

# zAI 图像语义分割建模指南

文档版本 2.0

更新时间 2020-2-18

## 目录

概述 .....	2
准备工作.....	2
基于几何标注的图像语义分割建模.....	3
自动化几何标注.....	8
基于像素蒙版的图像语义建模.....	9
基于 VOC2012 开放数据集制作样本.....	9
基于 Photoshop 制作样本 .....	17
几何标注与像素蒙版的相互搭配.....	19
真正的大规模样本制作.....	19

# 概述

图像语义分割在大数据提炼(大规模标注, 大规模采集), 视觉(图像和视频)合成(换脸, 换背景), 自动驾驶, 等等领域, 应用非常之多。

zAI 的图像语义分割使用的 DNN 框架是 cuda, 第一次发行是在 1.20 的版本, 版本详情请参看 1.20 更新日志。

zAI 的图像语义分割使用了两种数据结构, 分别是带有包围+塌陷的几何描述数据, 以及使用颜色蒙版作为区分的位图数据, 这两种数据结构在 zAI 中均有完整的技术支持体系, 整个体系无任何水分, 稍作提炼即可应用(zAI 针对语义分割同步提供了用于提炼的技术体系)。

图像语义分割从 DNN+视觉体系的设计来说, 相比对象检测和图像分类器, 更接近真实人类视觉, 它有替代对象检测和图像分类的性质, 图像语义分割的数据提炼和建模工作也相对更加专业和繁琐, 对 GPU 的性能要求也更上层楼。

# 准备工作

图像语义分割系统只存在于 zAI 1.20 或则后续的版本中, 在早期版本是不存在的

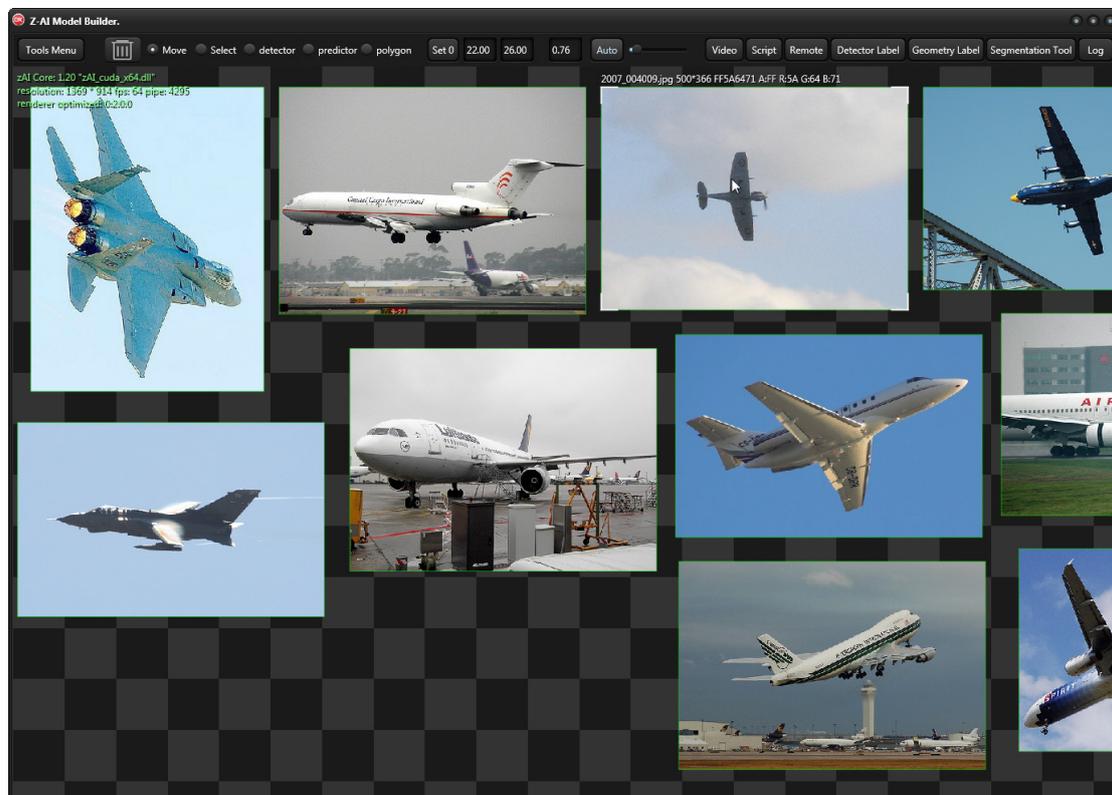
zAI 1.20 最低要求使用 nvidia cuda 10.1 update 1 并且需要与它对应的 cudnn。Cuda 需求文档请参考 [第一次接触 ZAI 需要的准备工作.pdf](#)

对于硬件的要求, 图像语义分割即使 100 张图片的分割, 也会是小型超算规模的需求。诸如要求: 使用 Titian 级的深度学习卡, 并且应该做好搭建 NVLINK 桥的准备(上 NVLINK 桥的 GPU 设备目前性价比最优屌丝卡是 Titian RTX, NV 官方指定 JD 购买, 16999 一张, 确定是只能桥 2 张卡, 如果 3 张卡, 需要考虑 NV 的 P,V 系列)。

