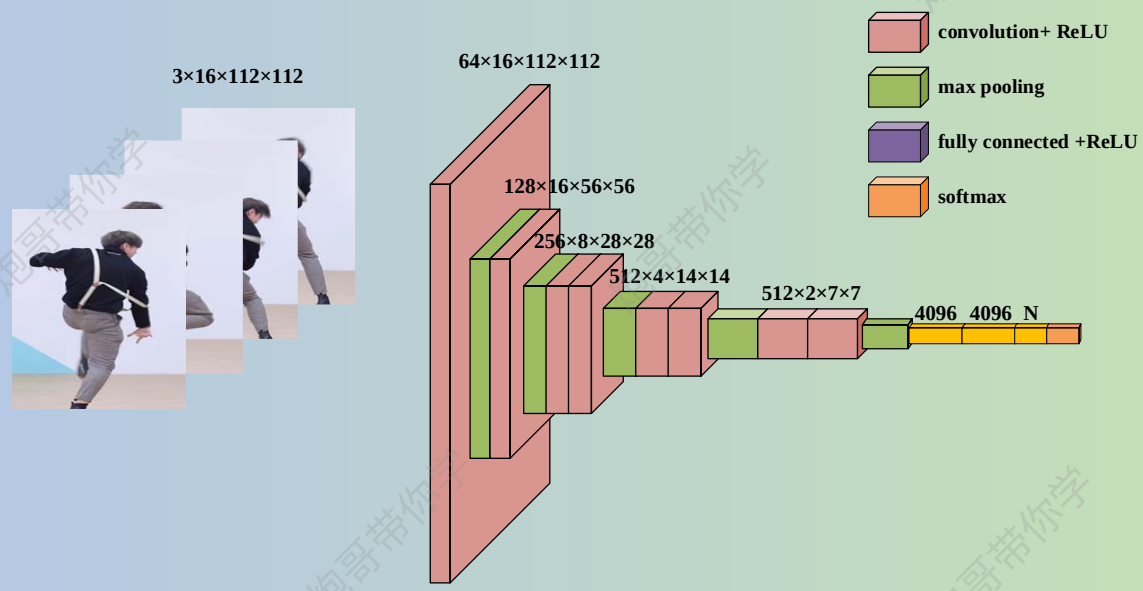
PyTorch



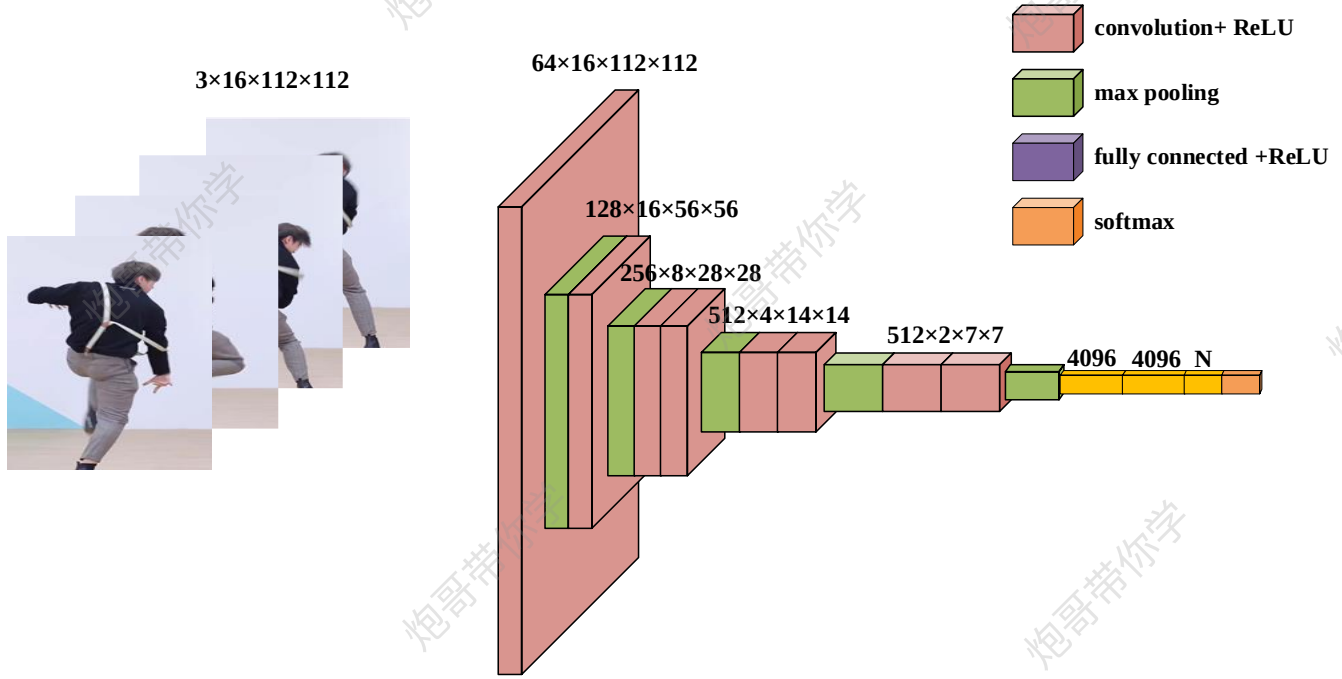
第4章 C3D模型与行为识别实战

C3D网络诞生背景

发明者：C3D是由Google 的研究人员

提出时间： 2014年提出的

研究目的：应用包括视频分类、行为识别、视频描述等领域。

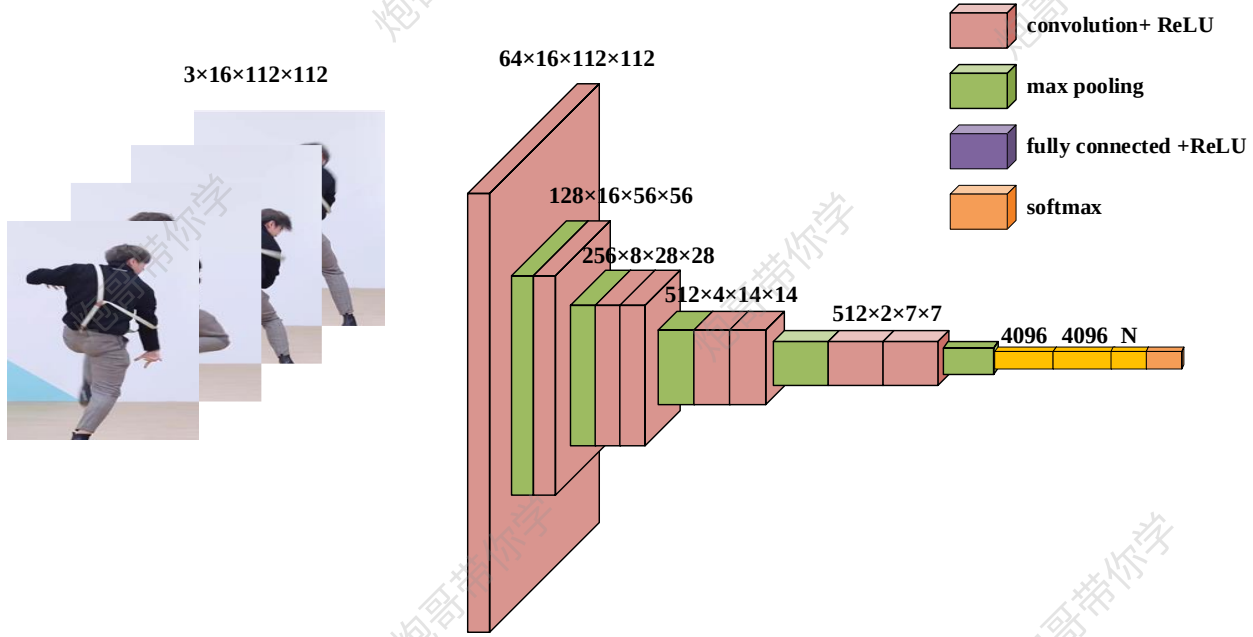


C3D网络结构

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Table 2: **Number of parameters** (in millions).

