



## 移动机器人系列课程

### 机器人视觉技术与深度学习课程

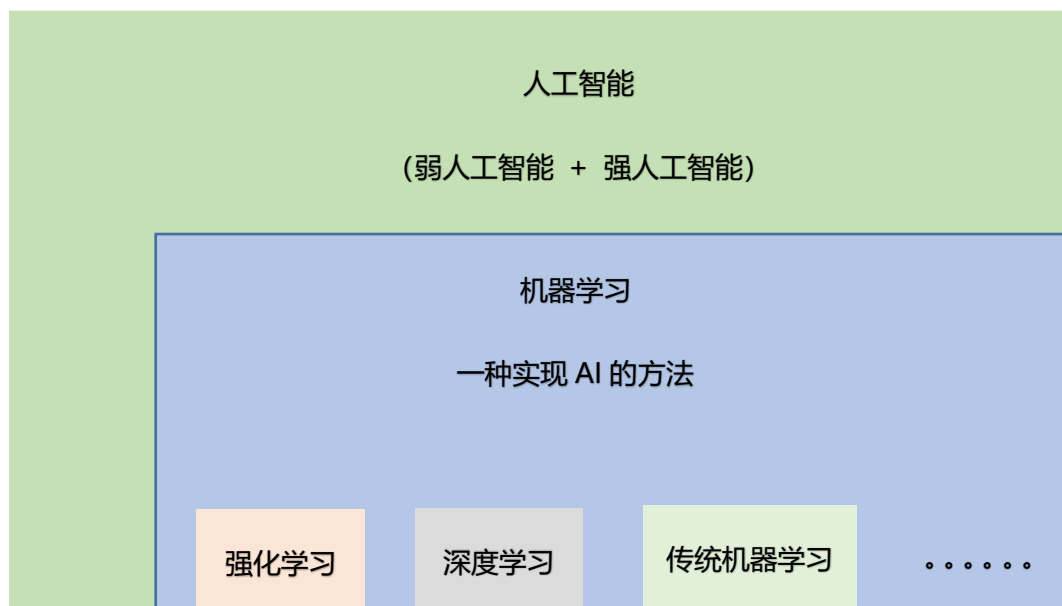


深圳市史河机器人科技有限公司

## 深度学习与数据标注

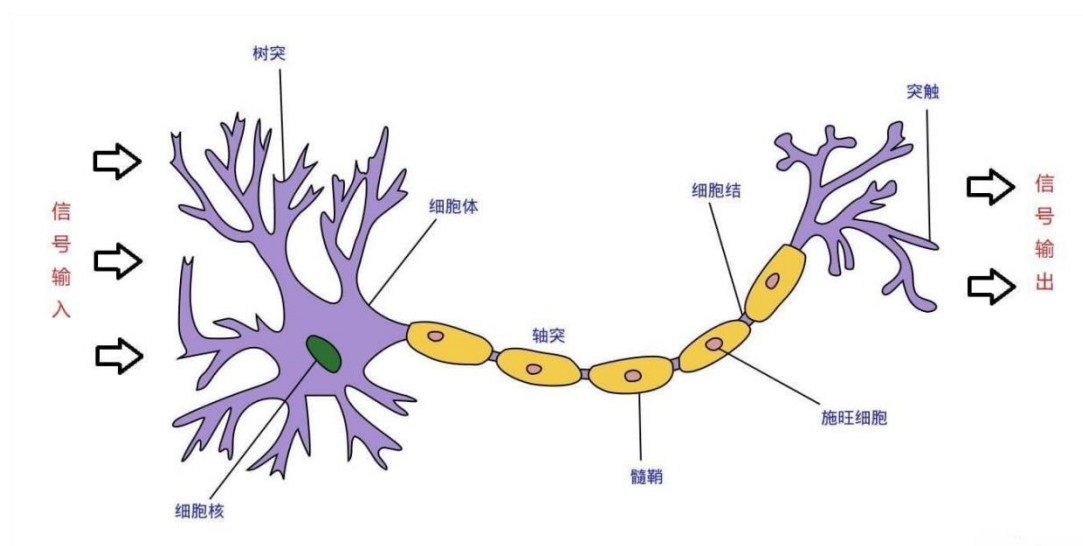
### 1. 深度学习简介

在介绍深度学习之前，先和大家介绍一下人工智能、机器学习和深度学习这三者之间的关系。人工智能就是我们常说的 AI，在 1956 年被科学家们提出这一概念，科学家探讨能不能制造出一个可以像人类大脑的一样思考的机器，拥有人类的智慧，这就是人工智能。实现人工智能的一种方法我们称为机器学习(Machine Learning)，就是让机器从历史数据中学习规律，然后将规律应用到未来中。其实人类的行为也是通过学习和模仿得来的，所以我们就希望计算机和人类的学习行为一样，从历史数据和行为中学习和模仿，从而实现 AI。但是随着研究不断深入，传统机器学习算法在很多“智能”问题上效果不佳，如图片识别、语音识别和语义理解等。后来人们对传统的神经网络算法进行了优化提出了深度学习(Deep Learning)的概念，很好地解决了很多传统机器学习算法效果不佳的问题。综上所述，机器学习是一种实现 AI 的方法，而深度学习是机器学习中的一种技术，他们之间的关系图如下。



