1. 重训练流程
2. 首先对输入特征图进行剪枝并进行重训练，生成模型权重文件-.pth文件。
3. 其次针对上述.pth文件进行卷积核聚类剪枝，得到固定形状的权重分布，再进行重训练，此时的重训练也许带上输入特征图的剪枝，重训练的目的是改变卷积核中固定形状的权重的值。

二、main函数中变量含义与接口功能。

1. model\_name - 模型名称，举例:’LeNet’、’AlexNet’。
2. batch-size – 批处理图片数量。
3. img\_size – 单张图片大小。
4. ratio – 输入特征图剪枝率。
5. epochA – 针对输入特征图剪枝设定的重训练轮数或初始模型的训练轮数。
6. epochAW – 针对权重卷积核聚类剪枝设定的重训练轮数。
7. patternA – 针对输入特征图剪枝的三种类型的操作：
   1. patternA=’train’:训练初始模型，不添加任何剪枝操作，得到的权重文件存储在路径’./pth/model\_name/ratio=0/Activation/best.pth’中。
   2. patternA=’test’：测试网络模型精度。
      1. ratio=0，测试不添加任何剪枝操作的模型的精度。
      2. ratio!=0，测试特定的对输入特征图剪枝操作并重训练后的模型的精度。

c） pattern=‘retrain’:针对输入特征图剪枝后的模型进行重训练。

8. patternW – 针对卷积核聚类剪枝的两种类型的操作：

a) patternW=‘retrain’，在对经过输入特征图剪枝并重训练后的卷积核进行聚类剪枝并进行重训练。

b) patternW=’test’，测试经过输入特征图剪枝和权重聚类剪枝后模型的整体精度。

9. weightParameter – 进行卷积层聚类剪枝时剪枝参数。

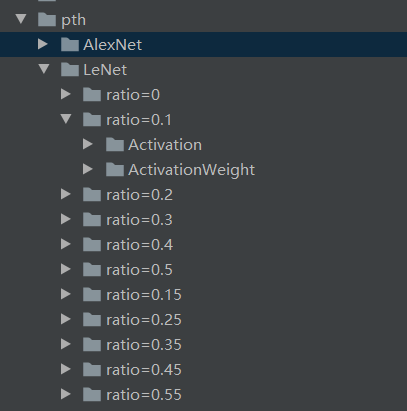
10.LinearParameter – 进行线性层聚类剪枝时剪枝参数。

11. 函数activationPruneModelOp - 对输入特征图进行剪枝。

12. 函数weightPruneModelOp – 对卷积核进行聚类剪枝。

三、权重文件的标准化保存目录

1. 以LeNet网络为例，如下所示

一级目录为pth，二级目录为模型名称，三级目录为ratio=0.x，四级目录包括两个并行目录，分别是Activation和ActivationWeight，前者主要存储只经过输入特征图剪枝并重训练过后的权重文件，后者存储的则是经过输入特征图剪枝与卷积核聚类剪枝并重训练过后的权重文件。

四、操作步骤

在训练时只需要确定main函数中前10个参数的值，已实现了标准目录自动生成与后续训练或测试过程中的模型文件的自动索引，无需考虑该方面的内容。如果是重训练的话，只需要将patternA和patternAW设置为’retrain’即可。