Network	A,A-LRN	B	C	D	E
Number of parameters	133	133	134	138	144



C3D网络结构

经典卷积神经网络的基本组成部分是下面的这个序列：

(1)带填充以保持分辨率的卷积层；

(2)非线性激活函数，如ReLU；

(3)池化层，最大池化层。

总结就是，一系列卷积层组成，后面再加上用于空间下采样的最大池化层。

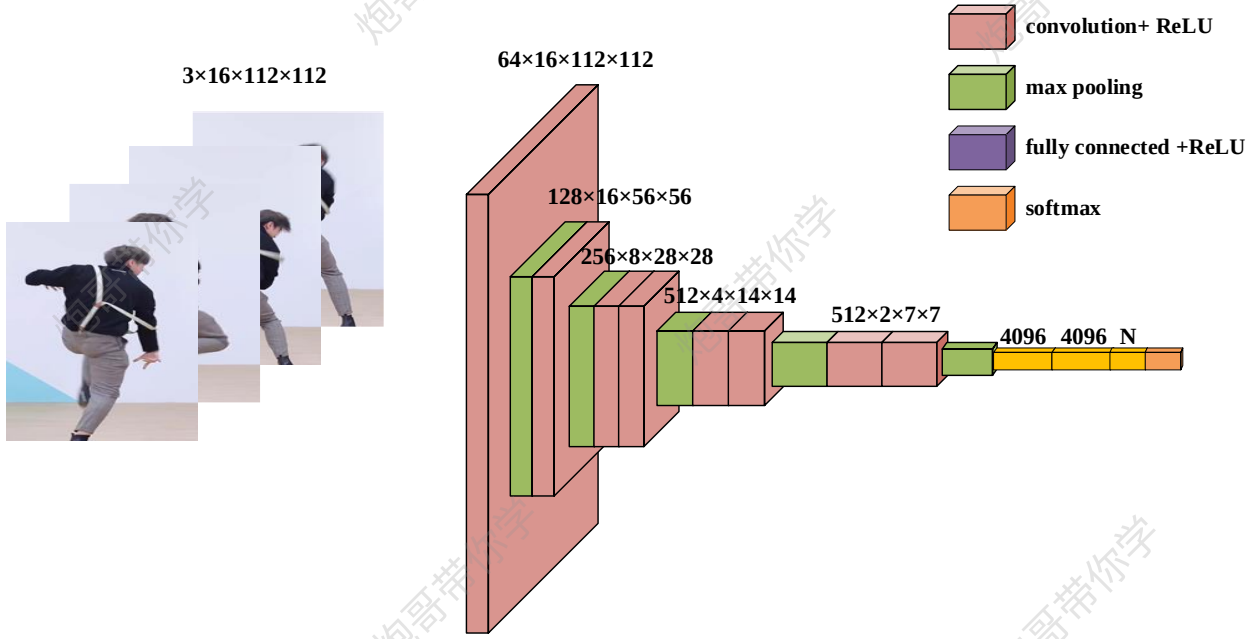
VGG特点：

vgg-block内的卷积层都是同结构的

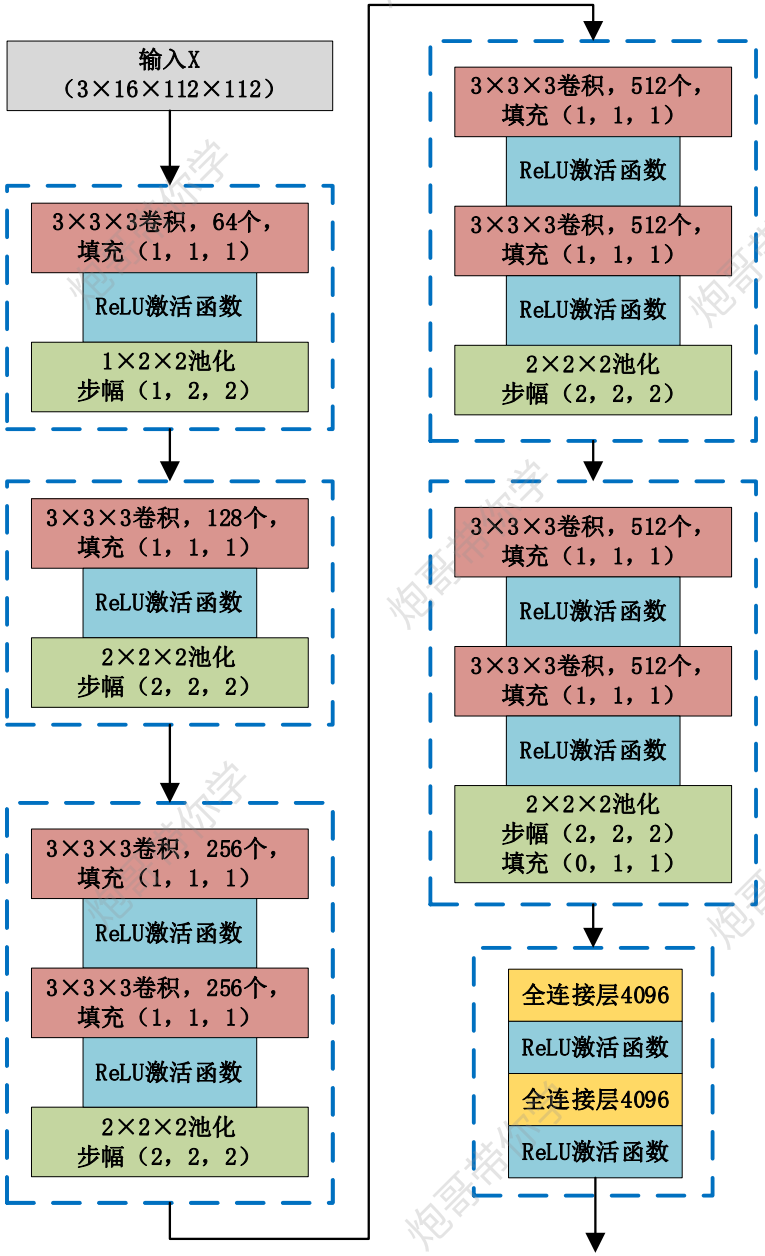
池化层都得上一层的卷积层特征缩减一半

深度较深，参数量够大

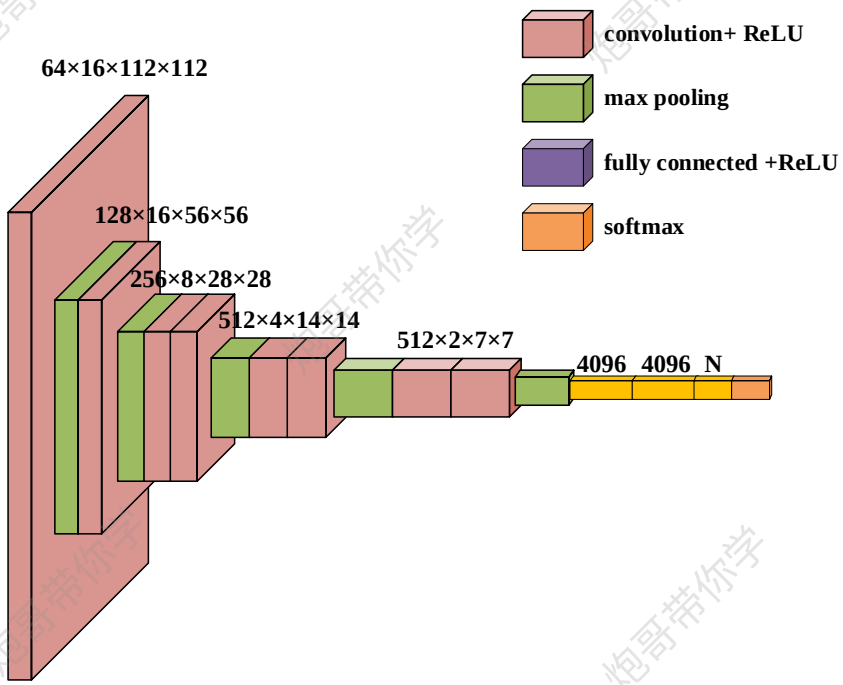
较小的filter size/kernel size



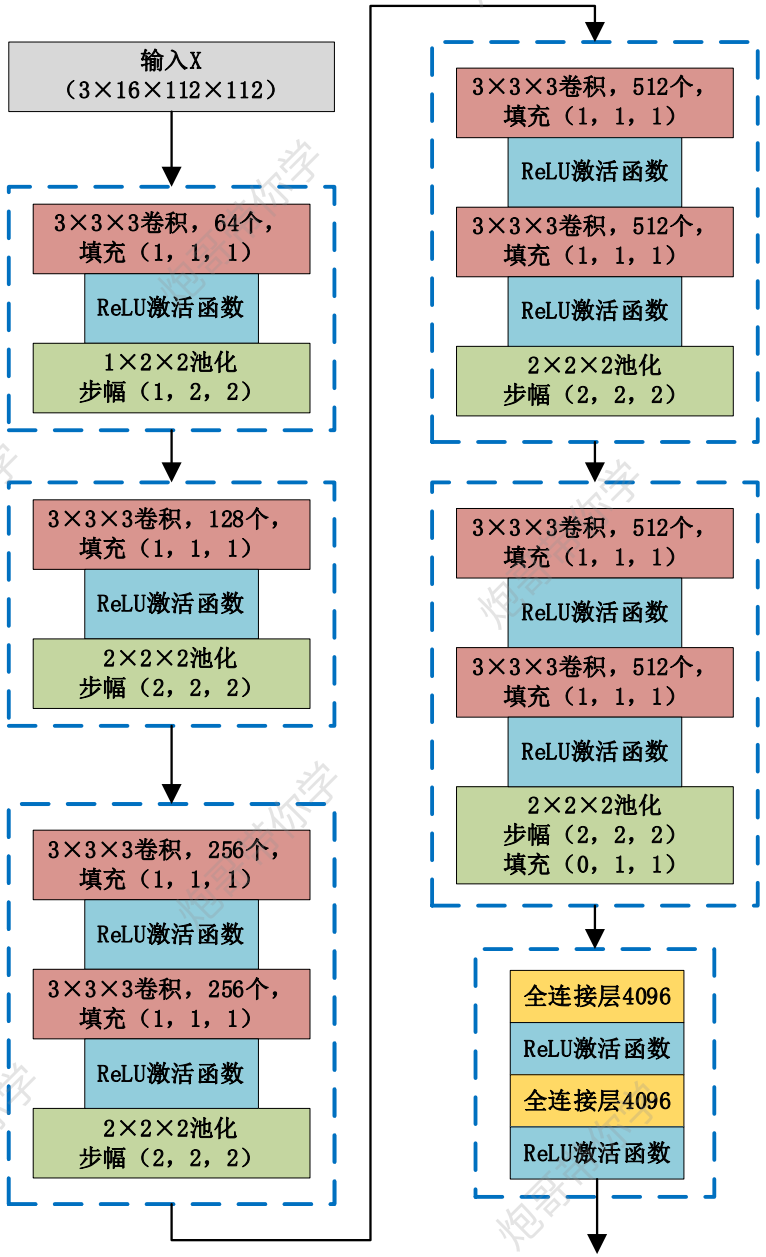
C3D网络参数详解


输入X
（3×16×112×112）
3×3×3卷积，64个，
填充（1，1，1）
ReLU激活函数
1×2×2池化
步幅（1，2，2）
3×3×3卷积，128个，
填充（1，1，1）
ReLU激活函数
2×2×2池化
步幅（2，2，2）
3×3×3卷积，256个，
填充（1，1，1）