## 1.1 深度学习与神经网络

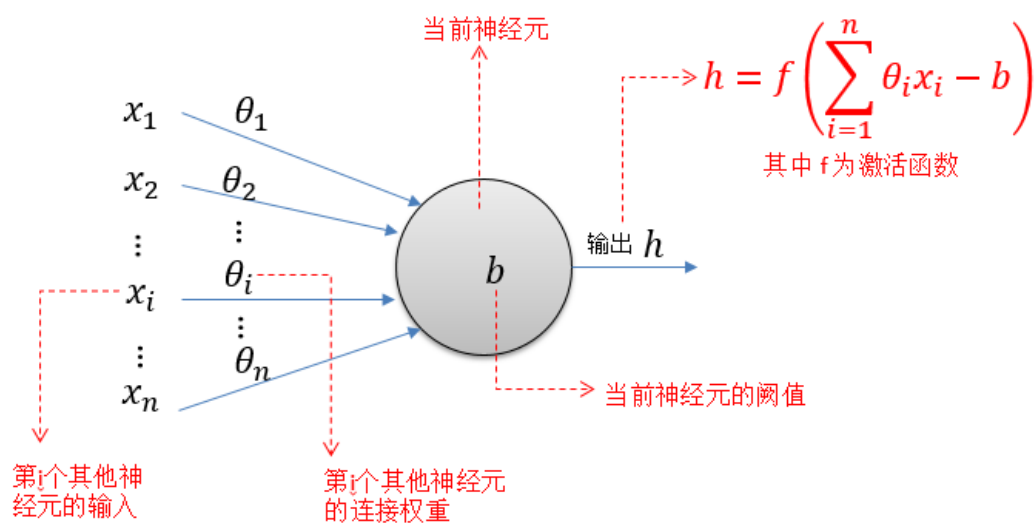
深度学习是基于神经网络算法的，人们从生物神经网络的运作机制得到启发，构建了人工神经网络。



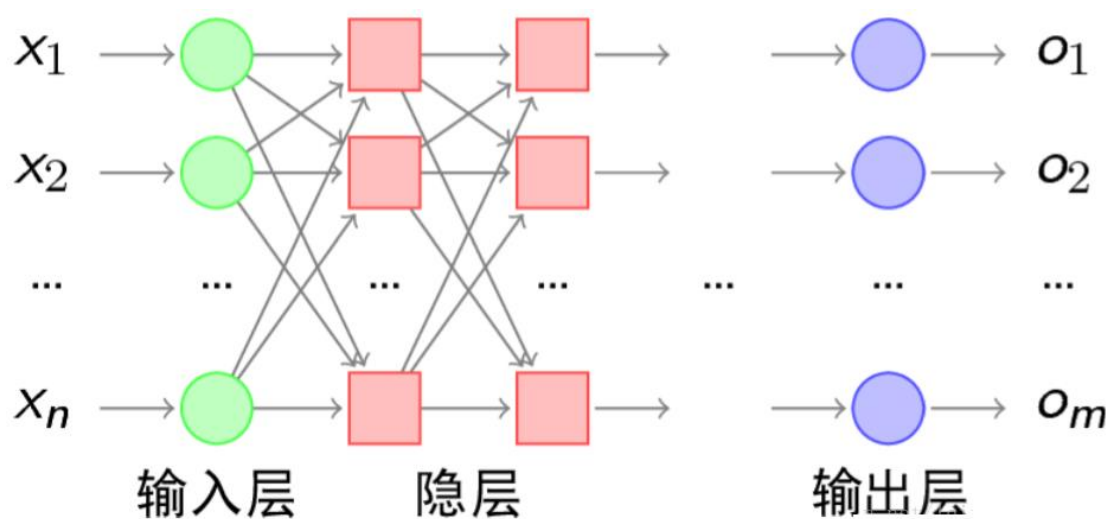
如上图是生物神经元的示意图，神经元通过接收外界信号，达到一定阈值，触发动作电位，通过突触释放神经递质，可以是兴奋或抑制，影响突触后神经元。