使用小型超算系统, 纯粹运行建模程序, 不包括人为干扰, 目前深度学习行业的图像语义分割平均建模运算时间为 2-4 周, 语义分割需要的样本规模非常大。

# 基于几何标注的图像语义分割建模

首先，我们往建模工具导入一批需要用于分割的图片样本，这里我是以飞机样本为例

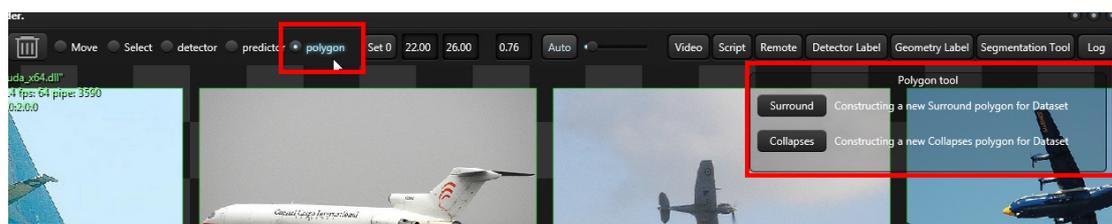


然后切换至 polygon 多边形建模工具

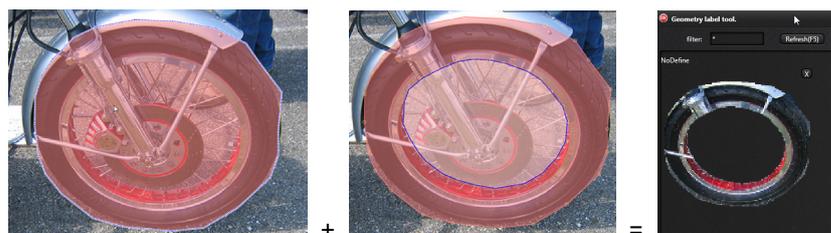
在红框中，有一个 Polygon tool 的透明小窗口

Surround，多边形包围描述，这是描述图像外形轮廓的多边形

Collapses，多边形塌陷描述区，这是描述在包围区内被掏空的外形



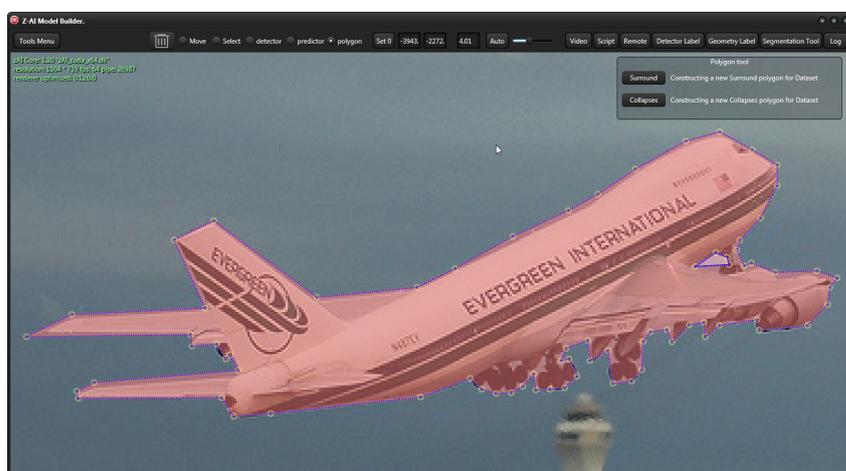
附图解释：包围是轮胎的外部轮廓，轮胎的内部轮廓就是塌陷，这是被掏空的部分，zAI 的光栅系统会自动根据这两种几何特性，产生可以输入到语义分割网络去学习的图像。



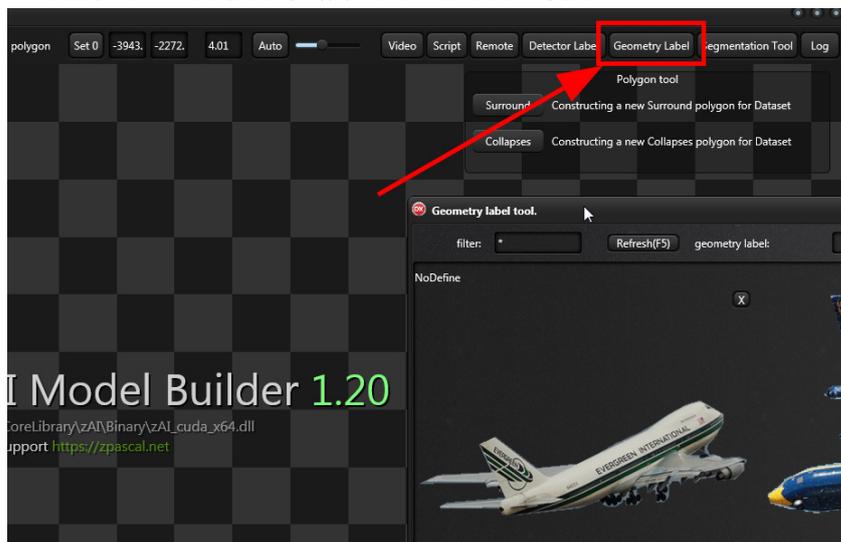
由于飞机几乎没有塌陷的内容，都是轮廓，我们使用几何包围方式即可描述完成飞机外形



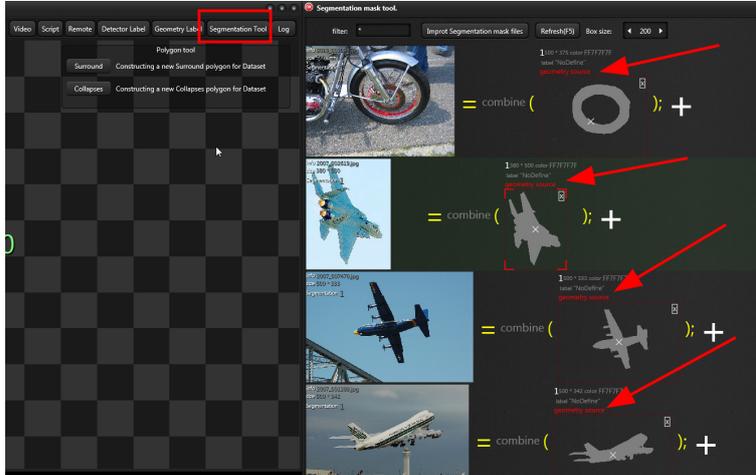
但是也不排除个别的飞机图像会有塌陷的部分，诸如下图



我们通过 Geometry Label 工具可以立即查看到通过几何方式描述出的飞机外形图像  
Geometry Label 不应该随时打开，因为内部程序会频繁的投影和刷新，会让建模速度变慢，正常情况下，当我们需要做几何标注时再去使用它。



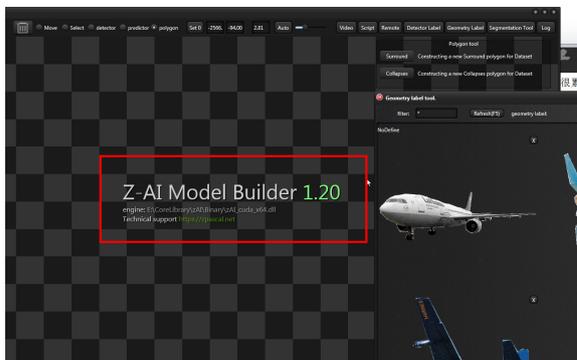
同样的我们 Segmentation tool 中也可以看到刚才我们通过包围+描述完成的飞机外观，这是由几何数据转换而出的是像素分割数据，相比 Geometry Label 的图像，这里的图形更抽象在 Segmentation tool 中，几何描述数据是以投影方式生成的像素蒙版，红色箭头所指的地方是指蒙版的数据来源，这些含有红色的 geometry source 的蒙版数据无法通过 X 按钮删除，即使删除，然后刷新，它还会再次出现，只能回到 Geometry Label 或则主视口才能真正的删除。



