

ZAI-1.18 版 LocalTrainingServer 可以在公网跑后台服务器训练

在 1.18 中，z_ai_model 的后台服务器训练的数据集被压缩到了以前 1/10。

假如在 1.16 版本，z_ai_model 的数据集是 2G，那么 1.18 的数据集就只会有 200M，这种文件尺寸可以在公网中有效请求远程训练

1.18 对数据集的压缩是采用 jpeg 分块压缩的，jpeg 分块会有 quality 参数，这是做为颜色分割使用，它以丢弃使用率最小的一批颜色来降低存储空间，而 ZAI 的内核在将图片 input net 时，都会做金字塔预处理，金字塔会使用高斯核做平滑，那些 jpeg 分块压缩导致丢弃的颜色，都会被平滑掉，简单来说，无论是否使用 jpeg 压缩，input net 的图片信息都是一致的。

我们通过工具链源码，可以自行改进 LocalTrainingServer，使它可以验证身份，甚至加密数据集和分别指派不同的 gpu 工作。

让我们开始搭建自己的机器学习超算后台吧。

By.qq600585

2019-3