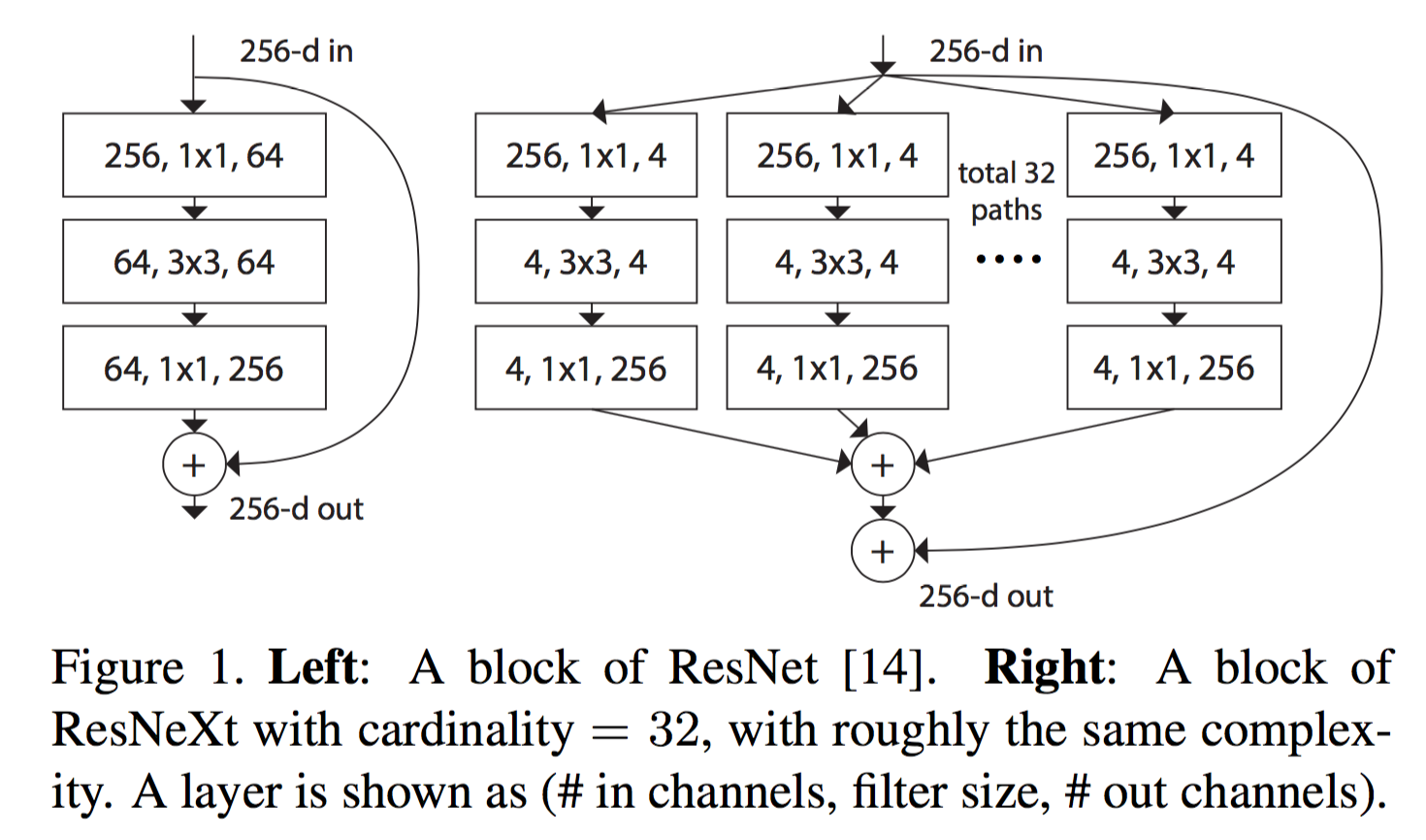
1. 摘要
2. 高度模块化：网络由多个重复的模块组成，每一个模块会聚了一个变换集合，每一个变换的拓扑形式都相同。
3. 除了网络的深度和宽度，本模型引入了一个新的超参数：基数（cardinality），即每一个集合中所含有的变换的数量。
4. 在维持模型复杂度不变的条件下，提升基数可以提高模型准确率；在可以提高模型容量的条件下，提升基数比提高深度和宽度的效果更好。
5. 简介

