本指南介绍如何使用YOLOv5训练您自己的自定义数据集。

在开始之前：

克隆此repo，下载教程数据集，然后安装requirements.txt依赖项，包括Python>=3.8

和PyTorch>=1.7

$ git clone https://github.com/ultralytics/yolov5 # clone repo

$ cd yolov5

$ pip install -r requirements.txt # install dependencie

制作数据集

1创建数据集

yamlCOCO128是一个小型教程数据集，由cocotrain2017中的前128个图像

组成。这些相同的128个图像用于训练和验证，以验证我们的训练管道能够进行过度拟

合。

data/coco128.yaml，如下所示，是数据集配置文件，它定义了1）用于自动下载的可

选下载命令/URL，2）指向训练图像目录的路径（或指向带有训练图像列表的\*.txt文

件的路径），3）与验证图像相同，4）类数，5）类名列表：

2创建标签

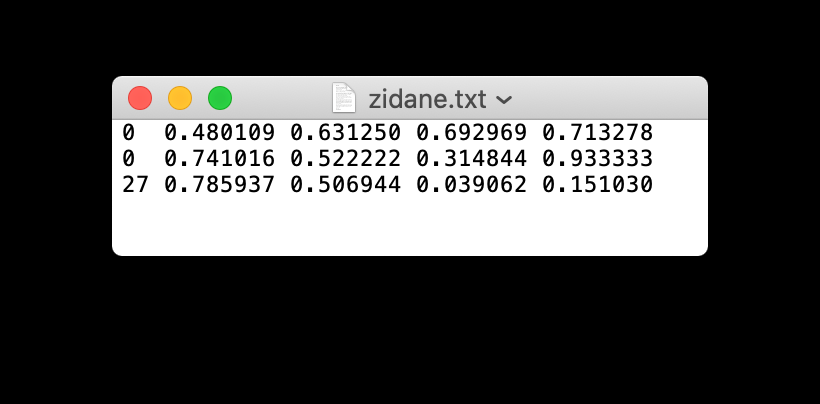
在使用了CVAT，makesence.ai或Labelbox为图像添加标签，将标签导出为YOLO格式，每个图像有一个\*.txt文件（如果图像中没有对象，则不需要\*.txt文件）。\*.txt文件规范为：

1. 每个对象一行(每张图一行)

