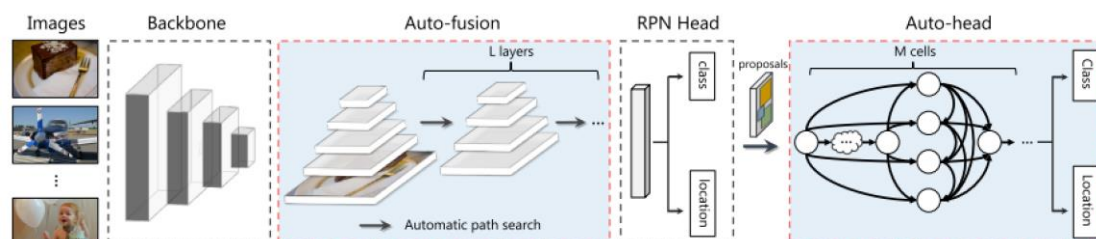
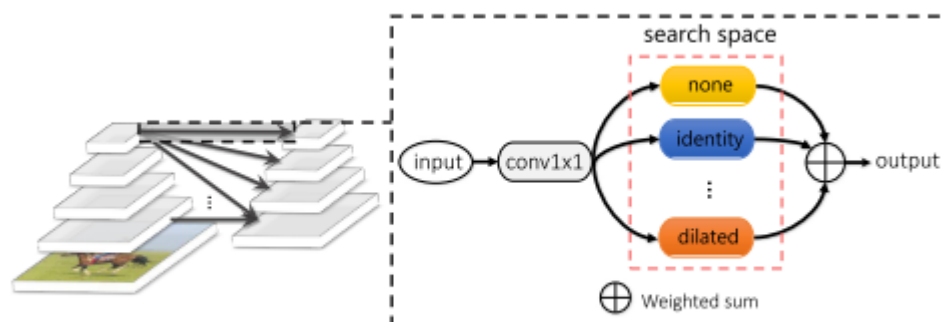


神经结构搜索（NAS）在图像分类中有很大的潜力来自动设计良好的 CNN 结构。作者在此研究了用于目标检测的 NAS。现存的工作主要集中于将从分类任务搜索到的结构迁移到检测任务，而其余的体系结构保持不变。然而，这种流水线不是特定于任务或面向数据的网络搜索，不能最佳适应于任何数据集。作者提出了的架构搜索框架，名为 Auto-FPN，专为检测而设计，而不是简单地一个分类搜索主干。



如上图所示，Auto-FPN 包括四个部分，其中 backbone 可以为任何一个 backbone，RPN Head 和 Faster RCNN 中 RPN 类似。另外两个则是模型需要自动化搜索网络结构。自动融合模块（Auto-fusion）旨在找到一种更好的体系结构，以利用来自所有特征层次结构的信息，并进行特征融合以获得更好的预测。自动头（Auto-Head）模块由不同的自动搜索单元构成，以执行更好的分类和 bbox 预测。



(a) Auto-fusion layer

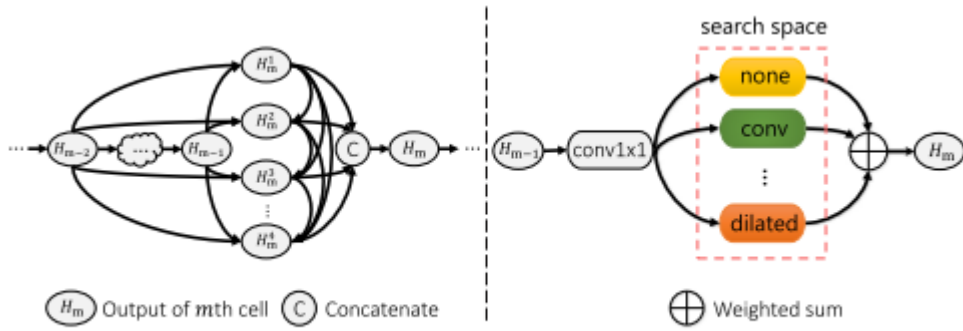
上图所示为 Auto-fusion 模块。每一层的 Auto-fusion 包含一系列的在特征之间的连接。每一个连接是由  $1 \times 1$  的卷积核和一组平行的操作算子  $\mathcal{O}_N$ . 其中， $\mathcal{O}_N$  包含下面几个操作：

- *no connection (none)*
- *skip connection (identity)*
- *3×3 dilated conv with rate 2*
- *3×3 dilated conv with rate 3*
- *5×5 dilated conv with rate 2*
- *5×5 dilated conv with rate 3*
- *3×3 depthwise-separable conv*
- *5×5 depthwise-separable conv*

定义  $P_i^l = O_{i \rightarrow j}(P_j^{l-1})$  为从第  $l-1$  层的第  $j$  个尺度特征到第  $l$  层的第  $i$  个尺度特征的操作，则  $O_{i \rightarrow j}^l$  的估计值通过下面公式计算：

$$\hat{O}_{i \rightarrow j}(P_j^{l-1}) = \sum_{O^k \in \mathcal{O}_N} \alpha_{i \rightarrow j}^{kl} O_{i \rightarrow j}^{kl}(P_j^{l-1})$$

其中：  $\sum_{k=1}^{|\mathcal{O}_N|} \alpha_{i \rightarrow j}^{kl} = 1$  并且  $\alpha_{i \rightarrow j}^{kl} \geq 0$ ，最终，通过  $\arg \max(\alpha_{i \rightarrow j}^{kl})$  选择最有可能的操作。



(b) Auto-head cell

上图为 Auto-head 单元。它包含 7 个节点：2 个输入节点，4 个中间节点和 1 个输出节点。对于第  $m$  个单元，第  $(m-2)$  个单元和第  $(m-1)$  个单元的输出作为其输入，输出节点是所有中间节点  $\{H_m^1, H_m^2, H_m^3, H_m^4\}$  的深度连接。类似于 Auto-fusion，需要学习每个单元的结构参数  $\beta_m$ 。对于操作算子  $\mathcal{O}_H$ ，添加了  $3 \times 3$  的卷积和  $5 \times 5$  卷积允许充分利用提取的特征，并且删除了倍率为 3 的空洞卷积，因为对于小的空间分辨率来说，空洞卷积不再有用。

最后，为了限制资源使用，加入了资源占用参数： $C(\alpha, \beta) = \sum_{i,j,k,l} \alpha_{i \rightarrow j}^{kl} C(O_{i \rightarrow j}^{kl}) + \sum_{i,j,k} \beta_{i \rightarrow i}^k C(O_{i \rightarrow i}^k)$ ，其中  $C(O_{i \rightarrow j}^k)$  是每个操作算子的计算消耗，其值为规范化的参数大小、FLOPs 和 MAC（内存访问成本）三者之和，都是可微的，支持反向传播。

	Search Method	No. arch searched	Search time (GPU days)	Average/Best mAP of searched arch
PASCAL VOC	Random	20	~ 18.3	76.4/80.3
	Evolutionary	60	~ 20.0	81.1
	SNAS[58]	1	~ 0.8	81.0
	Auto-FPN	1	~ <b>0.8</b>	<b>81.8</b>
MS-COCO	Random	10	~ 130.0	36.4/38.2
	Evolutionary	30	~ 68.3	38.6
	SNAS[58]	1	~ 16.0	37.9
	Auto-FPN	1	~ <b>16.0</b>	<b>40.5</b>

Table 8. More NAS baselines upon current search space on VOC and COCO.

结果比较，可以发现，search time 几乎最佳水平，mAP 同时也较有提升。