ReLU激活函数
3×3×3卷积，256个，
填充（1，1，1）
ReLU激活函数
2×2×2池化
步幅（2，2，2）
3×3×3卷积，512个，
填充（1，1，1）
ReLU激活函数
3×3×3卷积，512个，
填充（1，1，1）
ReLU激活函数
2×2×2池化
步幅（2，2，2）
3×3×3卷积，512个，
填充（1，1，1）
ReLU激活函数
3×3×3卷积，512个，
填充（1，1，1）
ReLU激活函数
2×2×2池化
步幅（2，2，2）
填充（0，1，1）
全连接层4096
ReLU激活函数
全连接层4096
ReLU激活函数


3×16×112×112


64×16×112×112
128×16×56×56
256×8×28×28
512×4×14×14
512×2×7×7
4096
4096
N
convolution+ ReLU
max pooling
fully connected +ReLU
softmax

C3D网络参数详解



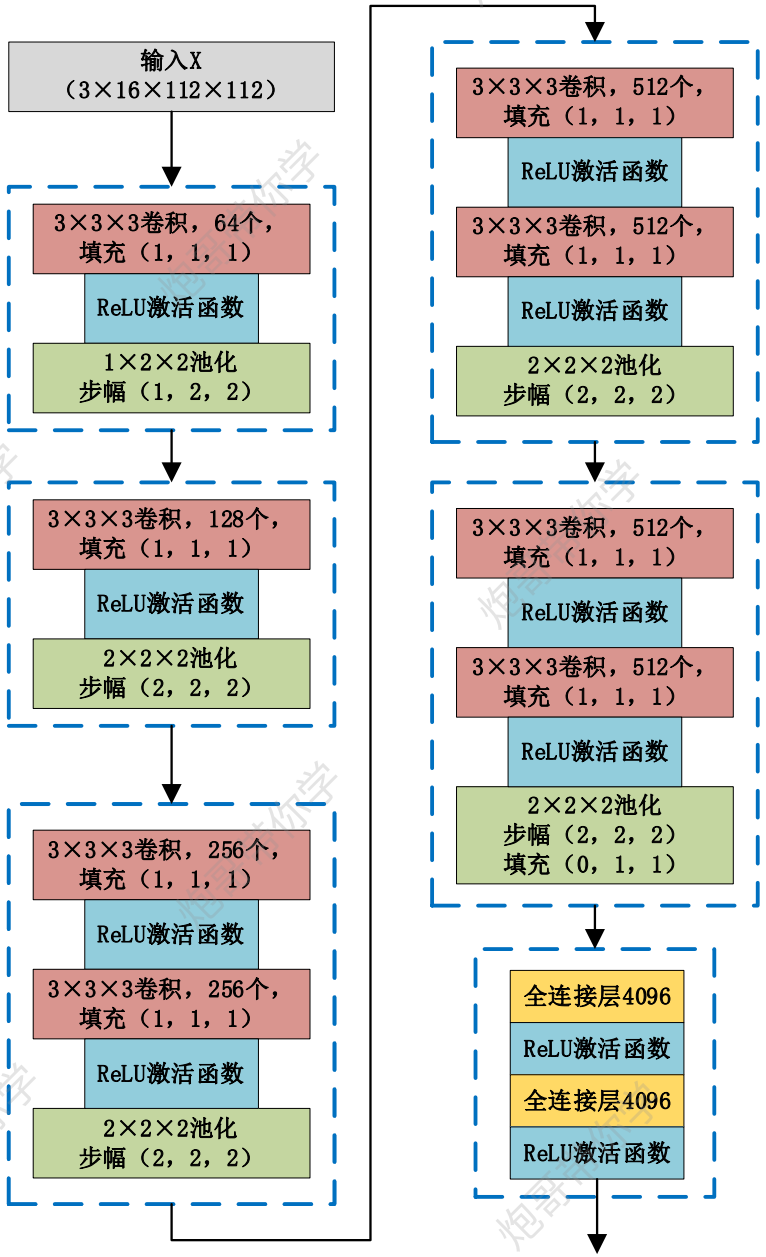
第1层输入层：输入为3×16×112×112的视频帧。

第1个块：

（1）输入为3×16×112×112，卷积核数量为64个；卷积核的尺寸大小为3×3×3×3；，填充为（1，1，1）；卷积后得到shape为64×16×112×112的特征图输出。