2. 每一行都是class x\_center y\_center width height这种格式。

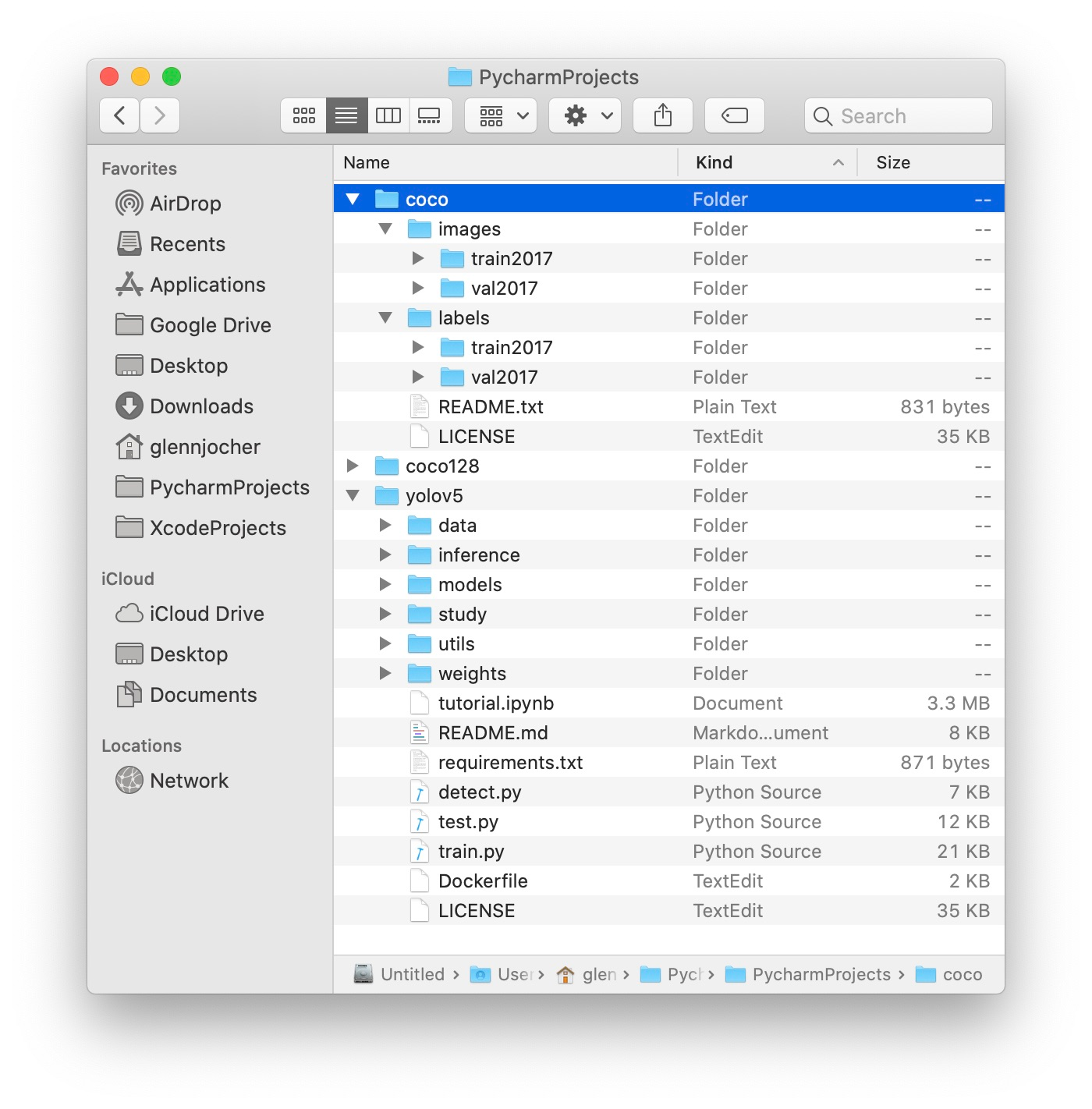
3. 框坐标必须是标准化的xywh格式归一化（从0到1）。如果框是以像素为单位的，请将x\_center和width除以图像宽度，将y\_center和height除以图像高度。