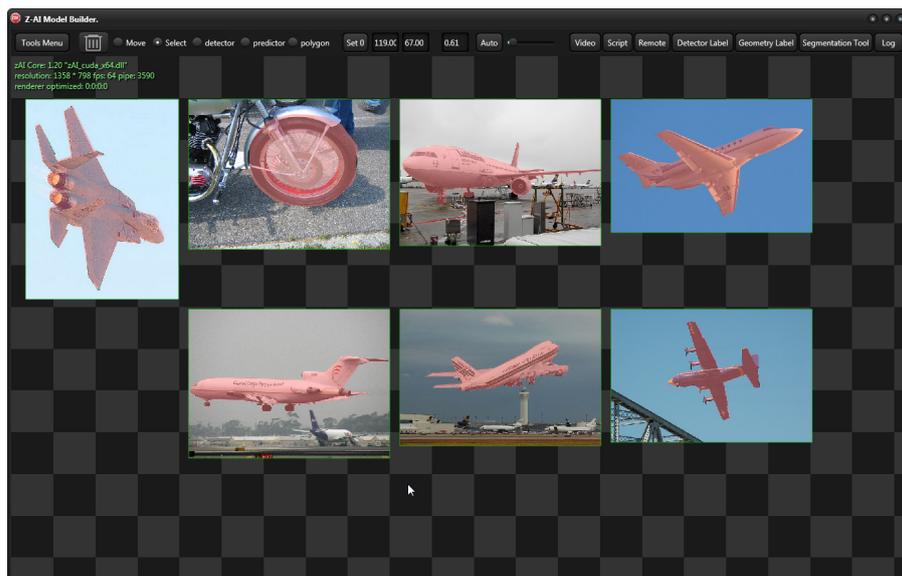
在我们以几何方式勾画图像的过程中，不必十分精细，高频率点鼠标会让手很累，我们只需要勾出大致轮廓，然后通过点，线，方式去修改即可。当鼠标移动到点和线附近，可以自动拾取就近的点和线，直接修改即可完成几何轮廓描述。



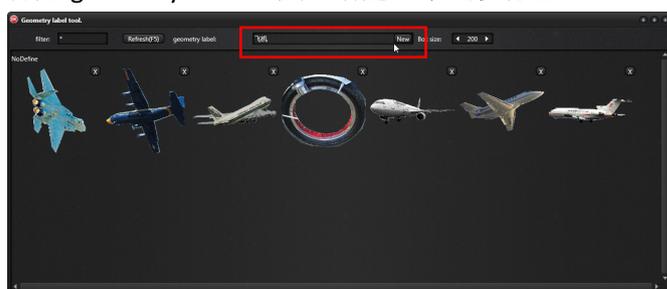
按上面的说明，我们可以再次使用 Geometry Label 工具来查看我们刚才描绘的这架飞机 zAI 在 1.20 版本新增的自动化显存回收机制，只要焦点离开主视口，在主视口就会看不到任何图像，只会看见红框中这几个字，待我们把焦点切回主视口后，才能看到图像。显存回收机制在应付数千张图片时，会自动把不可见图片占用的显存给清理掉，当我们离开主视口后，系统会自动进入显存回收的工作模式。这正是显示红框文字的原因。如果你觉得这样干有点影响自己的建模工作，可以以授权方式，找作者索要建模工具链的源码去改回来。



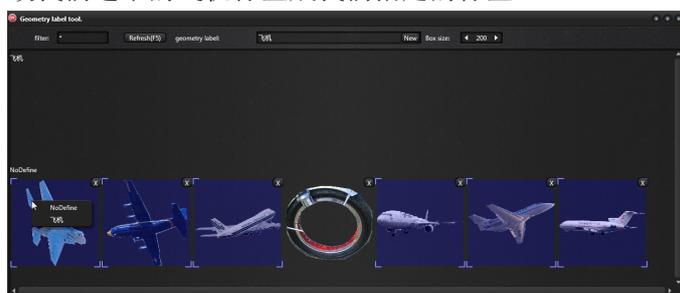
依照上面的修改编辑技巧，10分钟以后，我们完成了总共6架飞机的几何标注



打开 geometry label 工具，新建一个飞机标签



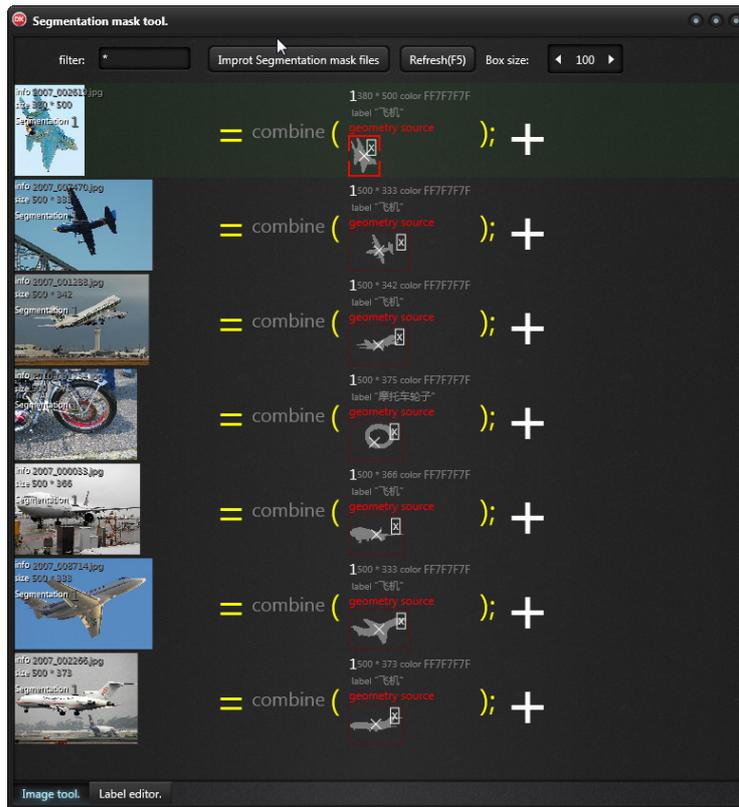
将 NoDefine 中的飞机拖放至飞机标签上即可，同样，我们也可以通过鼠标右键的菜单来修改我们选中的飞机标签成我们指定的标签



标注完成后



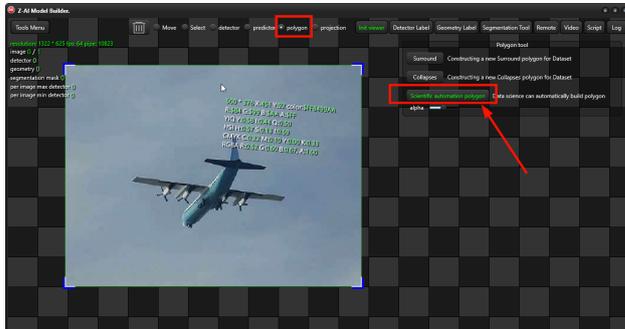
同样我们也可以通过分割蒙版的来查看我们的工作成果



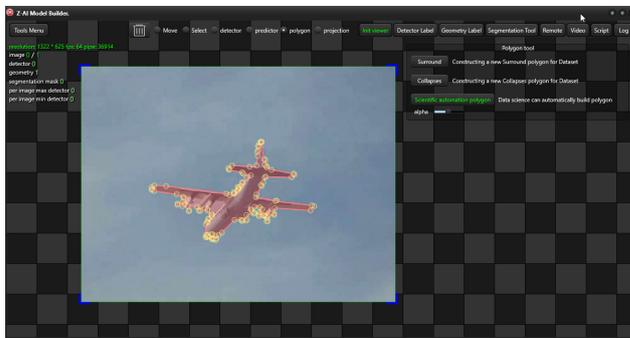
剩下的就是训练模型了

# 自动化几何标注

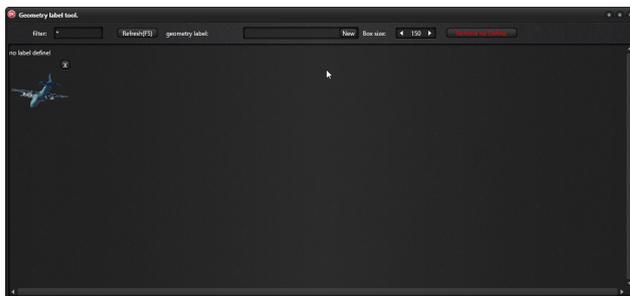
切换至 polygon 编辑器,使用 Scientific automation polygon