下图是经典的 MP 神经元模型，与生物神经元类似，输入端接受外界信号， $x$  和  $\theta$  等

都是参数值，经过激活函数  $f$  (Activation Function) 来判断是否触发“动作 电位”，对外释放的新的信号。



多个神经元组合在一起就形成了神经网络，下图便是人工神经网络的示意图。



可以在图中看到人工神经网络除了输入层和输出层，中间的统称为隐含层，隐含层可以有好多层，深度学习中“深度”的含义就是在输入层和输出层之间加更多的隐含层。

深度学习的概念源于人工神经网络的研究，但是并不完全等于传统神经网络。而很多深度学习算法中都会包含“神经网络”这个词，所以可以说深度学习是在传统神经网络基础上的升级，约等于神经网络。

## 1.2 深度学习与机器学习

### 1. 深度学习与机器学习的相似点。

他们都会对数据进行以下操作：

- 数据清洗
- 数据标签
- 归一化
- 去噪
- 降维

### 2. 深度学习与机器学习的区别

传统机器学习的特征提取主要依赖人工，针对特定简单任务的时候人工提取特征会简单有效，但是并不能通用。

深度学习的特征提取并不依靠人工，而是机器自动提取的。深度学习的可解释性很差，因为有时候深度学习虽然能有好的表现，但是我们并不知道他的原理是什么。

### 3. 深度学习优点

- 学习能力强
- 覆盖范围广、适应好
- 数据驱动、上限高
- 可移植性好

### 4. 深度学习缺点

- 计算量大、便携性差
- 硬件要求高

- 模型设计复杂
- 没有“人性”，容易存在偏见

## 1.3 常用的深度学习算法框架

因为深度学习发展至今，很多算法都已经是通用的，而且得到过验证的，有些公司就希望将一些标准算法一次性开发好，封装起来，后面再使用时直接调用引入即可，不需要大家从头再写一遍。常用的框架如：Tensorflow 是 Google 旗下的开源软件库，里面含有深度学习的各类标准算法 API 和数据集等；Pytorch 是 Facebook 旗下的开源机器学习库，也包含了大量的深度学习标准算法 API 和数据集等，Caffe 是贾扬清博士开发的深度学习框架，国内百度目前也有自己的深度学习框架 Paddle-Paddle，还有我们 BR280 中使用的 YOLO，作者 Joseph Redmon 自己写的深度学习框架 darknet。在上诉的各种框架内，我们可以实现各种具体的目标检测算法，如 R-CNN、Fast RCNN、YOLO、SSD 等等。

## 1.4 数据集

不管是深度学习还是传统的机器学习，都很依赖数据。数据集是一系列数据组成的集合。数据集是大部分人工智能算法得以有效运行的关键环节。数据集的数据标注越准确、标注的数据量越大，算法的性能就越好。

数据标注相当于人类从书本中获取已有知识的认知行为。具体操作时，数据标注把需要计算机识别和分辨的图片事先打上标签，让计算机不断地识别这些图片的特征，最终实现计算机能够自主识别。数据标注为人工智能企业提供了大量带标签的数据，供机器训练和学习，保证了算法模型的有效性。

目前，学术界尚未对数据标注的概念形成一个统一的认识，比较认可的是由王翀和李飞飞等人提出的定义。他们认为，标注是对未处理的初级数据，包括语音、图片、文本、视频等进行加工处理，并转换为机器可识别信息的过程。原始数据一般通过数据采集获得，随后的数据标注相当于对数据进行加工，然后输送到人工智能算法和模型里完成调用。

目前数据标注有 3 种常用的划分方式：

- (1)按照标注对象进行分类，包括图像标注、视频标注、语音标注和文本标注；
- (2)根据标注的构成形式，将其分为结构化标注、非结构化标注和半结构化标注；
- (3)根据标注者类型，分为人工标注和机器标注。

