

DSNet-ImageNet-COCO

Architecture

Method	Backbone	Plugin of every stage				Position of plugins
DSNet50-A	ResNet50	None	None	None	DSM*	after conv2
DSNet50-B		DSM	DSM	DSM	DSM	after conv1
DSNet50-C		None	None	DAM**	DSM	after conv2
FcaNet50		MSCA** *	MSCA	MSCA	MSCA	after conv3

- 我们在论文中提出的主要方法是 DSNet50-A，我们的出发点是为了更好的解决语义分割问题，所以 DSNet50-A 在语义分割上的实验优于 FcaNet50，但是分类和检测上稍弱于 FcaNet50。
- 为了更好分类和检测的结果，我们提出了 DSNet50-B，其架构更像 FcaNet50，但是参数量均少于 FcaNet50 和 DSNet50-A，且分类和检测性能位于这两者之间。

*DSM: Decoupled self-attention module

**DAM: Decoupled attention module

***MSCA: Multi-spectral channel attention

ImageNet1K

Method	Parameters	FLOPS	Train FPS	top1	top5
ResNet50	25.557M	4.109G	1677.51	76.48	93.17
FcaNet50	30.121M	4.112G	1036.70	77.50	93.79
DSNet50-A	27.369M	4.357G	1873.95	76.51	93.12
DSNet50-B	29.106M	4.208G	1424.30	77.28	93.60

- 以上实验均是在 mmpretrain 中进行实验的，所以与 FcaNet 论文报告的有差异（FcaNet 论文使用的是 Nvidia APEX mixed precision training toolkit，而我们没有使用混合精度 mixed precision）。

COCO2018

Model	Method	Parameters	FLOPS	Train FPS	AP	AP_5
Faster - RCNN	ResNet50	41.750M	187.20G	46.75	37.4	58.3
	FcaNet50	44.268M	187.31G	28.90	38.9	60.2
	DSNet50-A	45.302M	188.97G	44.95	37.8	59.4
	DSNet50-B	43.565M	194.27G	64.89*	38.2	59.6
	DSNet50-B	43.565M	194.27G	49.65**	38.3	59.7
	DSNet50-B*	43.565M	194.27G	41.34	38.3	59.4

- 我们的 DSNet50-B 的 Parameters 相比 FcaNet 较少，训练的 FPS 两倍于 FcaNet50

*可能有更多核 CPU 的优势.

**在 3090 上训练，其他都是在 4090 上训练.

语义分割

- 以下是两个**语义分割实验**，注意，我们的 DSNet50-A 的 Parameters 和 FLOPS 较大，是由于我们的 DSM 保留了较高的空间分辨率和通道维度，也因此获得更好的分割效果。

Synapse

Method	Parameters	FLOPS	Train FPS	mDice	mIoU
ResNet50	47.13M	197.86G	7.52**	82.92	—*
FcaNet50	49.62M	196.67G	81.60	84.06	76.1
DSNet50-A	50.67M	205.93G	60.4	85.07	77.61
DSNet50-B	48.94M	204.44G	95.72	83.8	75.76

*我们没有测试 mIoU.

**在单卡 A5000 上训练，其他都是 8 卡 4090 测试。

Cityscapes

Decoder	Method	Pretrained weights*	Parameters	FLOPS	Train FPS
FCN	ResNet50 V1C**	—	47.13M	395.76G	—
	FcaNet50	FcaNet50	49.65M	395.91G	48.9
	DSNet50-A	DSNet50-A	50.68M	411.91G	46.9
	DSNet50-B	DSNet50-B	48.95M	408.92G	49.7
	ResNet50 V1C	ResNet50 V1C	47.13M	395.76G	48.9
	FcaNet50		49.65M	395.91G	69.1
	DSNet50-A		50.68M	411.91G	70.1

*由于时间关系，我们没有在 ImageNet1K 预训练基于 ResNet V1C 的 FcaNet50 和 DSNet50，而使用了两种训练策略：（1）使用基于 ResNet(即 ResNet V1B)的 FcaNet50 和 DSNet50 的预训练权重；（2）都使用 ResNet50 V1C 的训练权重。

**Compare to ResNet, ResNet V1C replace the 7x7 conv in the stem with three 3x3 convs.

***由于时间关系，我们没有亲自训练，ResNet50 使用的是 mmseg 的 official results.

ResNet50:

- Dataset: ImageNet1K
- Input shape: (3, 224, 224)
Flops: 4.109G
Params: 25.557M
Activation: 11.114M
- Dataset: COCO2017
- Compute type: dataloader: load a picture from the dataset
Input shape: (1088, 800)
Flops: 187.20G
Params: 41.75M

DSNet:

- Dataset: ImageNet1K
- Input shape: (3, 224, 224)
Flops: 4.357G
Params: 27.369M
Activation: 14.172M
- Dataset: COCO2017
- Compute type: dataloader: load a picture from the dataset
Input shape: (1088, 800)
Flops: 194.27G
Params: 43.565M

Flops: 188.97T

Params: 45.302M

FcaNet:

- Dataset: ImageNet1K

- Input shape: (3, 224, 224)
Flops: 4.112G
Params: 30.121M
Activation: 11.13M
- Dataset: COCO2017
- Compute type: dataloader: load a picture from the dataset
Input shape: (1088, 800)
Flops: 187.31G
Params: 44.268M