（2）输入为64×16×112×112，池化核为1×2×2，步幅为（1，2，2），池化后得到尺寸为64×16×56×56的特征图。

C3D网络参数详解

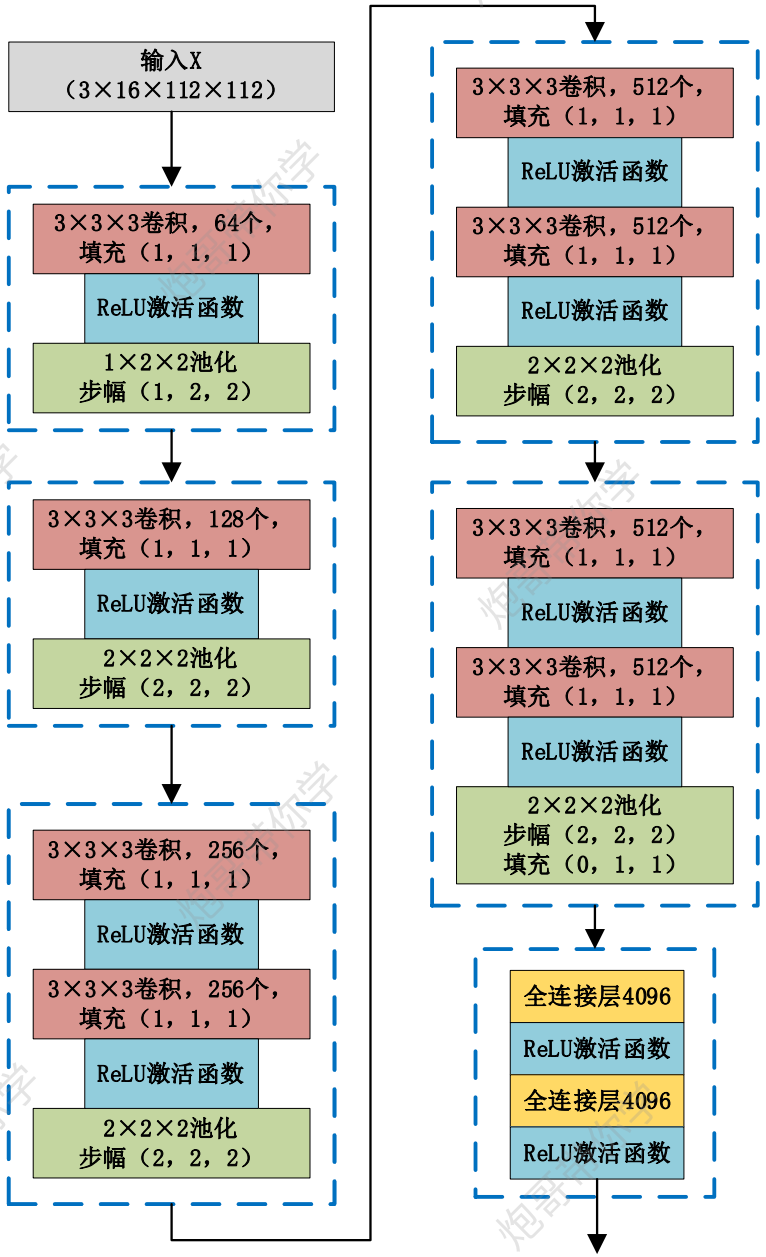


第2个块：

（1）输入为64×16×56×56，卷积核数量为128个；卷积核的尺寸大小为64×3×3×3；，填充为（1，1，1）；卷积后得到shape为128×16×56×56的特征图输出。

（2）输入为128×16×56×56，池化核为2×2×2，步幅为（2，2，2），池化后得到尺寸为128×8×28×28的特征图。

C3D网络参数详解



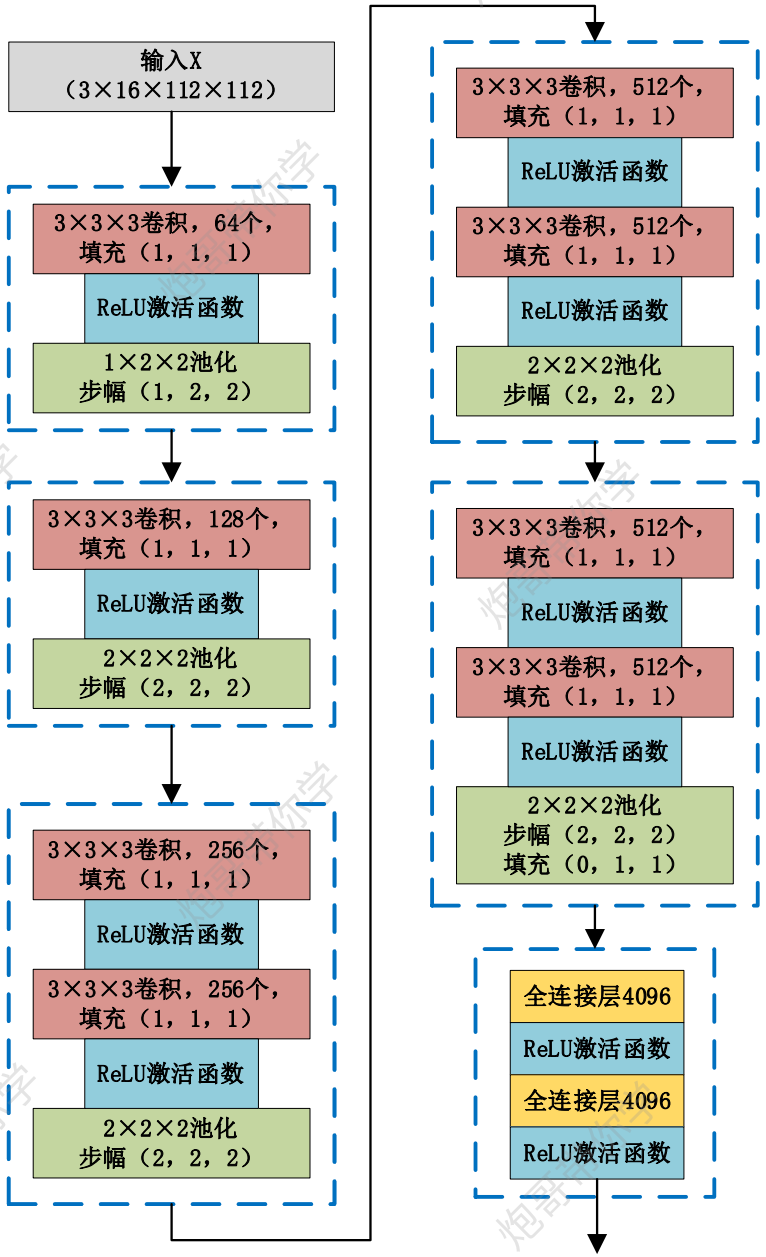
第3个块：

（1）输入为128×8×28×28，卷积核数量为256个；卷积核的尺寸大小为128×3×3×3；填充为（1，1，1）；卷积后得到shape为256×8×28×28的特征图输出。

