深度神经网络中的聚合残余变换

采用模型：ResNeXt

**ResNeXt：**ResNeXt是ResNet[2]和Inception[3]的结合体，不同于Inception v4[4]的是，ResNeXt不需要人工设计复杂的Inception结构细节，而是每一个分支都是采用相同的拓扑结构，ResNeXt的本质是分组卷积，通过变量基数来控制组的数量，组卷积是普通卷积和深度可分离卷积的一个折中方案，，每个分支产生的Feature Map的通道数为n.

**ResNet(残差网络)**

ResNet的结构可以极快的加速神经网络的训练，推广性比较好，可以直接用到InceptionNet网络中，主要思想是在网络中增加了直连通道，即Hightway Network的思想，允许原始输入信息直接传到后面的层中，这样的话一层的神经网络可以不用学习整个的输出，而是学习上一个网络输出的残差。

**VGG-nets(深度卷积神经网络)**

VGG网络的提出目的是为了探究大规模图像识别任务中，卷积网络深度对模型精确度有何影响，可以从图像中提取CNN特征。

（CNN主要用来识别位移、缩放极其它形式扭曲不变性的二维图形）

**卷积网络各个层的作用：**

1. 卷积层：用它来进行特征提取。
2. 池化层：对输入的特征图进行压缩，一方面使特征图变小，简化网络计算复杂度；一方面进行特征压缩，提取主要特征。
3. 全连接层：连接所有的特征，将输出值送给分类器。

**超参数：**

在机器学习的上下文中，超参数是在开始学习过程之前设置值的参数，而不是通过训练得到参数数值。通常情况下，需要对超参数进行优化，给学习机选择最优超参数，以提高学习的性能和效果。定义关于模型的更高层次的概念，如复杂性或学习能力，可以通过设置不同的值，训练不同的模型和选择更好的测试值来决定。

**GPU，每个GPU包含32个SGD，SGD 作用：**

主要是为了加速收敛，SGD主要是在每一个样本后都进行参数调节，与所有样本经过网络后再进行参数调节相比，可以加速收敛。

**特征映射：**

在迁移学习中，基于特征映射方法是把各个领域的数据，从原始高维特征空间到低维特征空间下，源领域数据与目标领域数据拥有相同的分布。

**Torch的基本功能：**

Torch：张量的相关运算，eg:创建、索引、切片、连接、转置、加减乘除等相关运算。

它具有CUDA的对应实现，可以在NVIDIA GPU上进行张量运算（计算能力>=3.0）

多维张量的数据结构以及基于基数上的多种数学运算，提供多种使用工具，其中一些可以更有效的对张量和任意类整理进行序列化工具。

**NVIDIA 的8个GPU：**人工智能计算公司，1999推出GPU，重新定义了现代计算机图形技术，彻底改变并行计算。

**CU DA:**是显卡厂商DVIDIA推出的运算平台，是一种通用并行计算架构，该架构使GPU能 够解决复杂的计算问题。

CIFAR：CIFAR数据集常以CIFAR-10命名，共包含6w张32\*32的彩色图像，（包含5w张训练图片，1w张测试图片），其中没有任何类型重叠的情况。

1. CNN（Region-cnn）是第一个成功将深度学习应用到目标检测上的算法，基于卷积神经网络（CNN），线性回归和支持向量机（SVM）等算法实现目标检测技术。

全局平均池：将特征图所有像素值相加求平局，得到一个数值，即该数值表示对应特征图。目的代替全连接层，效果是减少参数数量，减少计算量，减少过拟合。

卷积：分析数学的一种运算，通过两个函数f和g生成第三个函数的一种数学算子，表征f和g经过翻转和平移的重叠部分函数值乘积对重叠长度的积分。

1\*1的卷积层作用：1.升降维2.加入非线性3.跨通信信息交互

全连接层：全连接层的每一个结点都与上一层的所有结点相连，用来把前边提取到的特征综合起来。

池化层的作用：

1. 上采样：把图片放大，在图片识别过程中，需要对图形进行像素级别的分类，因此，在卷积提取特征后需要通过上采样将Feature Map还原到原图中。
2. 下采样：图像的缩小过程。
3. 降维：去除冗余信息，对特征进行压缩，简化网络复杂度，减小计算量，减小内存消耗等等，总的理解就是减少参数量。
4. 实现非线性。
5. 扩大感受野。（感受野：是卷积神经网络cnn每一层输出的特征图上的像素点在原始输入图像上映射的区域大小。）
6. 实现不变性：平移不变性、旋转不变性、尺度不变性。

池化主要有哪几种：

1. Max Polling(最大池化)：选图像区域的最大值作为该区域池化后的值，主要功能是下采样，不会损失识别结果，这意味着卷积后的feature map 中有对于识别物体不必要的冗余信息。
2. Average Polling（平均池化）计算图像区域的平均值作为该区域池化后的值。

**消融实验：**

**定义：**通常用于神经网络，尤其是相对复杂的神经网络，如R-CNN，我们的想法是通过删除部分网络并研究网络的性能来了解网络。

“消融”的原始含义是手术切除身体组织。在机器学习中，特别是复杂的深度神经网络的背景下，已经采用“消融研究”来描述去除网络的某些部分的过程，以便更好的理解网络的行为。

Val集：主要作用是来验证是否过拟合，以及用来调节训练参数等。

ReLU函数：

线性整流函数（修正线性单元），是一种人工神经网络中常用的激活函数，通常指代以斜坡函数及其变种为代表的非线性函数。

BN（BatchNorm）：在深度神经网络训练过程中使得每一层神经网络的输入保持相同分布的。

本论文中将ResNeXt-101与ResNet-101、ResNet-200、Inception-v3[39]、Inception-v4[37]、Inception-ResNet-v2[37]分别在224\*224、320\*320、299\*299的尺度上进行对比，显示错误率更低，效果更好一些。通过设置参数c（1~32）与d的值，最终得到c=32时ResNeXt的错误率最低。