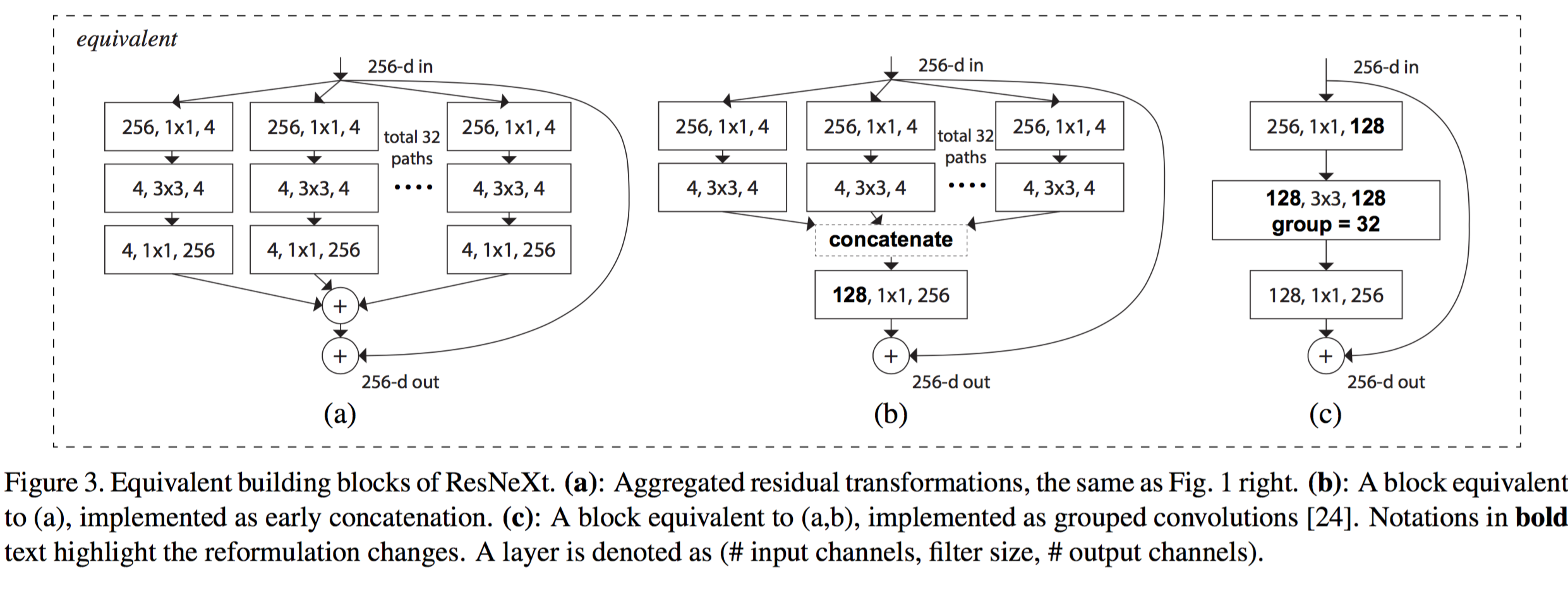
1、随着超参数的数量越来越多，网络的设计越来越困难

2、VGG/ResNet提供了一种简单有效的方法来设计深层网络：把相同结构的单元进行堆叠。这种方法减少了超参数选择的自由性，是的深度成为网络的必要参数；另一方面，这种简化也减少了超参数过度适应特定数据集的风险。

3、Inception模型：使用分解-转换-会聚（split-transform-merge）的方法，证明了通过合理地设计网络的结构也可以以很低的复杂度获得很好的准确率。　Inception模块的解空间是在高维embedding上操作的单个大网络的解空间的严格子集。Inception模块的分解-转换-会聚的行为可以以很低的计算代价获得大型密集网络的表示能力。

4、

我们提出的模型同时使用了：VGG的重复策略、ResNet的短链接策略、以及Inception的分解-转换-会聚策略。每一个模块包含一个变换集合，每一个变换在低维embedding上操作，且具有相同的拓扑结构

