**论文贡献**：

(1)设计了新的搜索空间，即NASNet search space，并在实验中搜索得到最优的网络结构NASNet

(2)提出新的正则化技术，ScheduledDropPath，是DropPath方法的改进版，可以大大提高了模型的泛化能力。

DropPath方法在训练过程中以随机概率p进行drop，该概率训练中保持不变。而ScheduledDropPath方法在训练过程线性的提高随机概率p。

**论文核心**

看完NAS论文和此篇论文，我将从以下几点描述论文的核心点。

**核心一**：延续NAS论文的核心机制使得能够自动产生网络结构；

**核心二：**采用resnet和Inception重复使用block结构思想；

**核心三**：利用迁移学习将生成的网络迁移到大数据集上提出一个new search space。

图中的预测步骤一共是5个步骤：

  Step1：从hidden states（前面block中产生的）中选择一个hidden state—hi-1；

  Step2：重复step的操作，选择一个hidden state—hi；

  Step3：为step1中选择的hidden state选择一个操作；

  Step4：为step2中选择的hidden state选择一个操作；

  Step5：选择一个方法去连接step3和step4中的输出从而产生一个new hidden state，并将其添加到hidden states中。

说明

考虑到计算资源的限制，论文限制了search space，设置B=5。

Step3和step4中选择的操作可以下面的这些选项选择

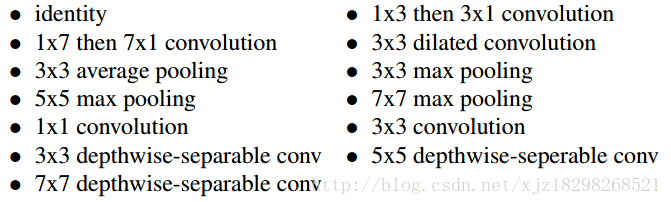


图7就是作者通过实验得到的表现最好的Normal Cell和Reduction Cell结构。每个convolution cell（比如Fiugre4中的Normal Cell和Reduction Cell）都是B个block的结果，在Figure4中，B=5，所以可以看到不管在Normal Cell还是Reduction Cell中都有5个add操作。注意图6展示的是论文称为NASNet-A的Cell结构，论文中的其他两种基于不同的search spaces得到的NASNet-B和NASNet-C的Cell结构可以看论文中最后的附录部分。可以看出虽然作者没有设计类似ResNet那样的residual connection结构（或者叫skip connection），但是这些Cell在训练过程中自己学会了这种skip connection（图7中的虚线连接），这个学习的过程就是图6中前面两个灰色矩形框的select过程（可以和ICLR2017的NAS中设计的skip connection结构对比，在那篇文章中采用的sigmoid结构表达层与层之间的连接关系），而且作者发现当手动插入residual connection时并没能提高模型的效果。

