

# 持续学习大作业说明文档

## 一、要求

### 1、前期准备

- (1) 深度学习基础知识、深度学习框架的搭建和使用。
- (2) 理解和学习持续学习概念和基本算法（如 iCaRL、EWC、LwF 等）。

### 2、目标

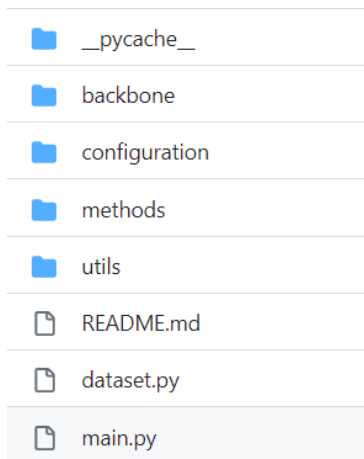
- (1) 根据提供的论文列表、持续学习框架，复现 1 个经典持续学习算法（复现多个可加分）。
- (2) 完成实验报告，汇报复现过程、实验结果、遇到的问题及尝试的解决办法，报告无字数要求，表明复现过程中的思考、尝试即可（能复现出论文精度最好，如果结果精度有一定差距，可以给出可能的原因以及调试过程）。
- (3) 作业提交：实验报告、代码、训练日志打包，发送至 [nju\\_ml@163.com](mailto:nju_ml@163.com)。

### 3、持续学习框架：<https://github.com/yizhibaiwuya/MLwork>

### 4、论文范围

- [1] (ECCV 2016) Learning without Forgetting
- [2] (CVPR 2017) iCaRL: Incremental Classifier and Representation Learning
- [3] (PNAS 2017) Overcoming catastrophic forgetting in neural networks
- [4] (ECCV 2018) End-to-End Incremental Learning
- [5] (CVPR 2019) Large Scale Incremental Learning
- [6] (CVPR 2020) Maintaining Discrimination and Fairness in Class Incremental Learning
- [7] (CVPR 2020) Memory-Efficient Incremental Learning Through Feature Adaptation
- [8] (CVPR 2020) Semantic Drift Compensation for Class-Incremental Learning
- [9] (CVPR 2021) Rainbow Memory: Continual Learning with a Memory of Diverse Samples

## 二、框架介绍



## 1、文件组成

(1) **configuration** 文件夹：用于设置模型参数，如果所复现的方法需要添加其他参数，加在 **config.py** 中。

(2) **backbone** 文件夹：提供 ResNet-18 模型。

(3) **logs** 文件夹（训练时自动创建）：用于存储训练日志，最后需要提交最终的模型训练日志。

(4) **methods** 文件夹：用于保存复现的方法。

## 2、环境配置

(1) 代码环境：Python3.8 及以上，[PyTorch](#)

(2) 数据集：CIFAR10、CIFAR100

## 3、复现内容

### (1) 新建文件

在 **methods** 文件夹下创建一个新文件来复现方法，如 **methods/icarl.py**。该文件中需要实现一个类，如 **ICARL**，其继承 **methods/finetune.py** 文件中的 **Finetune** 类（**Finetune** 类实现了一个最基本的 **replay** 模型）。

### (2) 实现训练流程

手动实现下面四个函数（如果与 **Finetune** 类的函数相同，可直接复用）：

- **before\_task**：负责在每个任务训练前，执行更换分类头、重置优化器、赋值变量等操作。
- **train**：负责利用 **train loader** 进行训练，控制训练过程。
- **evaluation**：负责利用 **test loader** 进行测试。
- **after\_task**：在每个任务结束时更新内存中保存的样本（**replay** 模型）。

### (3) 设置所需的模型参数

在 **configuration/config.py** 中设置实验所需参数，比如：

- 数据集 **dataset**：CIFAR10、CIFAR100
- 数据集路径 **dir**：比如数据集文件路径是 **/home/xyk/CIFAR100**，那么 **dir = "/home/xyk"**
- 保存旧样本的数量 **memory\_size**
- 模型 **backbone**：resnet12
- 优化器相关参数：**opt\_name**、**sched\_name**
- 训练参数：训练 **epoch** 数目 **n\_epoch**、学习率 **lr**、**batch\_size**

### (4) **main.py**

- **line24**：**save\_path** 设置保存训练日志的路径，最终会存放在 **./logs/DATASET/save\_path** 路径下。
- **line70**：**class\_of\_task** 为每个任务划分多少类别，比如为 CIFAR100 划分为 **[20, 20, 20, 20, 20]**，保证总和不超过总的类别数（CIFAR10 为 10，CIFAR100 为 100），每个任务的类别数应大于 0。
- **line78**：**method** 实例化你所复现的方法。

#### 4、其他

关于数据集下载、其他说明请参考 [Github](#) 上的说明文档。