2020cvpr 的一篇文章，提出了一种搜索 channel 配置的方法。
定义的问题是（problem）：

$$\mathbf{c}^{\star} = \arg\max_{\mathbf{c}} \mathrm{Acc}_{\mathrm{val}}(\mathbf{c}, \mathbf{W}^{\star}) \quad (1)$$
$$\text{s.t.} \quad \mathrm{FLOPs}(\mathbf{c}) = F_0, \quad \mathbf{W}^{\star} = \arg\max_{\mathbf{W}} \mathrm{Acc}_{\mathrm{train}}(\mathbf{c}, \mathbf{W})$$

C 是各层的 channel 数量配置，对应为一个向量。在限制 c 对应的 flops 一定的情况下，求最优的配置 c，使网络精度最大。

Method:
For i in range(T):
1 train the current network
2 compute the FLOPs utilization ratio(FURs) on each layer
3 increased the numbers of channels of a few layers with top-ranked FURs, and decrease the numbers of a few layers with bottom-ranked FURs.
FURs：

$$\mathrm{FUR}_l = \frac{\partial \mathrm{Acc}_{\mathrm{val}}(\mathbf{c}, \mathbf{W}^{\star})}{\partial \mathrm{FLOPs}(\mathbf{c})} = \frac{\partial \mathrm{Acc}_{\mathrm{val}}(\mathbf{c}, \mathbf{W}^{\star}) / \partial c_l}{\partial \mathrm{FLOPs}(\mathbf{c}) / \partial c_l} \quad (2)$$

求 Acc 相对于 channel 的梯度，flops 相对于 channel 的梯度，两者相除得到一个 flops 利用率。梯度就是相对变化，作者用 SpatialDropout 来去除一定的 channel 来获得 acc 和 flops 的变化，作为上式中的梯度，最终得到每层的 FURt。

Result:

Table 4. Top-1 accuracy (%) comparison on ImageNet between our approach and other competitors. Here, 'FT' indicates that the searched network is only fine-tuned, not trained from scratch, and 'KD' is for knowledge distillation. The mark 'imp' indicates our own implementation. Legend: LCCL [5], SFP [15], FPGM [16], TAS [6], AutoSlim [48].

Networks	Method	FT	KD	Top-1 (%)	Top-5 (%)	FLOPs (M)	Parameters(M)
ResNet-18	LCCL			66.33 (-3.65)	86.94 (-2.29)	1.19E3	-
	SFP			67.10 (-3.18)	87.78 (-1.85)	1.06E3	-
	FPGM	✓		68.41 (-1.87)	88.48 (-1.15)	1.06E3	-
	TAS		✓	69.15 (-1.5)	89.19 (-0.68)	1.21E3	-
	original (imp)			70.31	89.45	1.81E3	11.7
	searched			71.22 (+0.91)	90.07 (+0.62)	1.81E3	18.5
	searched 0.8×			69.41 (-0.90)	88.71 (-0.74)	1.17E3	11.8
	searched 0.8×		✓	71.00 (+0.69)	90.05 (+0.60)	1.17E3	11.8
MnasNet	original (imp)			74.21	91.83	312	3.89
	AutoSlim			74.60 (+0.60)	-	315	6.00
	ours			74.88 (+0.67)	92.15 (+0.32)	312	4.96

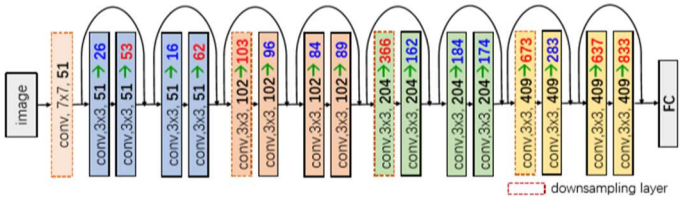

Figure 4. The configuration of ResNet-18 0.8× before and after the entire network adjustment process (10 rounds) on ImageNet. Red and blue indicate increased and decreased channel numbers, respectively.

0.8x 的意思就是将 resnet18 的 channel 数缩放到 0.8 倍来达到一定的 flops，和之前的对比方法在同一 level 的 flops 上对比性能。