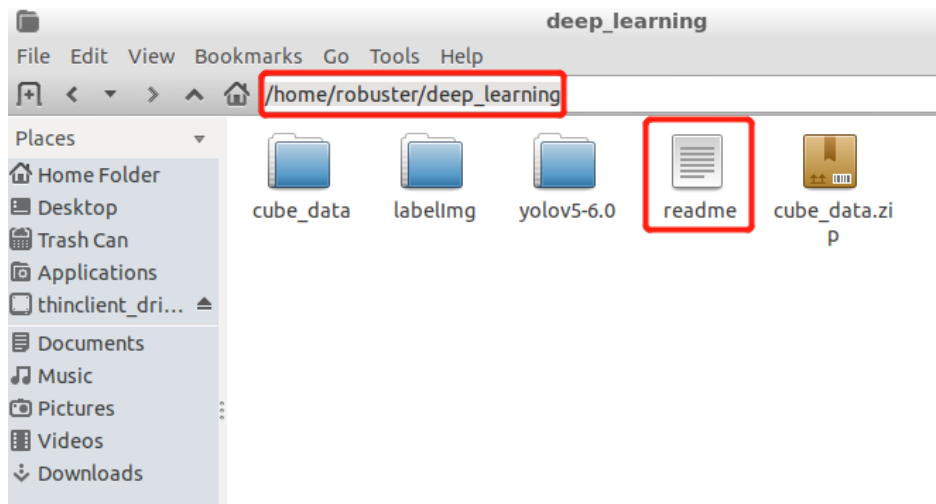
图像标注包括图像标注和视频标注，因为视频也是由连续播放的图像组成。图像标注一般要求标注人员使用不同的颜色来对不同的目标标记物进行轮廓识别，然后给相应的轮廓打上标签，用标签来概述轮廓内的内容。以便让算法模型能够识别图像中的不同标记物。

## 2. 数据集采集与标注实践

工欲善其事必先利其器，为了方便我们之后顺利的应用深度学习模型 YOLO，我们先在 BR280 中进行数据集采集与数据标注的工作。

### 2.1 数据集采集

1. 在 `/home/robuster/deep_learning` 目录下存在一个 `readme` 文件，里面书写了关于深度学习的所有操作指令，可自行打开查看。



2. 进行图片采集之前，首先将机械臂调整到下图所示位置。运行 `beetle_ai/scripts` 下的 `ready.py` 文件可以达到这个目的。

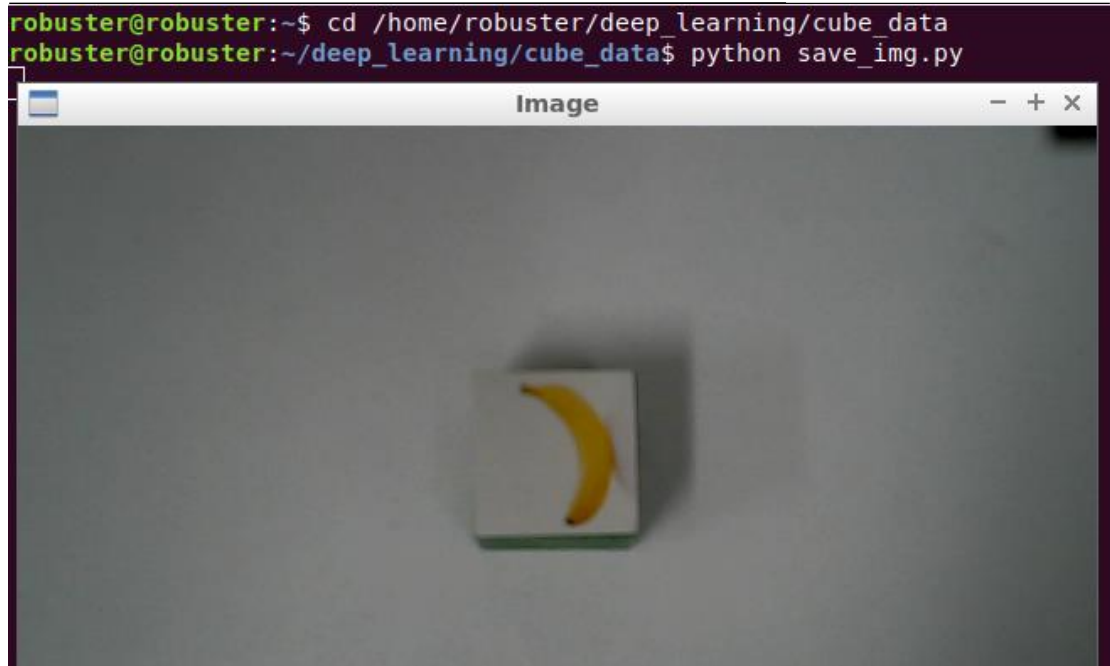
```
robuster@robuster:~$ cd beetle_ai/scripts/  
robuster@robuster:~/beetle_ai/scripts$ python ready.py
```