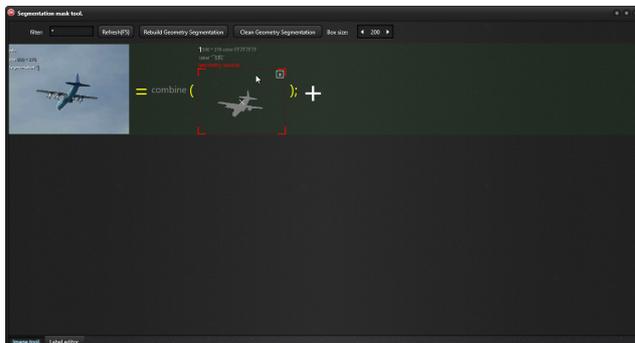
拉框,框住飞机,自动标注完成,自动标注对于特写镜头且目标显著性突出可以得到比较好自动标注.接下来



在几何标签编辑器给出标签



在分割蒙版工具中,点击 rebuild geometry segmentation,飞机的分割样本制作完成



# 基于像素蒙版的图像语义建模

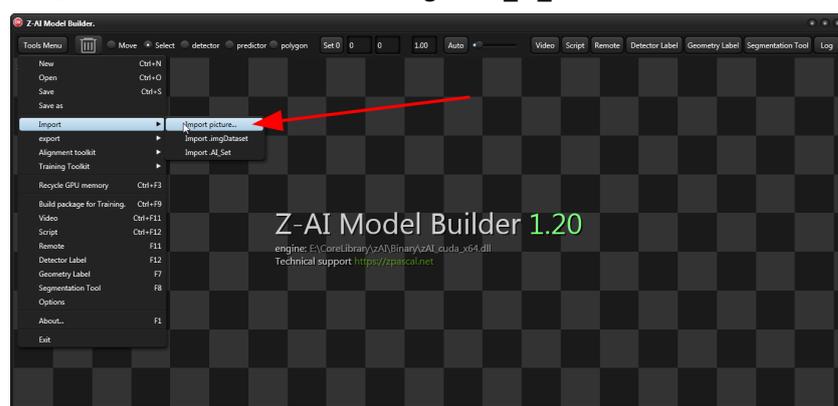
像素蒙版和几何标注的差异在于：像素蒙版用纯色表示平面物体，几何标注用几何(包围+塌陷)表示平面物体。

在机器视觉领域，图像语义分割都使用像素蒙版来输入数据，既：输入的数据都是一种纯色图像。因此，像素蒙版在 ZAI 中被定义成为一种接口其它数据样本的中间格式，比如，我们使用各种三方已经训练好的模型作为语义内容来做提炼，最后得到我们需要的“纯色标注蒙版”，一旦我们得到了“纯色标注蒙版”，zAI 的导入和标注就会变得非常简单。

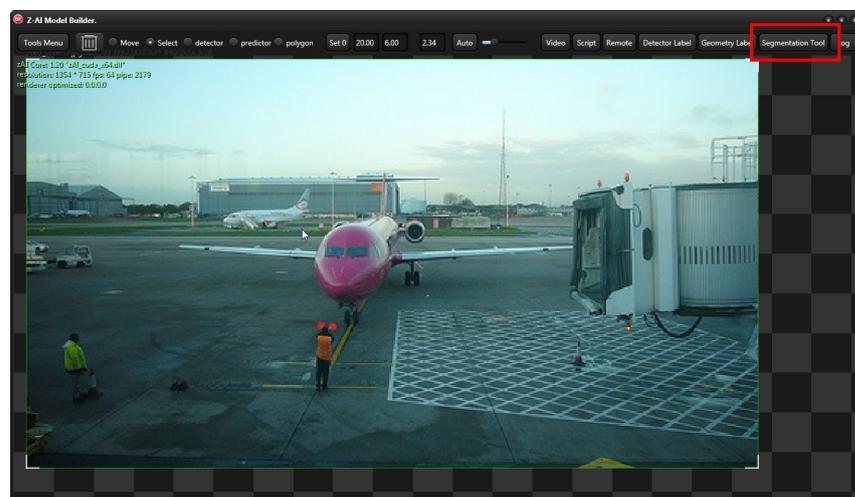
当然，我们也可以使用 Photoshop 这类工具来制作纯色标注，但是，它的工作效率远不如几何标注来的有效(几何标注在 1-2 分钟内即可标注完成一架飞机，使用 Photoshop，即使专业美工至少也要 10 分钟以上)，但是使用 Photoshop 的标注方式，仍然不失是一种全手工数据样本制作方法。