5、

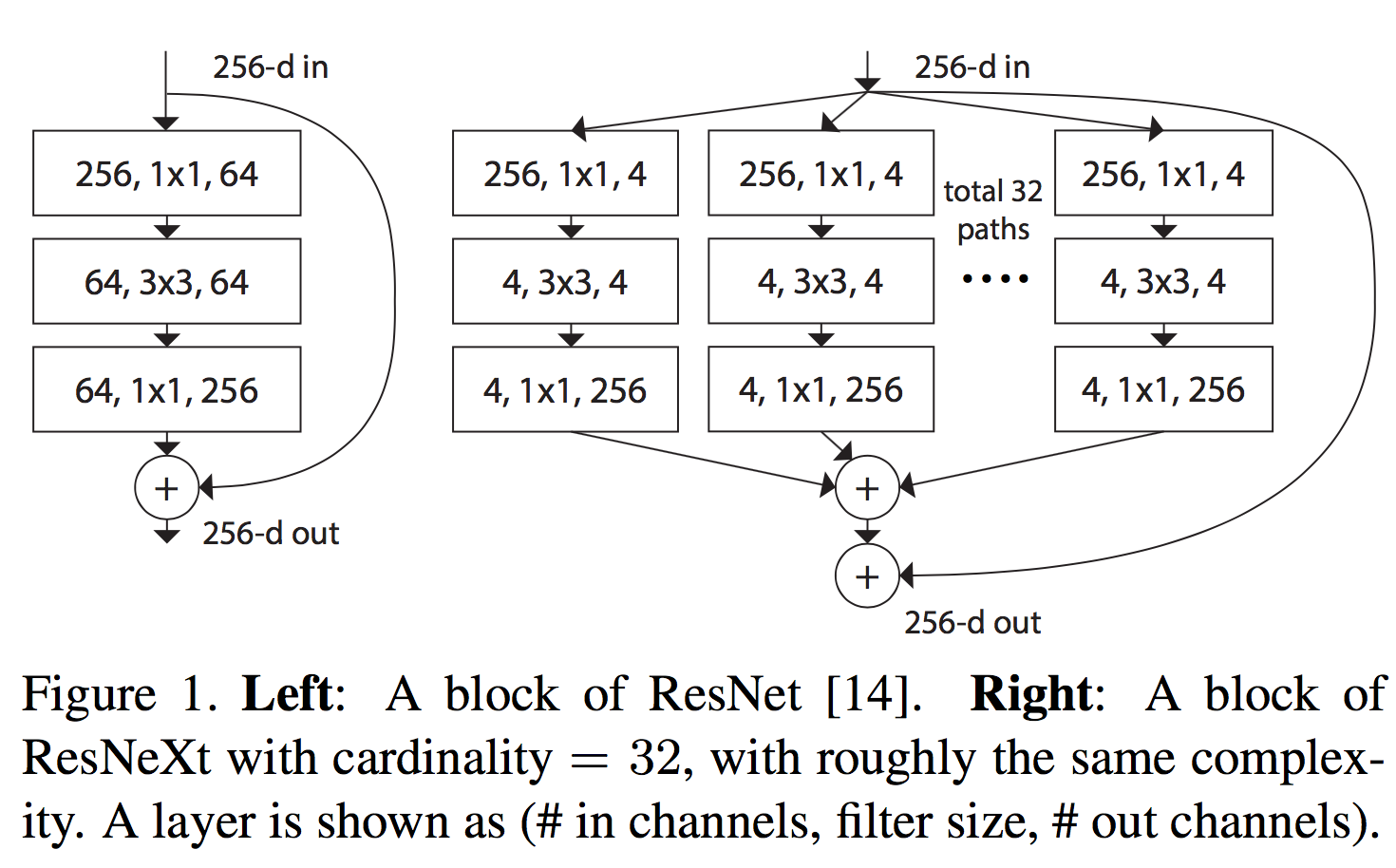
我们模型有两种等价形式，上图中的（b）的结构类似于Inception-ResNet结构，但是不同的是本模型中所有的路径使用相同的拓扑结构，从而可以把路径数作为单独的参数进行研究；上图中的（c）是一种工程上的折衷。

6、在保持计算复杂度和模型大小不变的前提下，本文提出的会聚模块比原始的ResNet的准确率更高。在模型容量不变时，能够提高准确率，说明新的模型在原理上优于原始模型。