（2）输入为256×8×28×28，卷积核数量为256个；卷积核的尺寸大小为256×3×3×3；填充为（1，1，1）；卷积后得到shape为256×8×28×28的特征图输出。

（3）输入为256×8×28×28，池化核为2×2×2，步幅为（2，2，2），池化后得到尺寸为256×4×14×14的特征图。

C3D网络参数详解



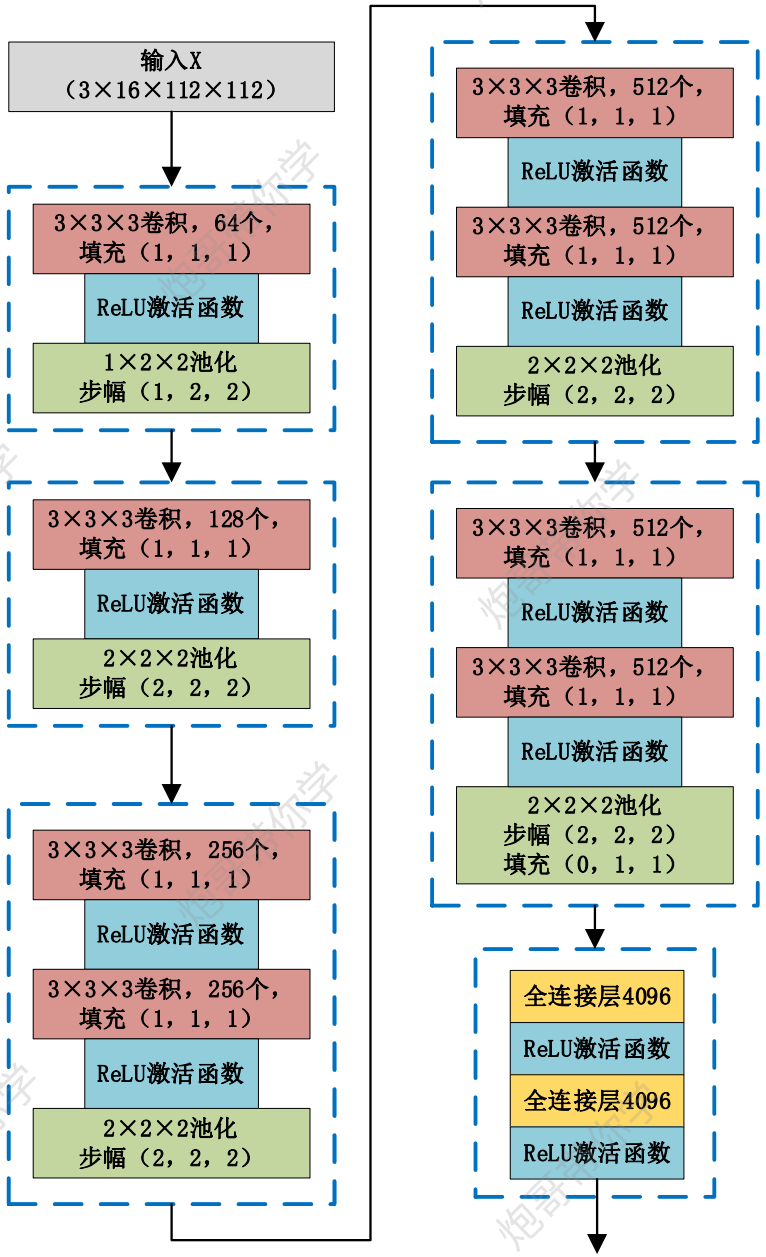
第4个块：

（1）输入为256×4×14×14，卷积核数量为512个；卷积核的尺寸大小为256×3×3×3；，填充为（1，1，1）；卷积后得到shape为512×4×14×14的特征图输出。

（2）输入为512×4×14×14，卷积核数量为512个；卷积核的尺寸大小为512×3×3×3；，填充为（1，1，1）；卷积后得到shape为512×4×14×14的特征图输出。