## 基于 VOC2012 开放数据集制作样本

首先我们导入一批 voc2012 的 image 到 z\_ai\_model，并且，确保从主视口能观察到结果

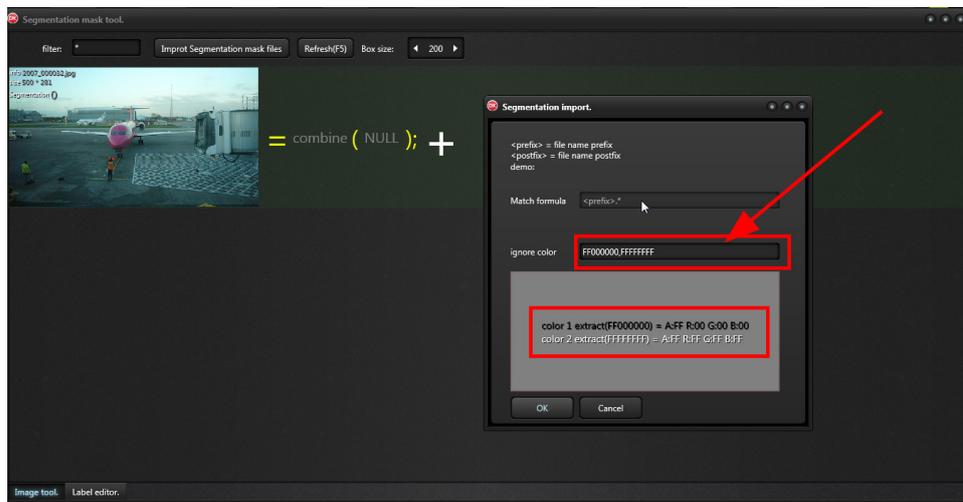


Voc2012 数据集有数十万张图片，全部导入会导致内存+显存爆炸，随意抽样导入一批即可。本文只导入了一张图片。然后我们打开红框中的分割工具。





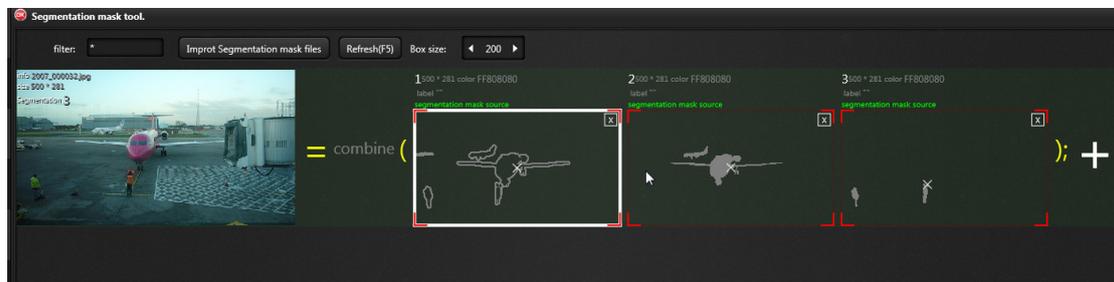
接下来，我们会看到一个弹窗，Match formula(文件自动匹配规则)，这在我们使用+号导入文件时是不可编辑的，它是用于自动化匹配大批量纯色分割蒙版文件名用的。我们只需要注意箭头指向的红框内容，ignore color，它的含义是无用的分割色