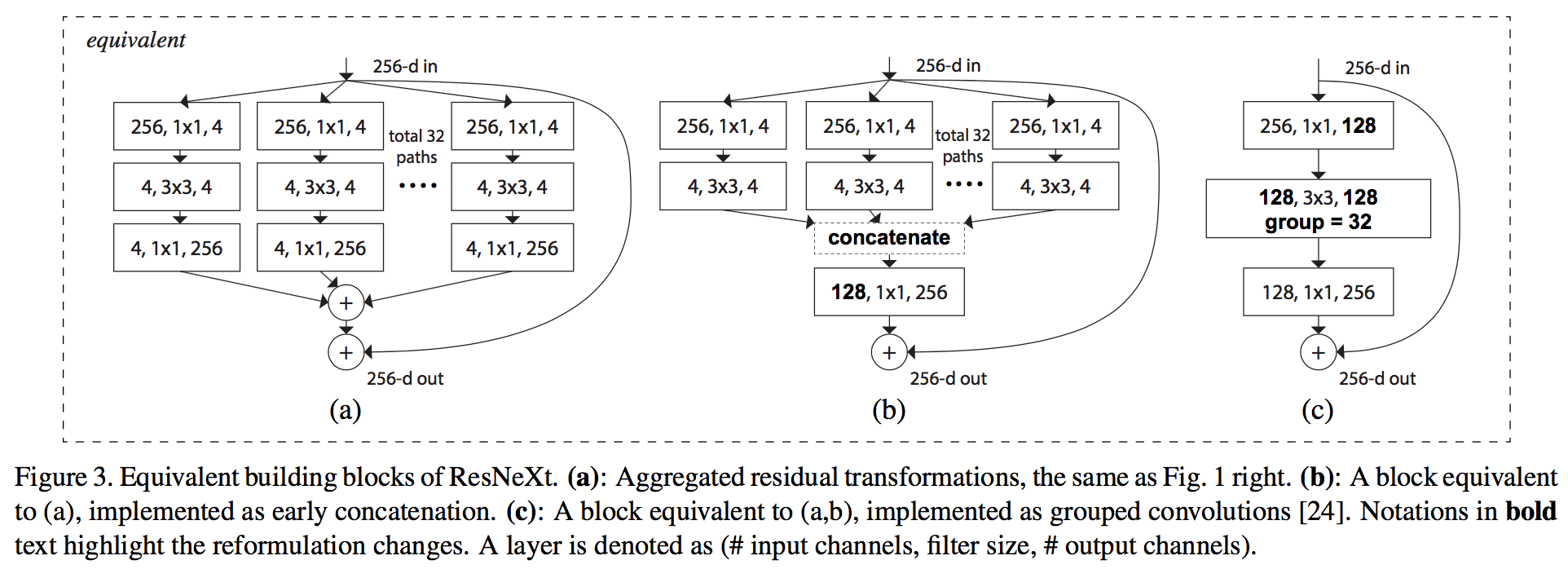
7、本文提出的cardinality维度是具体可度量的，相比深度和宽度而言，提高cardinality能够更有效的提高准确率，尤其当提升深度和宽度开始产生下降的效果时。

1. 方法
2. 模板
3. 本模型使用了一种高度模块化的设计：整个网络由多个残差模块堆叠而成，每一个模块有相同的拓扑结构。模块结构的设计遵循两个原则：一是如果该模块产生的feature map的size不变，则模块的超参数也不变；二是如果该模块产生的feature map的size减半，则模块的宽度要乘以2。这两个原则保证了所有模块的复杂度几乎相同。
4. 根据这两个原则，我们只需要设计一个模板，从而相应地生成所有的模块。由此，这两个原则大大减小了设计空间，让我们可以集中于少量的关键因素。
5. 机构演变过程