（3）输入为512×4×14×14，池化核为2×2×2，步幅为（2，2，2），池化后得到尺寸为512×2×7×7的特征图。

C3D网络参数详解



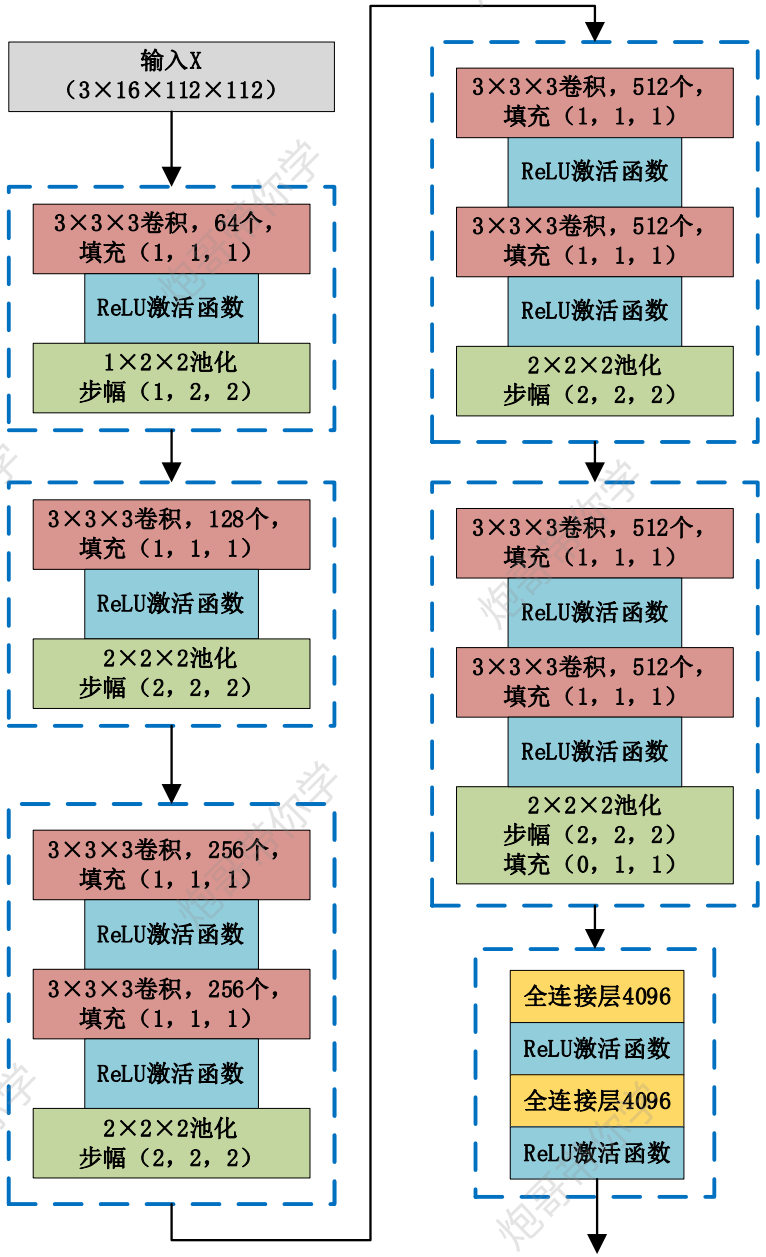
第5个块：

（1）输入为512×2×7×7，卷积核数量为512个；卷积核的尺寸大小为512×3×3×3；，填充为（1，1，1）；卷积后得到shape为512×2×7×7的特征图输出。

（2）输入为512×2×7×7，卷积核数量为512个；卷积核的尺寸大小为512×3×3×3；，填充为（1，1，1）；卷积后得到shape为512×2×7×7的特征图输出。