3. 通过使用单目相机拍摄图片来获取想要的数据集，为后面的训练模型做准备。在 PC 端工控机打开终端，执行以下指令。

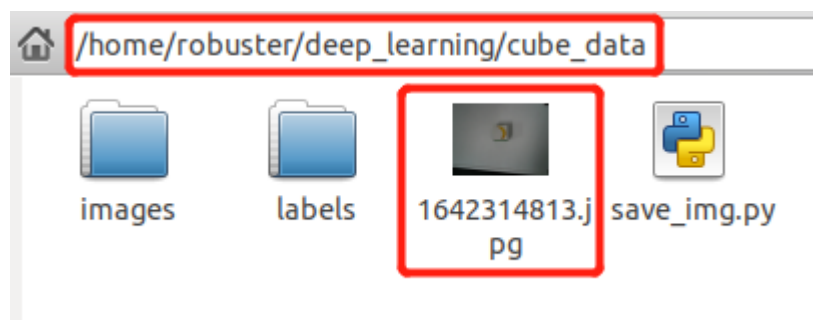
```
cd ~/deep_learning/cube_data  
  
python save_img.py
```





点击画面窗口，然后按下 Enter 键可以刷新画面窗口。

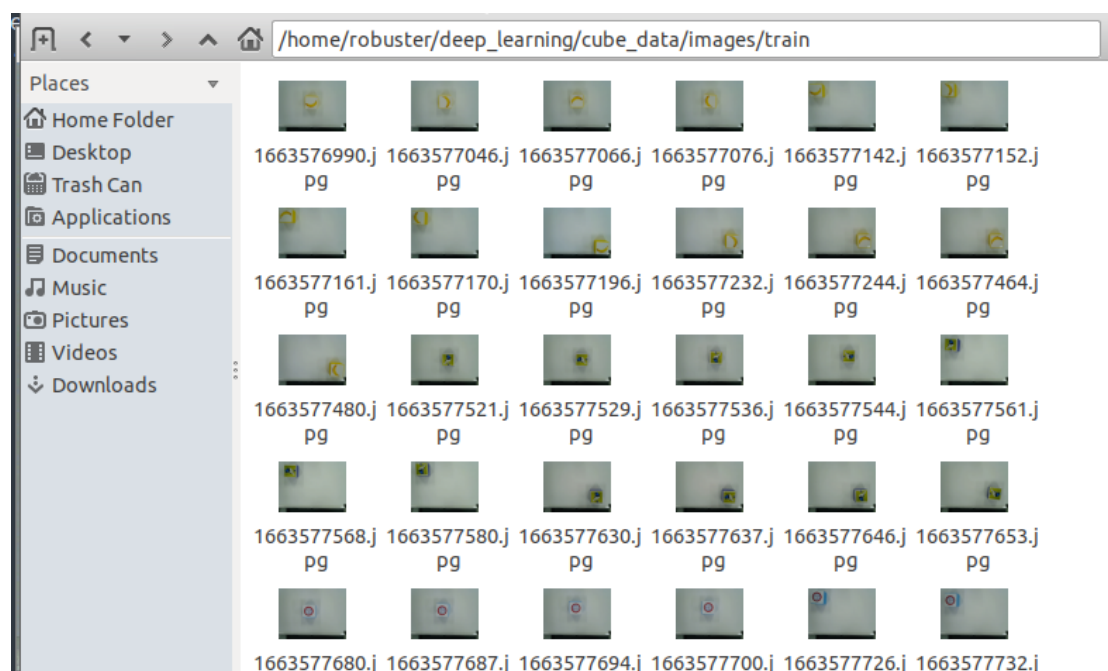
4. 按 S 键保存图片（保存图片之前，需要用鼠标点一下 Image 界面），此时图片被保存在/home/robuster/deep\_learning/cube\_data 目录下。



5. 将积木旋转 90°，点击 Enter 键直到界面刷新，然后按 S 键保存。



6. 重复上述操作，得到 4 张不一样的图片，每次旋转  $90^\circ$ 。
7. 将积木分别放置在左上角和右下角，重复上述操作，每个位置拍摄 4 张图片。也就是说，每个物体都要拍 12 张图片来作为训练的数据。
8. 将拍摄的图片拷贝到 `/home/robuster/deep_learning/cube_data/images/train` 目录下。因为我们要识别的图片有 5 类，所以我们总共要拍摄 60 张图片。复制操作成功后，把 `cube_data` 目录下的原照片删除，给对照样本清出空间，方便操作。



9. 每个物体还需要拍摄 3 张不同的图片，作为对照使用。注：这 3 张图片不能与上面已经拍摄的图片相同，你可以将积木摆在右上角和左下角，亦或是你喜欢的位置，但是不能与前面的重复。具体操作方法与上面相同。将拍好的 15 张图片拷贝到 `/home/robuster/deep_learning/cube_data/images/val` 目录下。复制完成后将 `cube_data` 下的源文件删除。



