Z-AI1.33RC2 更新日志

目录

计算引擎大更新.....2
重大 bug 修复.....2
低等级 bug 修复.....3
    MemoryRaster 光栅引擎.....3
工具链和 Demo.....3
    新增一个 normal training 创建工具.....3
    新增一个针对儿童手写、早教卡片识别和预览的技术方案.....4
    针对儿童早教卡片，新增自动旋转标注.....5

计算引擎大更新

- 彻底抛弃以前按 windows sdk 版本号构建的方式区分计算引擎，不再区分 1806,1903,2004 这类 sdk 版本
- 计算引擎按 cuda 版本进行区分，分别支持了 cuda9.2/10.0/10.1/10.2/11.1，一共 5 个 cuda 版本，可以兼容几乎任何三方项目，例如 tf,mindspore， Z-AI 主动支持各种 cuda 版本
- MKL 不再区分 sequence/parallel 这两种框架，只需要区分 IA32/x64 两种架构，MKL 可以主动检测和支持 2 代志强 avx512
- 裸 IA-32/X64 架构的计算引擎只有一个 dll，不在依赖于三方 opencv/eigen3 这类动态库，可以和应用一起简而美的发布，例如我们只用一下 modelbuilder，不需要计算模型，那么 IA-32/x64 架构的计算引擎就是首选

重大 bug 修复

- **卷积网络**：修复 GDCNIC/GNIC 两个网络不支持 prepare raster 的问题，我们在训练参数中会看到新的 snapshot 这个参数，给 true 它会直接使用框注样本，否则使用整张光栅，此参数看似不起眼，但它能让 GDCNIC/GNIC 从冷宫解放出来，成为我们常用的分类器网络之一，并且 GDCNIC/GNIC 建模非常快，15 秒-5 分钟间即可完成一个模型，GDCNIC/GNIC 可以支持笔记本，游戏卡这种算力薄弱的 gpu。
- **CPU 性能临界**：在多 gpu 的 hpc 中，修复 DNNThread 可能发生的 gpu 太快，输出的数据过多，以至于 cpu 不堪重负，一直 100%拉满处理 gpu 的输出问题。修复以后，gpu 算完会根据 cpu 性能临界做合适的等待。简单来说，cpu 如果不够快，会拖累多 gpu 的运算能力，但不会再出现长时间 100%CPU 处理等待的情况，有效避免算法编程中出现疏忽。
- **GPU IO 超时响应**：在多卡 gpu 系统中，pci 有个延迟机制，当我们发送一个 io 请求后，在额定时间内，gpu 未响应，将返回错误，这与我们的 bios 设置有一定关系，当我们多 gpu 进行大规模计算时，gpu 负载达到 100%，这种延迟将会很大，如果再发送 10G 规模的数量请求，会高几率出现计算错误。在 cuda 内核中，延迟检测从 1 次错误，加大到了 3 次，当延迟错误出现 3 次才会真正触发计算错误，可以对付大多数 hpc 配置不到位的问题。
- **GPU 性能临界**：DNNThread 现在不会再无节制的扔数据给 GPU 计算，如果达到 GPU 性能临界，SendData to GPU 这一步将会在线程进行等待。因为我实际使用语义分割时，4 个 GPU 每秒 40fps 计算像素分割，输入数据无穷等待，达到数百 GB，这是一个地雷，稍不注意内存用完。

低等级 bug 修复

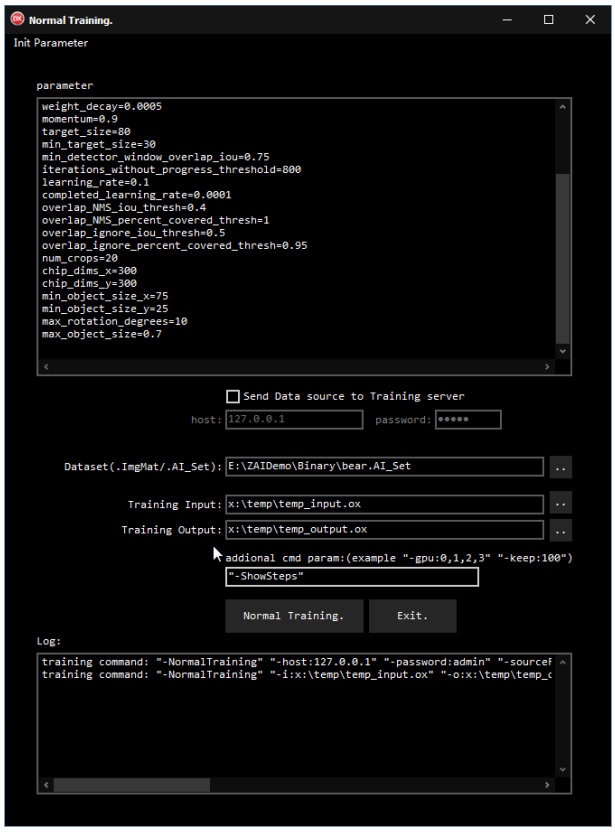
MemoryRaster 光栅引擎

修复在光栅数据使用硬盘转存以后，投影时不会载入硬盘数据的问题（低级错误），这个问题导致了外面跑计算测试，只有 10%的正确率，几乎全部报错，最开始我还以为是计算引擎出的问题，诸如升级 cuda11.1 以后驱动有问题，折腾半天，是我的低级错误!!

