed approach. The overall network structure is shown in (a), where the pooling size of the Partial Class Activation Attention (PCAA) module is set to 4. (b) illustrates the process of partial class activation attention calculation. It utilizes local class centers to compute similarity maps and uses global representations for feature aggregation.

Framework

PCAA:

1.局部类中心：将CAM分块得到PCAM之后，计算每一个块的局部类中心

$$\hat{\mathbf{F}}_l^{(i)} = \tilde{\mathbf{S}}_c^{(i)} \cdot [\sigma_s(\tilde{\mathbf{A}}_c^{(i)})^\top \times \tilde{\mathbf{X}}_{in}^{(i)}]. \quad (5)$$

$i \in 1, 2, \dots, N_p - 1$

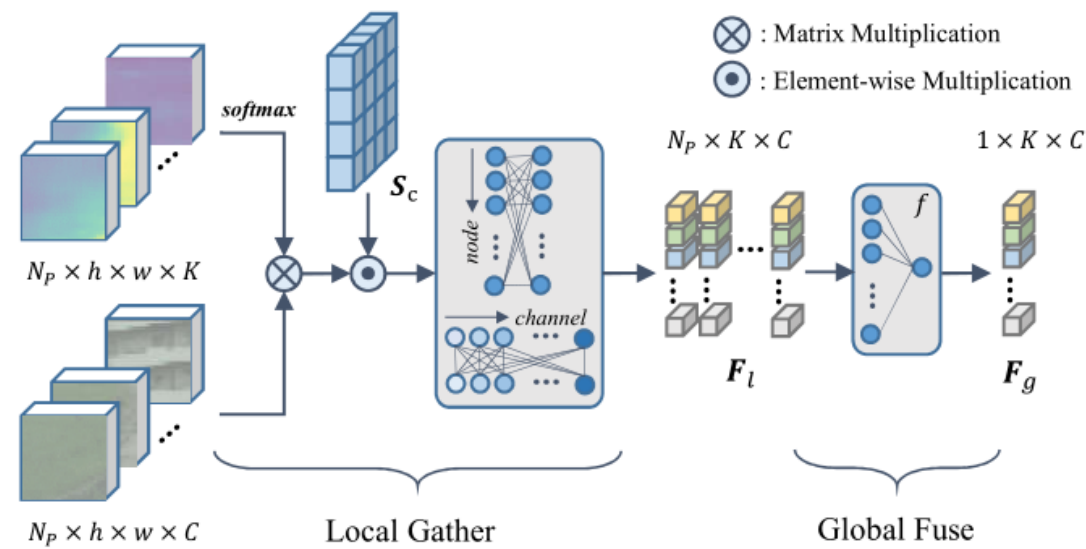


Figure 3. Illustration of gathering local class centers and fusing them as global representations.

Framework

PCAA:

2.全局类中心:

$$\mathbf{F}_g = \sum_i f_i \mathbf{F}_l^{(i)}, \quad (7)$$

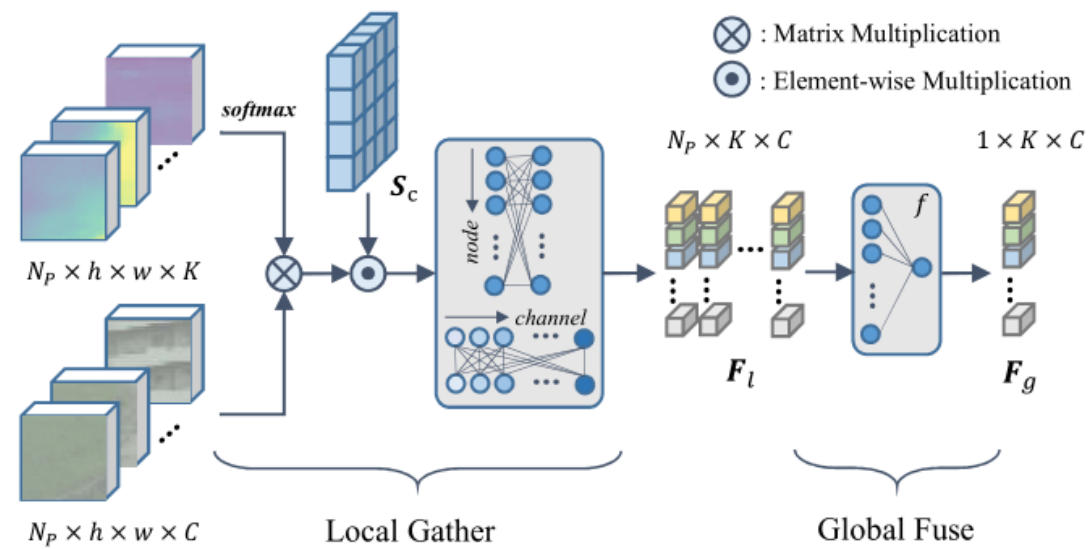


Figure 3. Illustration of gathering local class centers and fusing them as global representations.