10.拍摄完所有的图片之后，按 Q 键退出图片采集。（先点击一下 Image 界面）

## 2.2 数据标注

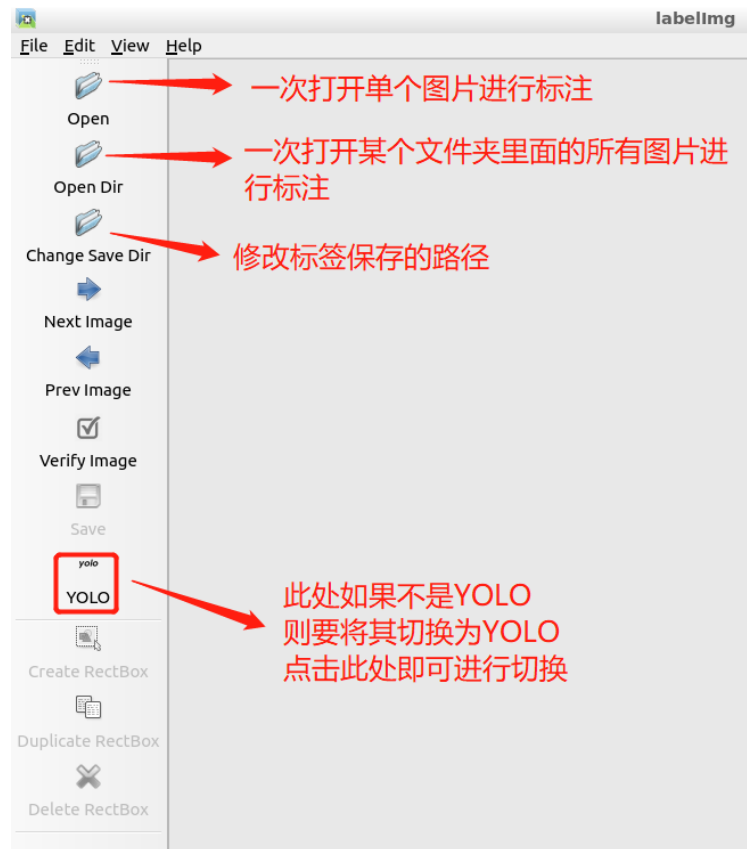
1. 打开终端，执行以下指令，开始进行数据标注。

```
cd ~/deep_learning/labelImg
```

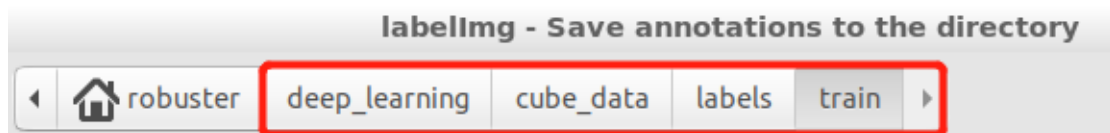
```
conda activate yolov5
```

```
python labelImg.py
```

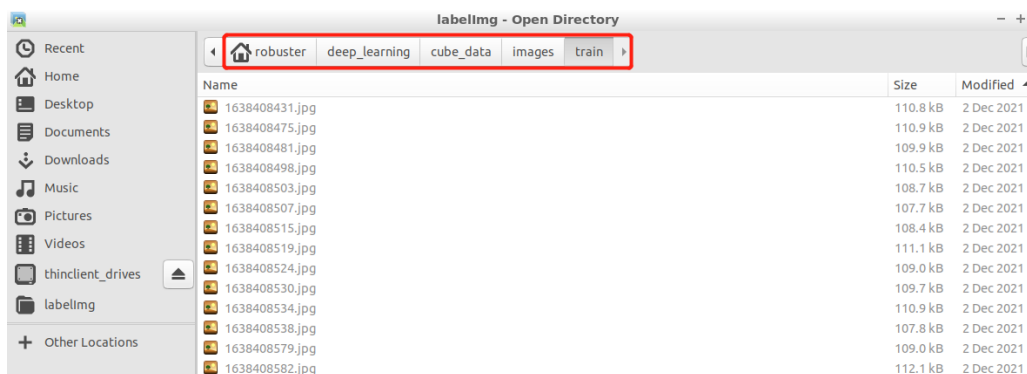
```
robuster@robuster:~/deep_learning/cube_data$ cd ~/deep_learning/labelImg
robuster@robuster:~/deep_learning/labelImg$ conda activate yolov5
(yolov5) robuster@robuster:~/deep_learning/labelImg$
```