工具链和 Demo

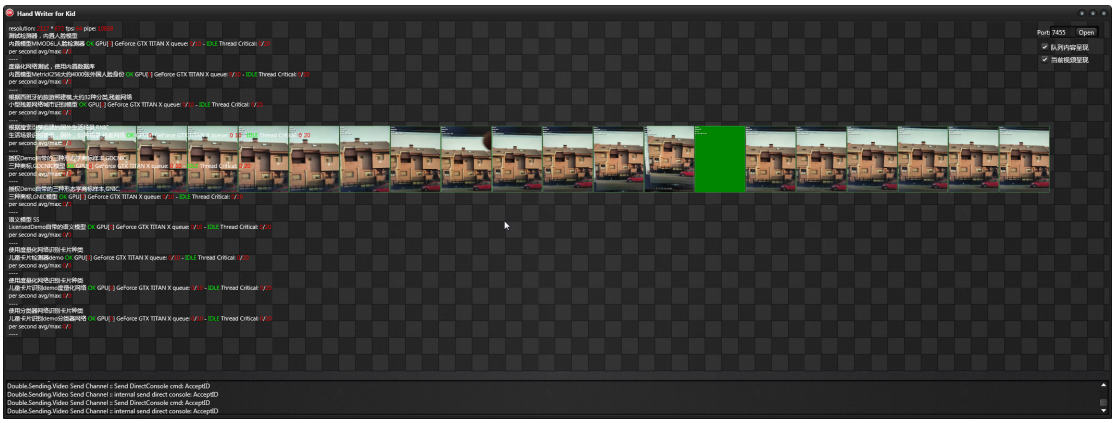
新增一个 normal training 创建工具

我们不想使用训练服务器，也不想重新打开 modelbuilder 这类工具，我们可以直接该工具在本机/网络中训练模型，在 hpc 中测试验证模型它更方便

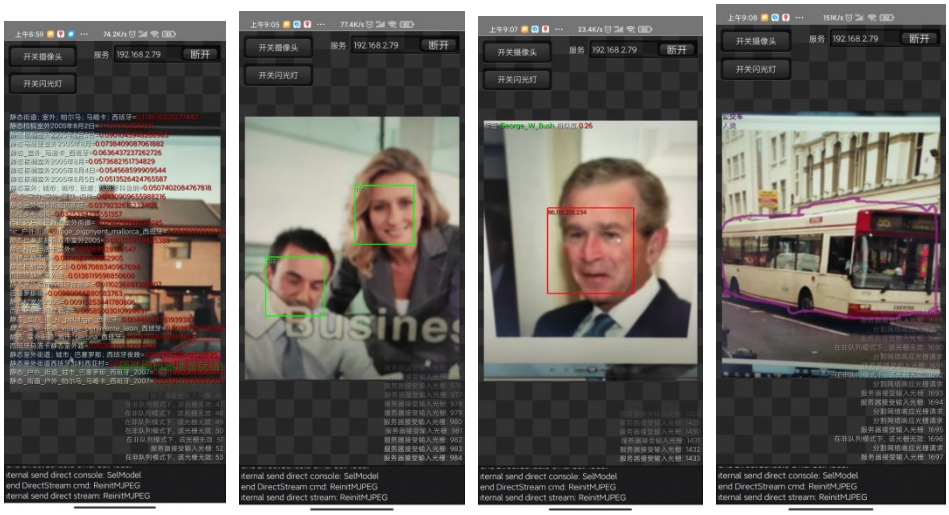


新增一个针对儿童手写、早教卡片识别和预览的技术方案

后台服务器，这套服务器带有内部队列的机制呈现功能，它会把后台数据直接画出来
这套方案是近期做给林涛公司验证模型使用，它几乎能适应任何 dnn 模型，包括：检测器，度量化网络，分类器，分割器。后续版本它极大可能会被纳入主工具链的体系中。