Framework

PCAA:

3. 特征聚合:

$$\mathbf{P}^{(i)} = \sigma_c(W_q(\tilde{\mathbf{X}}_{in}^{(i)}) \times W_k(\mathbf{F}_l^{(i)})^\top),$$

$$\tilde{\mathbf{X}}_{out}^{(i)} = \mathbf{P}^{(i)} \times W_v(\mathbf{F}_g). \quad (9)$$

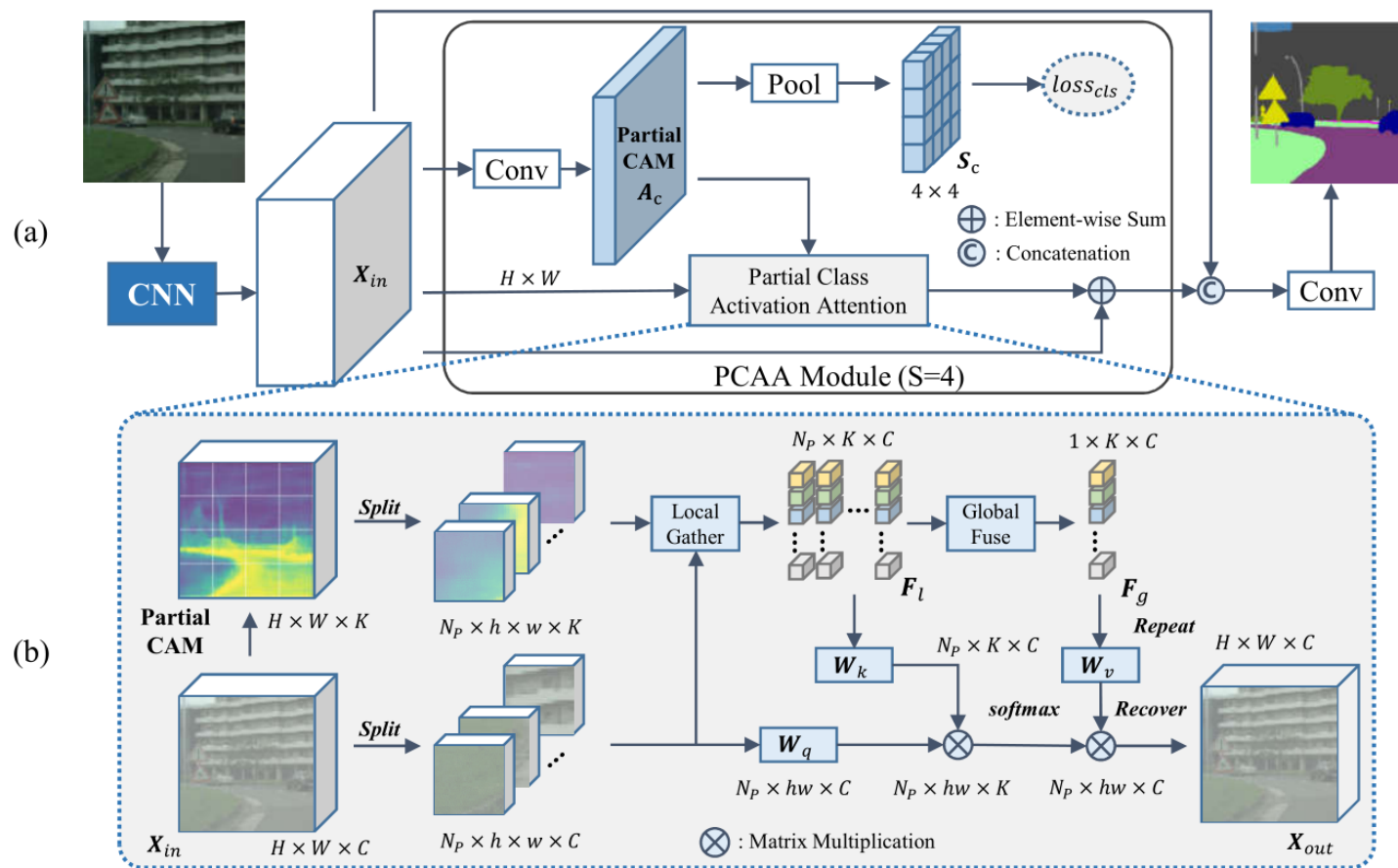


Figure 2. Detailed architecture of the proposed approach. The overall network structure is shown in (a), where the pooling size of the Partial Class Activation Attention (PCAA) module is set to 4. (b) illustrates the process of partial class activation attention calculation. It utilizes local class centers to compute similarity maps and uses global representations for feature aggregation.

Results

消融实验：

S=1	S=2	S=4	S=8	S=16	mIoU(%)
✓					77.47
	✓				78.68
		✓			79.22
			✓		79.00
				✓	78.25
	✓	✓			78.46
		✓	✓		79.29
			✓	✓	78.70

Table 1. Ablation study on pooling size in PCAA module.

Results

消融实验:

Key	F_l	F_l	F_g	F_l	F_l
Value	F_l	F_l	F_g	F_g	F_g
GCU		✓	✓		✓
mIoU(%)	75.36	78.89	78.68	78.93	79.22

Table 2. Ablation study on attention calculation.

Results

:

Method	Backbone	mIoU(%)
FCN (Baseline)	ResNet-50	74.68
+ASPP	ResNet-50	78.34
+NL	ResNet-50	78.65
+OCR	ResNet-50	78.86
+PCAM	ResNet-50	78.84
+PCAA	ResNet-50	79.22
+PCAA	ResNet-101	80.70

Table 3. Experimental results on the Cityscapes validation set.