大家好，我是北小菜，今天介绍 **视觉模型训练第2讲，**今天主要介绍如何收集样本，如何标注样本，如何将标注的样本转换为yolo格式的训练数据集

**下面的表格是视觉模型训练教程的目录**

|  |  |  |
| --- | --- | --- |
| **视觉模型训练教程** | **大致内容** | **视频录制时间** |
| 视觉模型训练第1讲 | 介绍了首次发布开源模型训练框架：  <https://gitee.com/Vanishi/BXC_AutoML> | 2024/4/27 |
| 视觉模型训练第2讲 | 主要介绍如何收集样本，如何标注样本，如何将标注的样本转换为yolo格式的训练数据集 | 2024/9/1 |

**关于我对模型训练的看法**：人工智能发展至今，无论是比较小的视觉模型还是大语言模型，大部分公司的模型网络结构差别都不会特别大，当下没有哪一个公司的模型结构真的是遥遥领先，也包括此前风靡全球的chatgpt。之所以有一点区别存在，最主要的因素就是样本集。所以，当下样本才是决定模型最重要的因素，超越调参。

今天会用到的工具如下：

1. 视频分割图片工具hs
   1. <https://gitee.com/Vanishi/BXC_hs>
2. 样本标注工具labelme
   1. 网盘下载：[https://pan.quark.cn/s/e255a8c4dd2e](https://pan.quark.cn/s/e255a8c4dd2e" \l "/list/share)
   2. github下载：<https://github.com/labelmeai/labelme/releases>
3. labelme标注的样本转换为yolo数据集的工具 labeltools
   1. <https://gitee.com/Vanishi/BXC_AutoML/tree/master/labeltools>

今天的讲解视频会尽可能，从最简单的角度出发，争取让足够多的用户学会标注样本

接下来讲解步骤如下：

1. 准备一份视频文件（用作制作样本）
2. 使用工具hs将准备的视频文件分割成图片文件夹
3. 使用工具labelme将分割好的图片文件夹标注成样本
4. 将上一步labelme标注的样本，转换为yolo格式的训练数据集