现在，我们可以什么都不用做，直接 OK 开始导入  
现在，我们可以看到 3 个分割颜色的蒙版

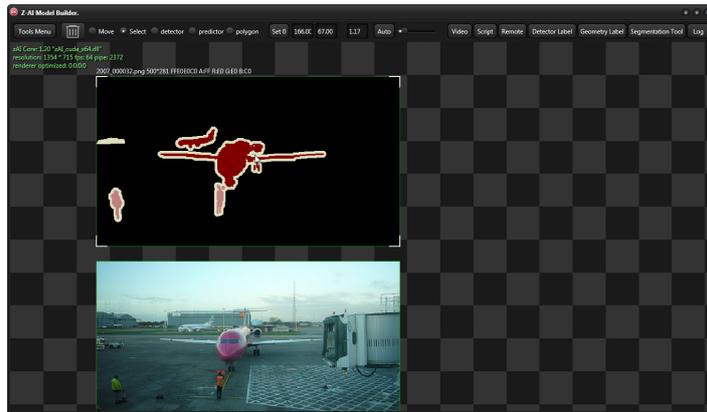
- 1, 是飞机和人类的轮廓
- 2, 是飞机的分割
- 3, 是人类的分割

对 zAI 来说，轮廓是无用的分割色，因为 zAI 在做 input 处理前，会自动生成对象轮廓，避免 DNN 跑去无节制学习对象周围的颜色。

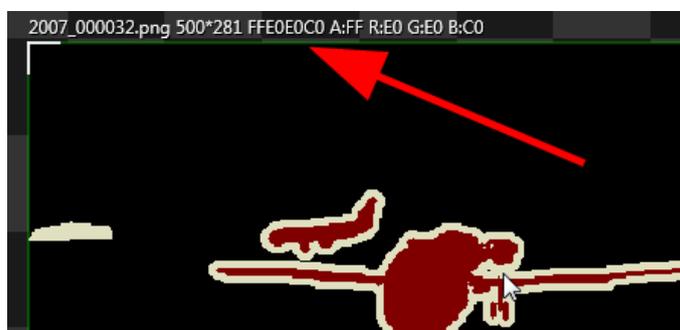


怎样确定无用的分割色？

最简单的办法是把分割的图片导入到主视口中，然后将鼠标移动到轮廓色带上

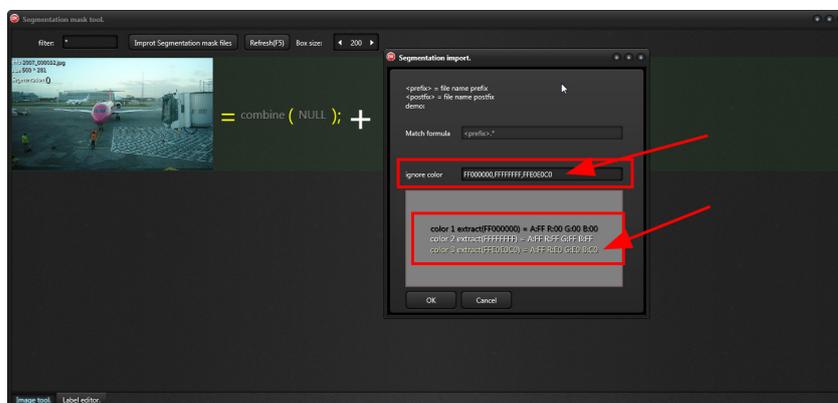


在图片上方，在分辨率后面，会有一个像素信息，其中，FFE0E0C0 就是包围带的颜色

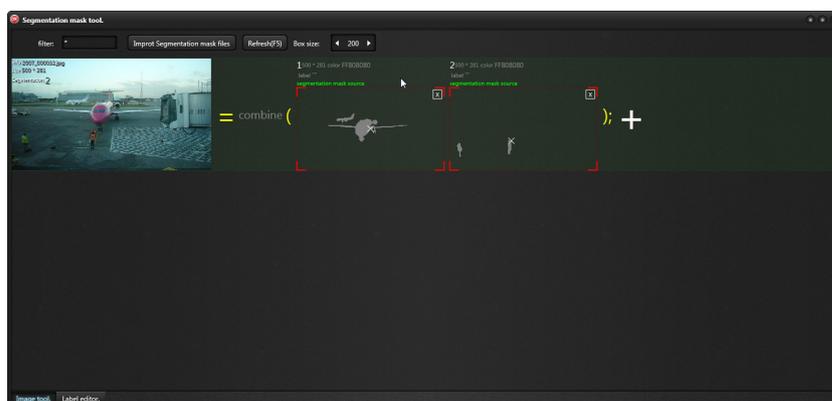


当确定了 FFE0E0C0 以后，我们可以直接用 del 键删除图片，然后再次导入

在 Ignore color 后面输入 FFE0E0C0，然后，可以在下面看到黄色的 FFE0E0C0 解码效果,点 OK



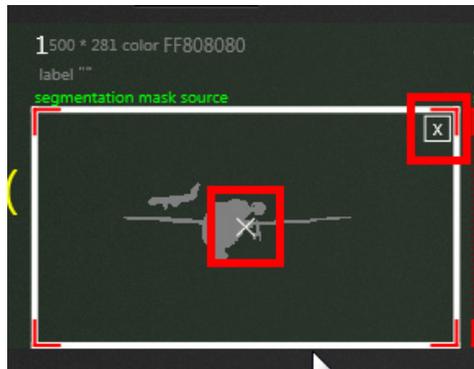
这时我们就成功导入了飞机和人类的分割蒙版



两个 x 的说明，在红框中的两个 x

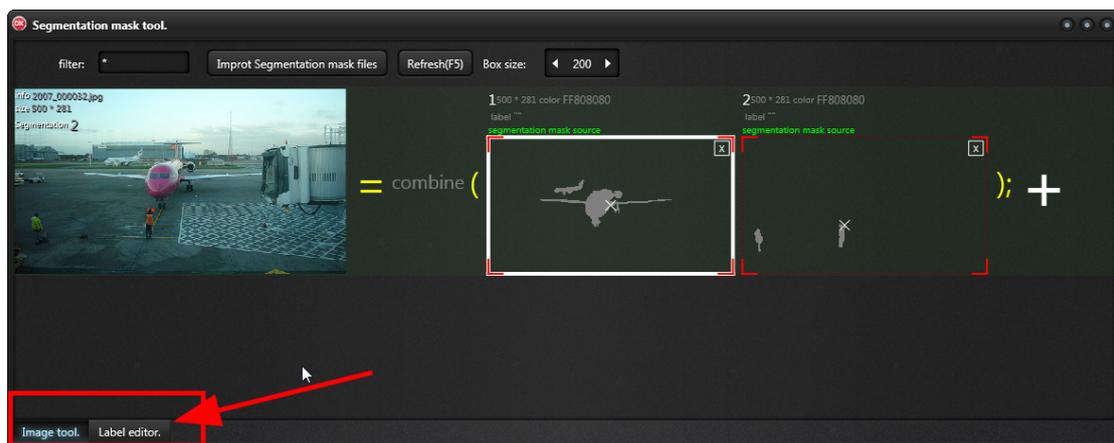
飞机上的 x 表示颜色定位坐标，在使用 Photoshop 纯人工制作颜色分割蒙版才会用到，如果导入的是已经制作好颜色分割蒙版，这个坐标是自动化计算的，不用管它。

右上角白框内的 x 表示删除，如果我们觉得这张颜色分割蒙版不对，点下就删除了



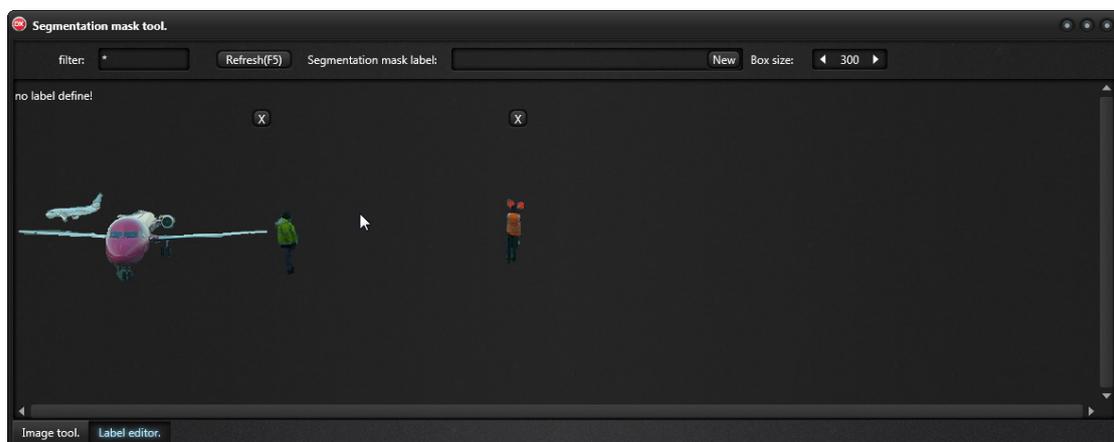
现在，让我们来做分割标签

在 segmentation mask tool 中，Image tool 表示按颜色分割图像工具，旁边的 label editor 则是标签工具了



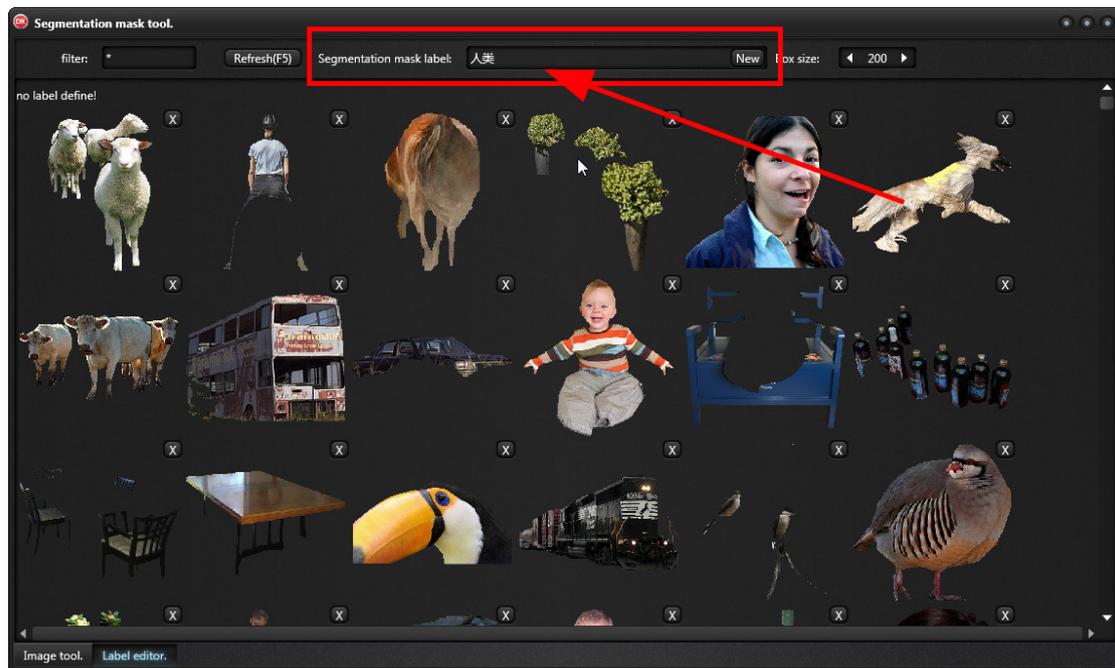
我们切换至 Label editor.