（3）输入为512×2×7×7，池化核为2×2×2，步幅为（2，2，2），填充为（0，1，1），池化后得到尺寸为512×1×4×4的特征图。

C3D网络参数详解



第1~3层全连接层：

首先进行维度转化将512×1×4×4转化为（8192，1）；

第1层全连接层，神经元4096。使用relu激活函数后还使用了Dropout。

第2层全连接层，神经元4096。使用relu激活函数后还使用了Dropout。

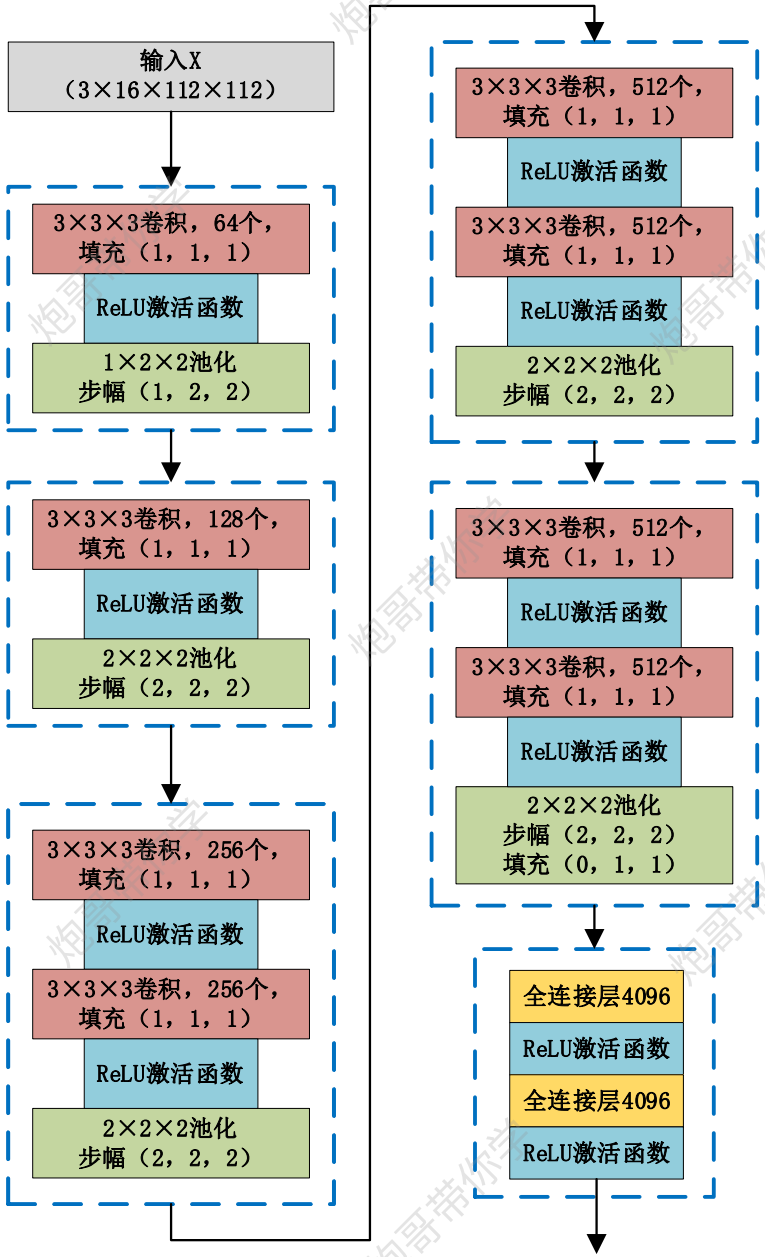
第3层全连接层，神经元输出类别数。

C3D网络总结

C3D网络总结：

1、采用VGG网络架构，因此VGG该有的优势，C3D网络网络结构也都有。

2、应用3D卷积运算从视频数据中提取空间和时间特征以进行动作识别。 这些3D特征提取器在空间和时间维度上操作，从而捕获视频流中的运动信息。

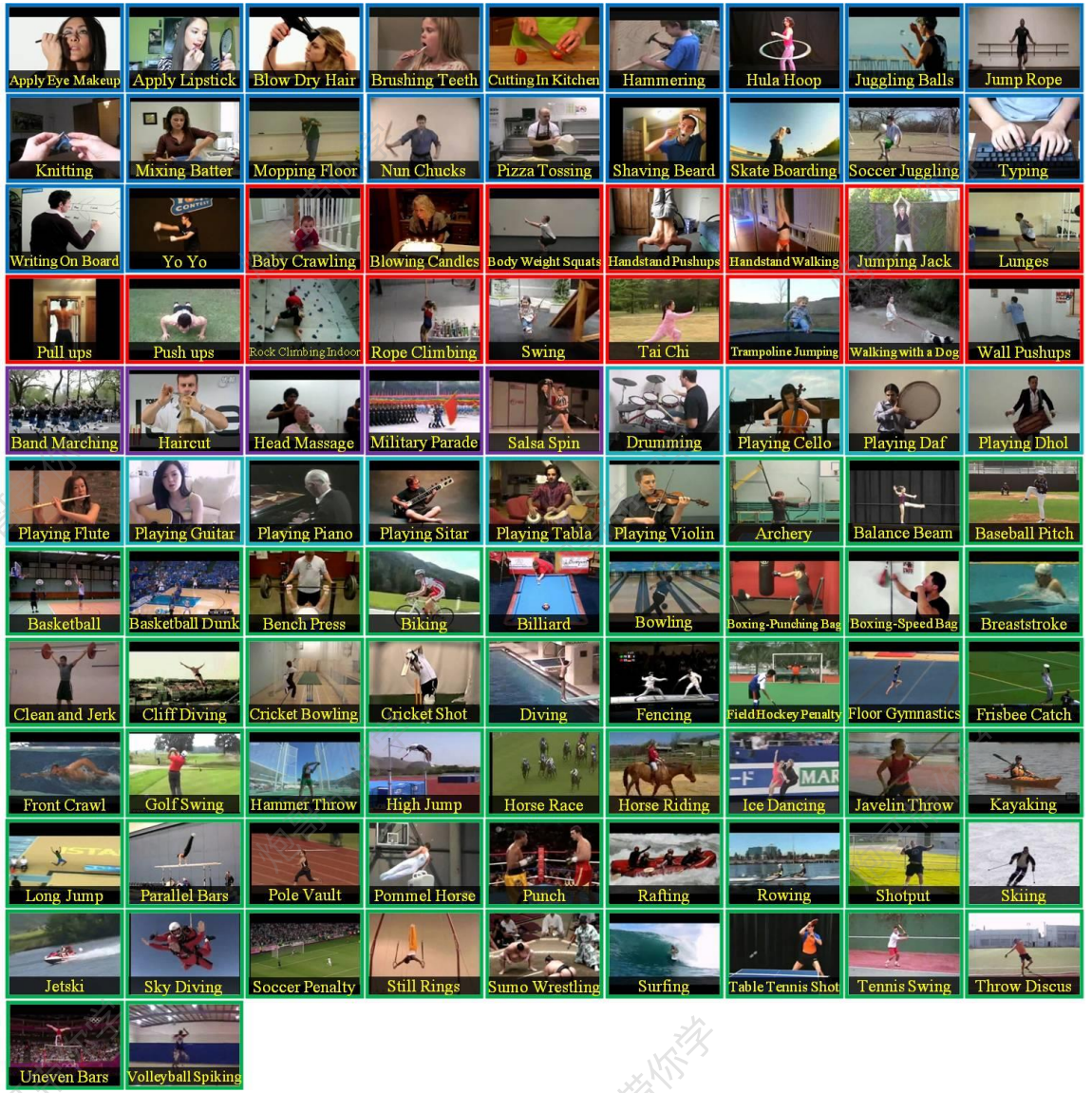


数据集介绍

数据集简介：

UCF101是一个现实动作视频的动作识别数据集，收集自YouTube，提供了来自101个动作类别的13320个视频。官方网站

- 数据集名称：UCF-101（2012）
- 总视频数：13,320个视频
- 总时长：27个小时
- 视频来源：YouTube采集
- 视频类别：101 种
- 主要包括5大类动作：人与物体交互，单纯的肢体动作，人与人交互，演奏乐器，体育运动
- 每个类别（文件夹）分为25组，每组4~7个短视频，每个视频时长不等。