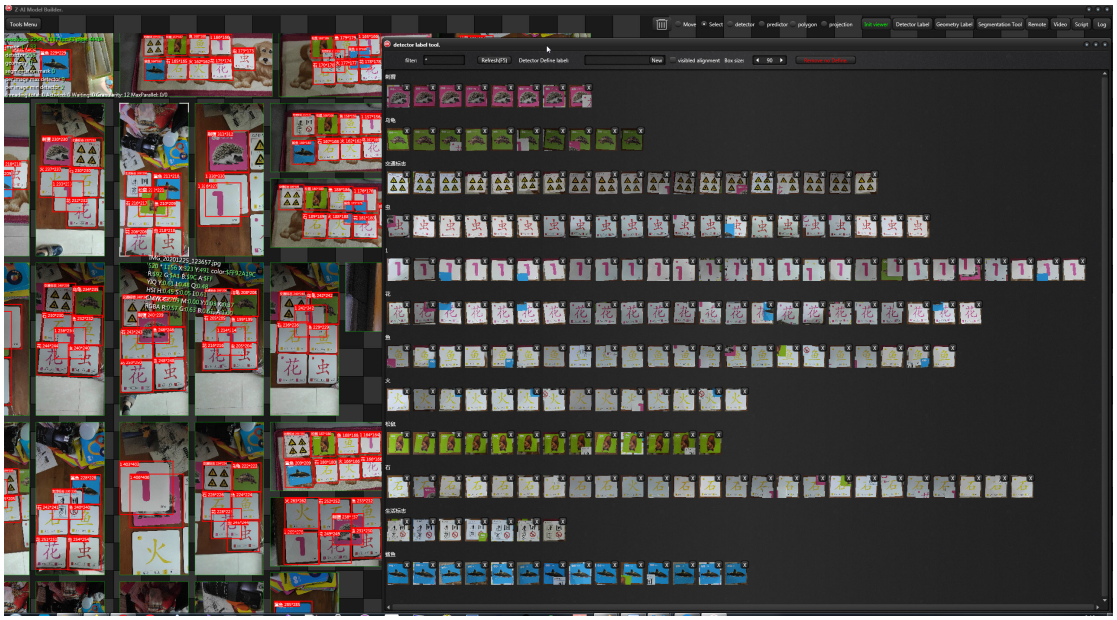


前端预览效果:分类器/检测器/度量化网络/分割器(可放大观看)

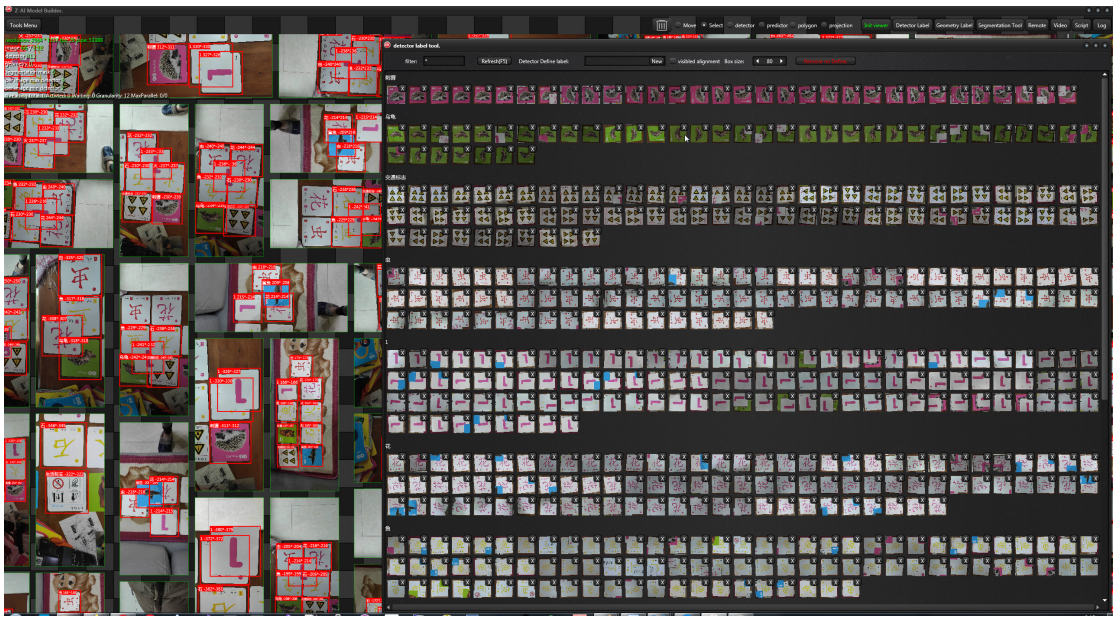


针对儿童早教卡片，新增自动旋转标注

卡片很少，也就 10 来个种类，正面标注如下
正面标注表示卡片没有旋转，是摆正的方式



当使用自动旋转以后，会按左、右、180 方向转动数据集和标注，从而立即增加 3 倍样本数据，对真实场景颠倒卡片的识别起到更聪明的识别效果



自动标注旋转不是简单的向量空间旋转，因为数据样本都是光栅，光栅是数据的，不是向量的，因此处理投影和坐标系，是慢慢推敲出来的。旋转效果很好，坐标系也算的很准。

完
by.qq600585