我们会看到和 Geometry Label 工具类似的图像，我们现在开始定义两个标签：飞机+人类

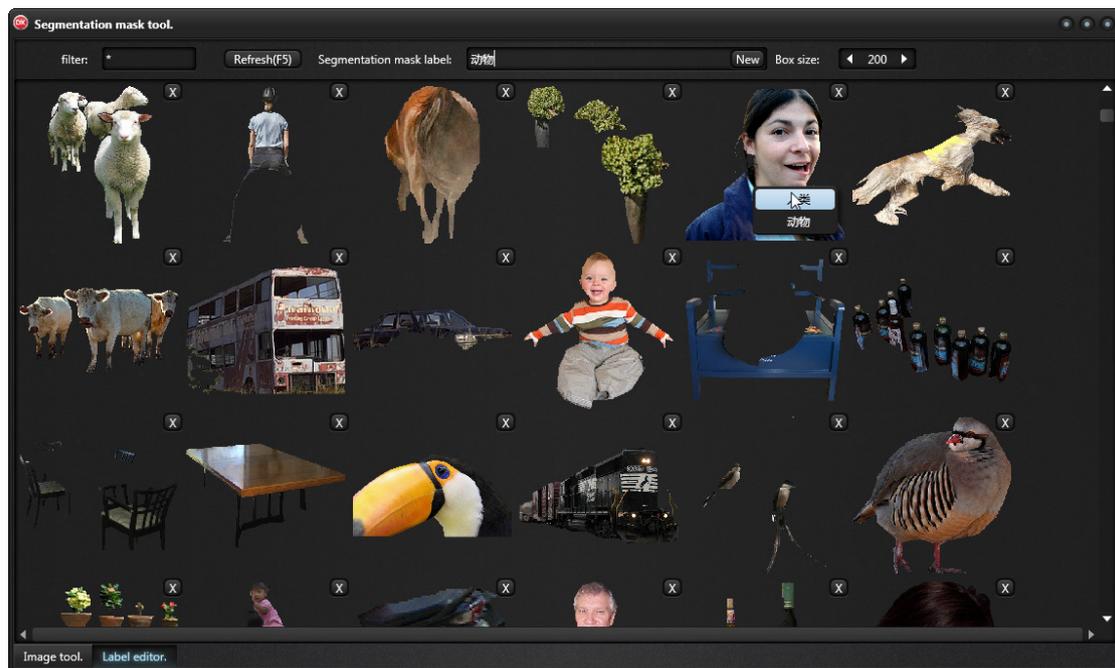




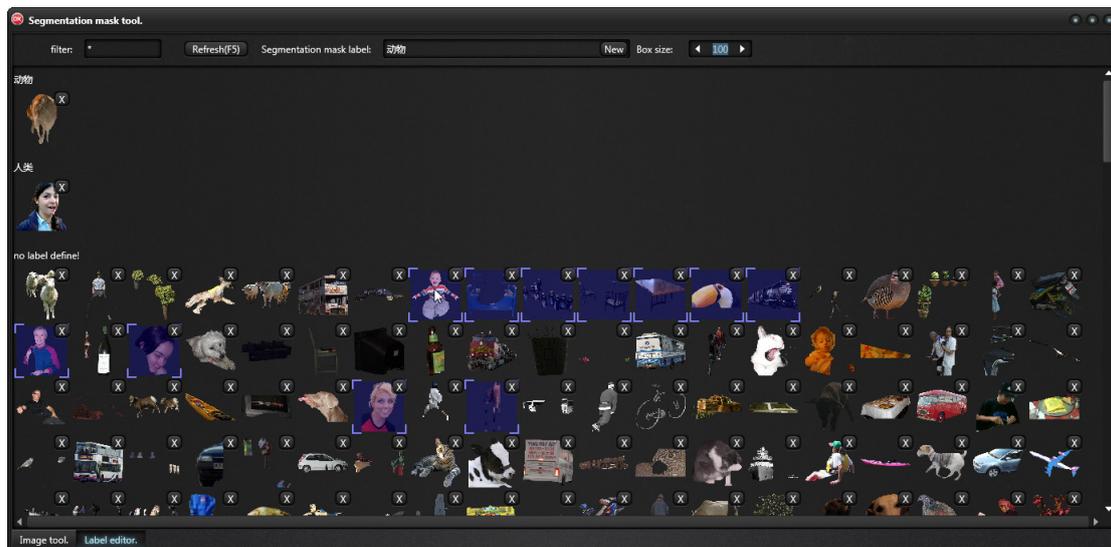
Label editor 对付大规模数据集的分类方法  
先建好我们的分类，人类+动物