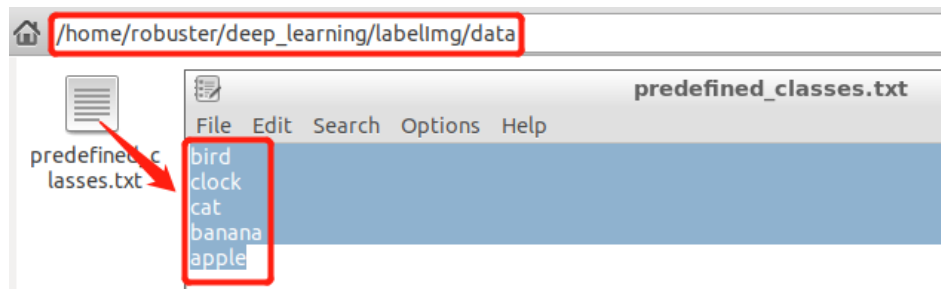
2. 先标注 train 文件里面的 60 张图片，点击 Change Save Dir，修改标签的保存路径，将路径修改为/home/robuster/deep\_learning/cube\_data/labels/train



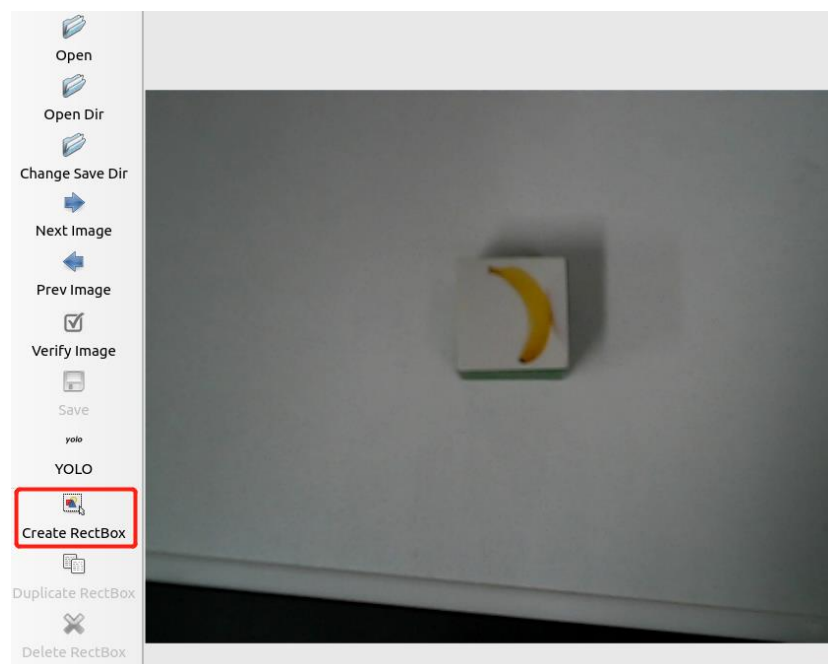
3. 点击 Open，找到你将要标注的图片，这些图片被保存在了 /home/robuster/deep\_learning/cube\_data/images/train 路径下。也可以选择 Open Dir，一次打开某个文件夹里面的所有图片，然后逐一进行标注。



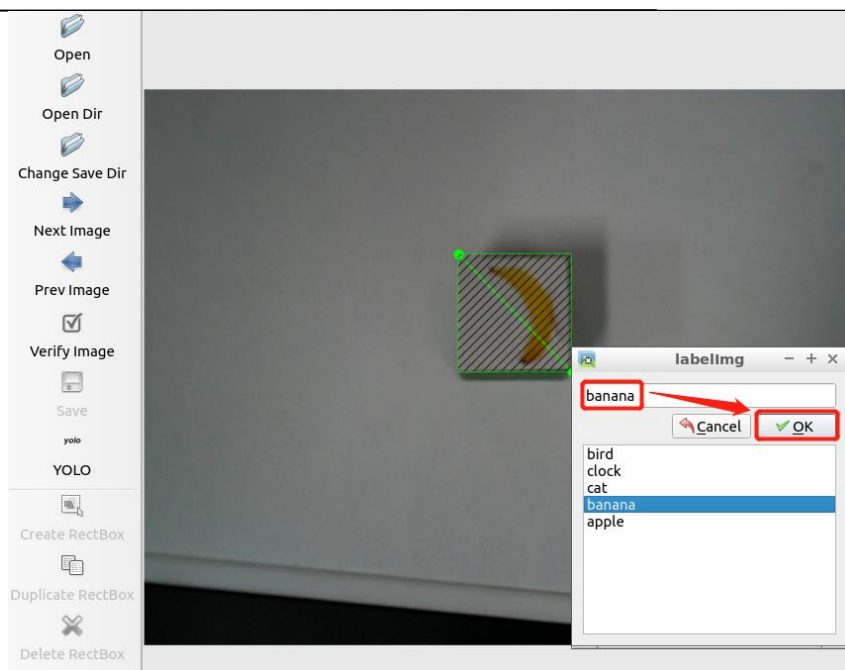
4. 进行标注之前，需要给图片分类，因为只拍摄了 5 种物体的图片，所以设置 5 类标签，标签的存放路径为/home/robuster/deep\_learning/labelImg/data。