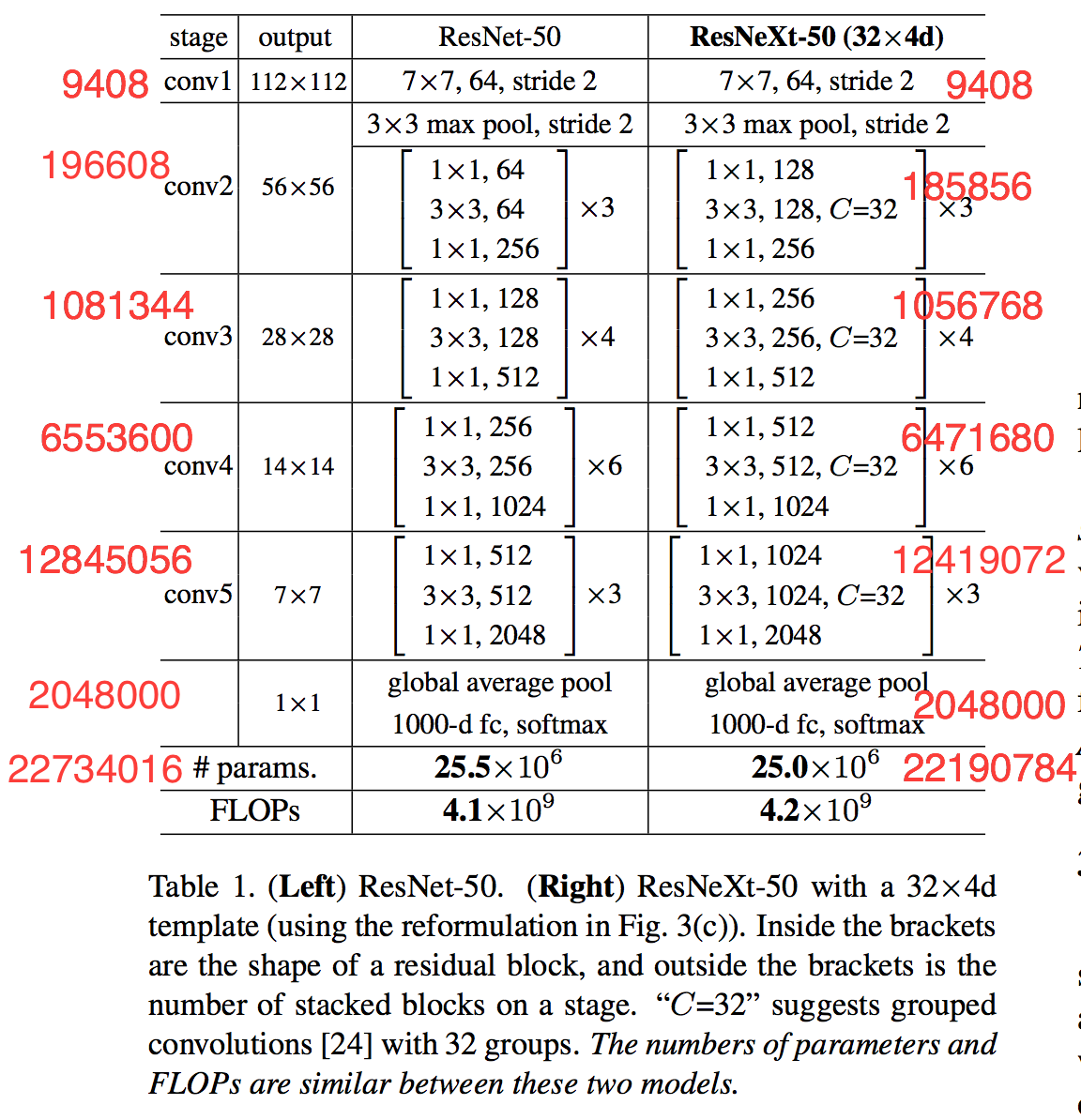
原始模块设计：



等价模块设计：



网络结构设计和参数数量：



1. 通过基本神经元结构对本模型的split-transform-merge结构进行对比解释。一方面说明在保持计算复杂度的前提下，cardinality比深度、宽度这些超参数更有效；另一方面，通过本模型的等价结构，说明和Inception-ResNet、Grouped Convolution的关系。
2. 模型容量
3. 本文通过同时调整cardinality参数和bottleneck宽度来保证模型容量（参数数量和计算复杂度）不变，从而探索cardinality对模型的影响；
4. 实验结果表明，在模型容量不变的条件下，ResNeXt比ResNet的准确率更高；在模型容量允许变化的条件下，提高cardinality比提高深度和宽度更有效。