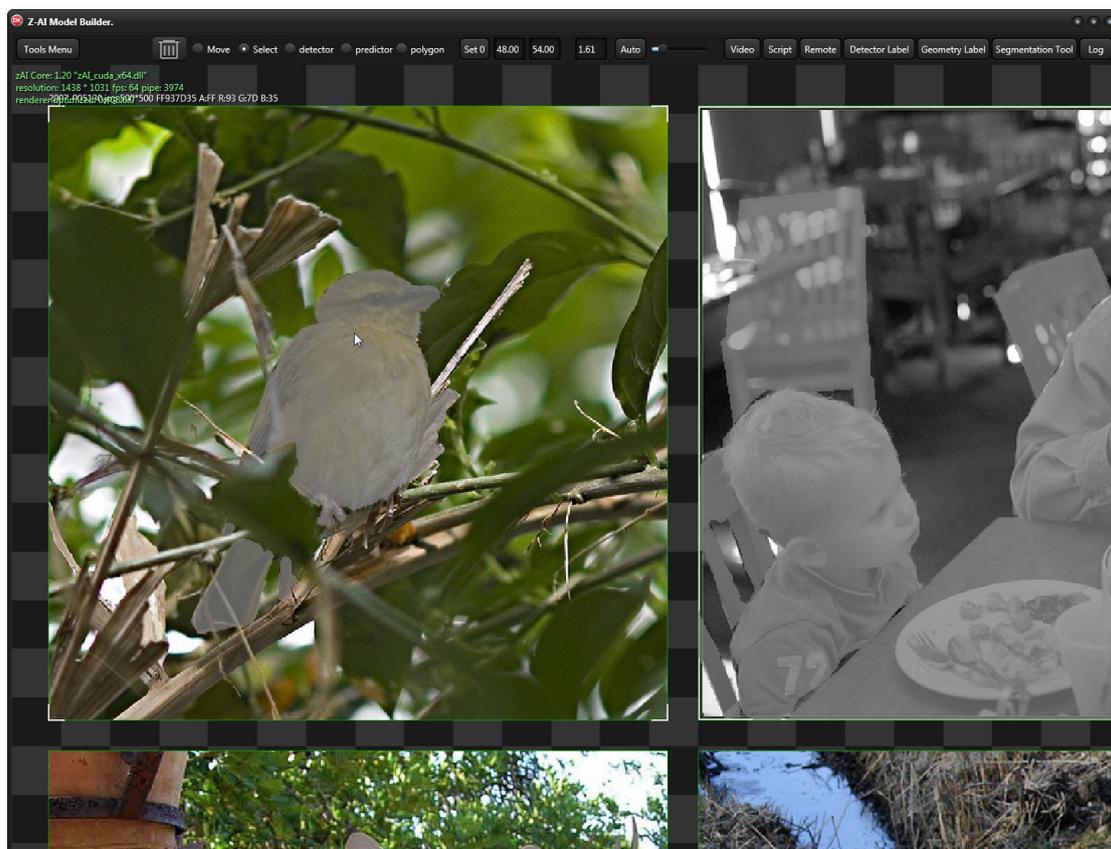
然后通过右键弹出菜单，直接指定



另一个办法是配置 2 台 5K 的显示器，将两个屏幕相连，通过拖拽来进行分类，这是最高效的 label editor 使用方式

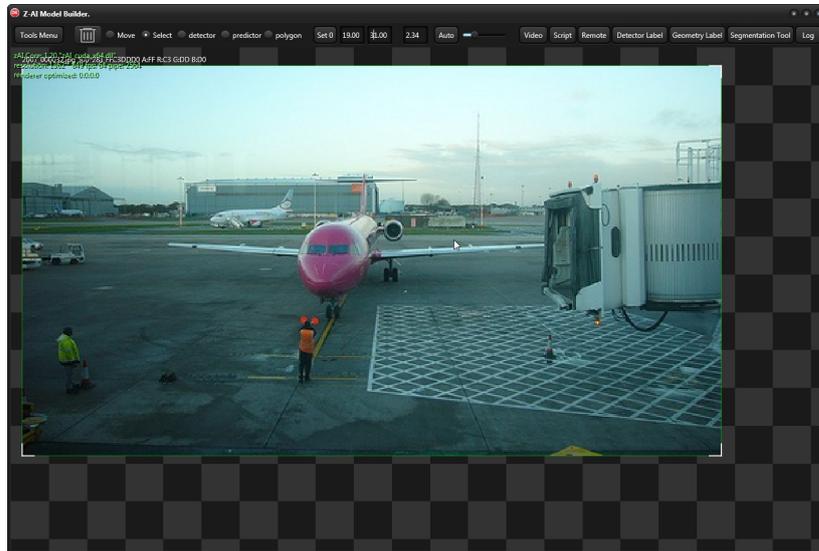


说明：当 `ai_set` 的样本中存在颜色分割数据，我们在主视口看到的图像会有明显的灰色感，这和我们在 `segmentation tool` 中的看到的灰色是一致的，在主视口中，颜色分割是以蒙版的方式画上去的。总之养成一种视觉感受习惯，只要是不正常变灰的就是颜色分割蒙版。

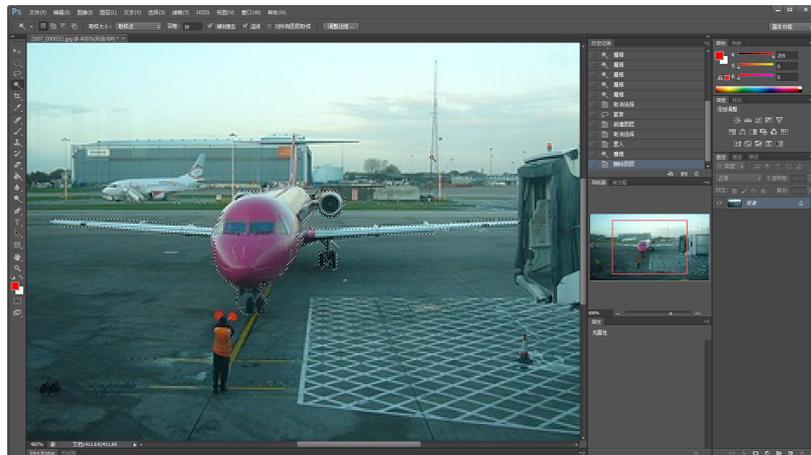


# 基于 Photoshop 制作样本

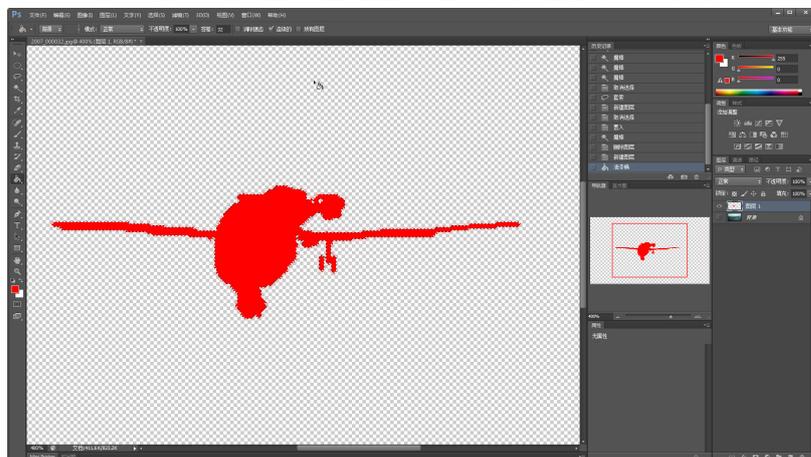
先把图片导入到主视口



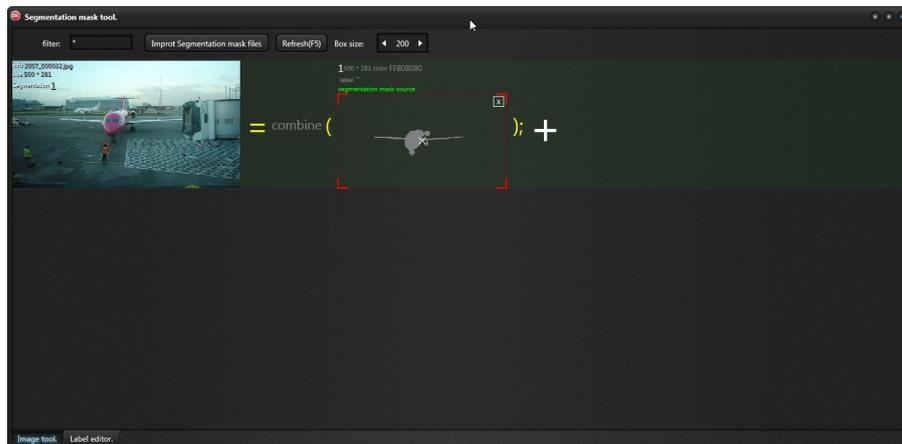
然后使用 photoshop 打开同名文件，使用魔棒，套索这类工具制作好蚂蚁线，这一过程等同于高精度抠图，需要细心，耐心，技术，需要专业美工才能制作。如果缺乏这类技术加点，选择使用几何标注来建模即可，这是 zAI 针对小白程序量身设计的制作工艺。



在新图层中按蚂蚁线颜色填充纯色，这一步注意要消除锯齿



Ps 编辑完成后，导出成.bmp,.png 这类无走样像素格式的文件，再导入进来这就完成这个纯色分割样本了



# 几何标注与像素蒙版的相互搭配

几何标注倾向于所见即所得，随手画标注，轻松的制作分割色彩标注。几何标注是先绘制几何体，然后 zAI 再将几何转换成像素蒙版，转换过程是自动化瞬间完成的。

像素蒙版更倾向于提炼从现有的数据集，以及自动标注。

zAI 图像语义分割输入的数据都是像素蒙版。即使使用几何标注，最后都是转换成像素蒙版后再 input DNN。

## 真正的大规模样本制作

真正的大规模样本会到达 nTB 的数据体量，单单使用 z\_ai\_model 是不行的，必须将数据系统以 imgMatrix 方式来保存。而制作 imgMatrix，需要依靠编程，imgMatrix 的支持都在 zAI\_Common 库中，z\_ai\_model 制作的是一个.ai\_set，导出来是.imgDataset，而 imgMatrix 则是一个无限制的.imgDataset 的大合集，它可以使用 LargeScale 技术来驱动零散的.imgDataset 建模。

一句话总结大规模建模：依靠编程生成.imgMat 数据库，别无他法，这一步没有自动化工具。具体实现技术和 demo，可以通过邮件或则 qq 留言找作者询问。

2020-2-18

By.qq600585