4. 类号索引为零（从0开始）



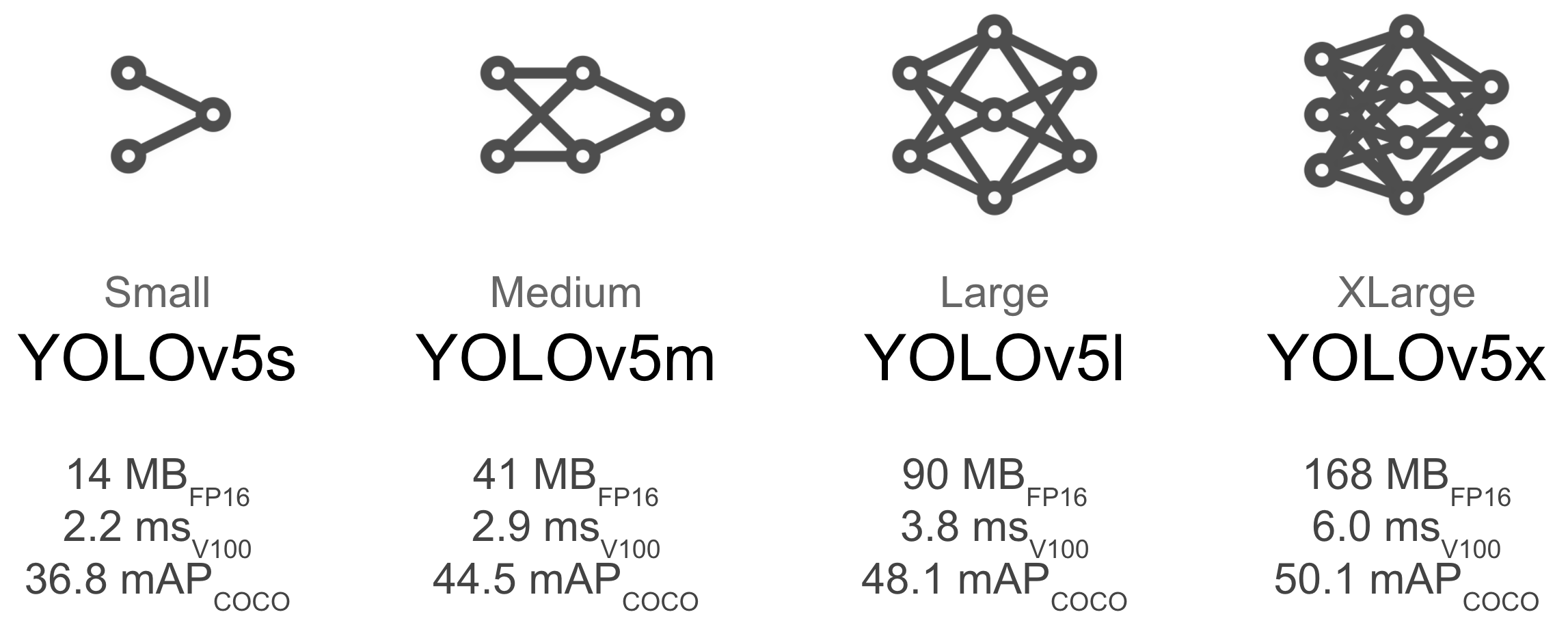
三。组织目录

根据下面的示例,组织train和val图像和标签。在本例中，我们假设/coco128位于/yolov5目录旁边。YOLOv5通过将每个图像路径中的最后一个/images/实例替换为/labels/来自动为每个图像定位标签。例如：



4选择模型

选择一个预训练模型开始训练。在这里，我们选择YOLOv5s，最小和最快的模式可用。请参阅我们的自述表以获得所有模型的完整比较



5训练

通过指定数据集、批量大小、图像大小以及预训练的--weights YOLOv5s.pt（推荐）或随机初始化的--weights''--cfg YOLOv5s.yaml（不推荐），在COCO128上训练YOLOv5s模型。预训练的重量是从最新的YOLOv5版本中自动下载的。

所有训练结果都保存到runs/train/中，运行目录递增，即runs/train/exp2、runs/train/exp3等。有关更多详细信息，请参阅我们的Google Colab笔记本的“训练”部分。

Requirements.txt内容：

# pip install -r requirements.txt

# base ----------------------------------------

Cython

matplotlib>=3.2.2

numpy>=1.18.5

opencv-python>=4.1.2

Pillow

PyYAML>=5.3.1

scipy>=1.4.1

tensorboard>=2.2

torch>=1.7.0

torchvision>=0.8.1

tqdm>=4.41.0

# logging -------------------------------------

# wandb

# plotting ------------------------------------

seaborn>=0.11.0

pandas

# export --------------------------------------

# coremltools>=4.1

# onnx>=1.8.1

# scikit-learn==0.19.2 # for coreml quantization

# extras --------------------------------------

thop # FLOPS computation

pycocotools>=2.0 # COCO mAP

可视化：

重量和偏差记录(新的）

权重和偏差（W&B）现在与YOLOv5集成，用于训练运行的实时可视化和云记录。这允许更好的运行比较和内省，以及改进团队成员之间的可见性和协作。要启用W&B日志记录，请安装wandb，然后正常训练（第一次使用时将指导您进行设置）。

$ pip install wandb

在训练期间，您将在https://www.wandb.com/，您可以使用W&B报告工具创建详细的结果报告。

本地日志记录：

默认情况下，所有结果都记录到runs/train，并为每个新训练创建一个新的实验目录runs/train/exp2、runs/train/exp3等。查看train和test jpg以查看马赛克、标签、预测和增强效果。注：马赛克数据加载器用于训练（如下所示），这是Ultralytics开发的新概念，首次出现在YOLOv4中。

训练损失和绩效指标也记录到Tensorboard和自定义结果.txt打印为结果.png（下）培训结束后。在这里，我们展示了从零开始（蓝色），从预训练的-weights YOLOv5s.pt（橙色），经过COCO128到300个历代训练的YOLOv5s。

环境

YOLOv5可以在以下任何一个最新验证的环境中运行（所有依赖项包括CUDA/CUDNN、Python和PyTorch都已预装）：

1. 带有免费GPU的Google Colab和Kaggle笔记本电脑：

2. 谷歌云深度学习虚拟机。请参阅GCP快速入门指南

3. 亚马逊深度学习。请参阅AWS快速入门指南

4. Docker图像。请参阅Docker快速入门指南Docker Pulls