

# Colaboratory

## 简介：

Colab 是谷歌提供的可供外部开发者使用的一个开发平台。

## 1.其包括三大运算资源.

CPU 适合拿来做数据预处理工作会比GPU快

GPU 适合拿来做深度学习运算的工作，一般提供单机单卡K80s, T4s, P4s, P40,P100,V100计算资源，16G显存。中高等大小image based model足够， 中等大小video based modal够用。

TPU 适合跑专门的计算项目 可用资源少，需要程序适配（不考虑）

## 2.操作系统：

18.04版本的ubuntu系统，常规常用的软件git,vim,docker都用，不需要单独配置。如果运用GPU资源，其cuda版本为固定11.2，nvcc -V为11.0. 这个cuda配置较高，不过目前torch1.9和tensorflow都是适配的，TF已经安装，无需新安装。

## 3.权限设置：

每次连接后端服务器会分配一台云服务器，拥有管理员权限，可sudo，无密码。

## 4.储存：

colab的储存主要分为两个板块，云服务器储存，改区域的储存是每次连接后端自动分配的，连接机器变化后会被清除（不建议存放具有较强time dependency的文件如数据集，代码等）。谷歌云盘储存，需要挂载云盘从而获取云盘上储存的数据，可存放长期数据。挂载一次后会自动挂载。

## 5.谷歌云盘的数据上传下载：

上传速度非常取决于vpn质量和网络带宽。不会太慢，5g数据稍微快点的vpn就几分钟。如果网络不稳定会很久，且容易失败。下载速度相对稳定，相较慢一点。（但一般也不需要下载太多东西）

## 优点：

1. 不用实体的机器，避免了一些库的安装和编译，特别是opencv-contrib,ffmpeg这种。也不需要anaconda这种管理环境但新手容易遇见各种error的env 管理器。基本在colab pip 解决所有问题。大多深度学习和机器学习的库都有（我只遇见过一次tensorrt库的自动编译失败的问题）

2. 假设环境真的很难配，配毁了直接restart 后端连接分配新gpu资源 (Doge+1)
3. 专业版十美元一个月每个账户，蛮便宜的。专业版同时可开四个进程，至多两三个V100 GPU(通常一张)，其余为P100。推理速度大概P100 1.75倍实测.计算资源还是不错的
4. jupyter based 主要用.ipynb 文件即notebook文件，对初学者可视化和debug都比较友好。（个人喜欢拿来做数据集预处理）
5. 还有就是nb文件的特性了，魔法命令执行命令行非常方便，基本不用开terminal就算要run .py文件。
6. 7.国外服务器，下载安装所需环境和依赖很快

## 缺点：

- 1.需要翻墙，需要谷歌账号
- 2.有一定的每个会话的时间限制，具体记不得了8h+以上吧（基本是够用了，除非做一些大数据集科研项目了，应该没有这个需求的）。