数据集介绍

数据集标签：

涂抹眼妆，涂抹口红，射箭，婴儿爬行，平衡木，乐队游行，棒球场，篮球投篮，篮球扣篮，卧推，骑自行车，台球射击，吹干头发，吹蜡烛，体重蹲，保龄球，拳击沙袋，拳击速度袋，蛙泳，刷牙，清洁和挺举，悬崖跳水，板球保龄球，板球射击，在厨房切割，潜水，打鼓，击剑，曲棍球罚款，地板体操，飞盘接球，前爬网，高尔夫挥杆，理发，链球掷，锤击，倒立俯卧撑，倒立行走，头部按摩，跳高，跑马，骑马，呼啦圈，冰舞，标枪掷，杂耍球，跳绳，跳跃杰克，皮划艇，针织，跳远，刺，阅兵，混合击球手，拖地板，修女夹头，双杠，披萨折腾，弹吉他，弹钢琴，弹塔布拉琴，弹小提琴，弹大提琴，弹Daf，弹Dhol，弹长笛，弹奏锡塔琴，撑竿跳高，鞍马，引体向上，拳打，俯卧撑，漂流，室内攀岩，爬绳，划船，莎莎旋转，剃胡子，铅球，滑板溜冰，滑雪，Skijet，跳伞，足球杂耍，足球罚球，静环，相扑摔跤，冲浪，秋千，乒乓球拍，太极拳，网球秋千，投掷铁饼，蹦床跳跃，打字，高低杠，排球突刺，与狗同行，墙上俯卧撑，在船上写字，溜溜球。剃胡须，铅球，滑冰登机，滑雪，Skijet，跳伞，足球杂耍，足球罚款，静物环，相扑，冲浪，秋千，乒乓球射击，太极拳，网球秋千，掷铁饼，蹦床跳跃，打字，不均匀酒吧，排球突刺，与狗同行，壁式俯卧撑，船上写字，溜溜球。剃胡须，铅球，滑冰登机，滑雪，Skijet，跳伞，足球杂耍，足球罚款，静物环，相扑，冲浪，秋千，乒乓球射击，太极拳，网球秋千，掷铁饼，蹦床跳跃，打字，不均匀酒吧，排球突刺，与狗同行，壁式俯卧撑，船上写字，溜溜球