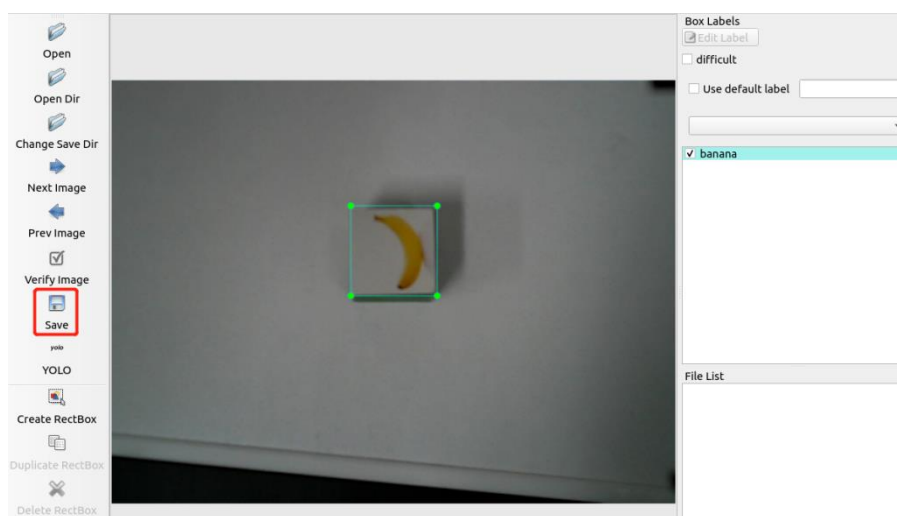
5. 点击 Create RectBox 开始标注。



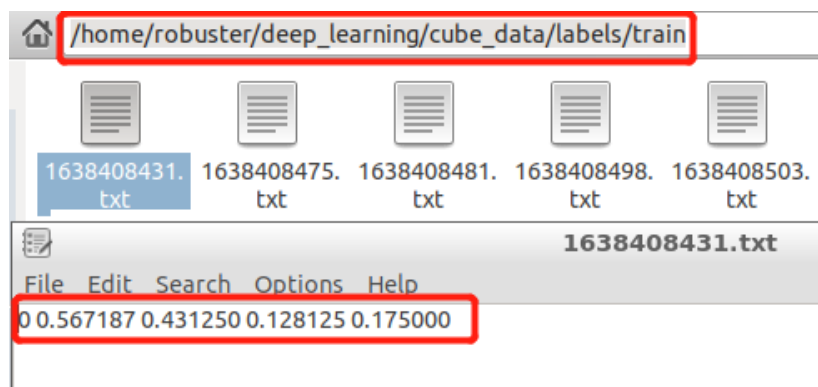
6. 创建边框，然后选择对应的标签分类。



7. 点击 Save 保存标签。在/home/robuster/deep\_learning/cube\_data/labels/train 里可以找到你保存的标签，它们以.txt 文件的格式存在。



里面的数字分别代表了分别表示：类别代号、标注框横向的相对中心坐标  $x\_center$ 、标注框纵向的相对中心坐标  $y\_center$ 、标注框相对宽度  $w$ 、标注框相对高度  $h$ 。注意  $x\_center$ 、 $y\_center$ 、 $w$ 、 $h$  为真实像素值除以图片的高和宽之后的值。



8. 标注完 train 文件里面的 60 张图片，再标注 val 文件里面的图片。将标签保存路径修改为 `/home/robuster/deep_learning/cube_data/labels/val`。重复上述操作，完成所有图片的标注工作。

### 3. 总结

本节主要介绍了深度学习相关概念与数据集采集与标注的内容，关于数据标注的工具还有很多，用法大致相同，想了解更多的数据标注工具可以点击下面链接。

<https://zhuanlan.zhihu.com/p/114404456>

请根据本节内容，完成下面练习。

1. 完成 BR280 中五个图片积木 60 张训练样本数据的采集工作；
2. 完成 BR280 中五个图片积木 15 张对照样本数据的采集工作；
3. 完成 60 张图片的标注工作；
4. 完成 15 张图片的标注工作。





— 深圳史河机器人科技有限公司 —



深圳史河机器人科技有限公司

[www.robotplusplus.com.cn](http://www.robotplusplus.com.cn)

深圳市龙华区龙华街道清华社区建设东路青年创业园C栋4层412-415房