visio图像转角 - {x: 76.1, y: 345.5}

部署：

onnx：

netron.start('f.onnx')

|  |
| --- |
| 对shape / size使用int转换 |
| 上采样使用scale\_factor指定参数 (而不是size) |
| reshape、view操作的batch维度为-1 |
| export的dynamic\_axes只指定batch维度 |

|  |  |  |
| --- | --- | --- |
| 模型导出torch.onnx： | | |
| export(model, args, f, input\_names, output\_names) | | 导出模型的onnx文件 |
| args：模型的输入 | f：导出文件名称 |  |
| torch.onnx.is\_in\_onnx\_export() | | 检查是否在导出模型 |

|  |  |  |
| --- | --- | --- |
| 模型调用onnxruntime： | | |
| get\_available\_providers() | | 获取可用设备 |
| InferenceSession(path\_or\_bytes) | | 实例化onnx文件的推断模型 |
| 实例方法 | get\_inputs() | 返回输入变量列表 |
| get\_outputs() | 返回输出变量列表 |
| run(ouput\_names, input\_feed) | 返回推理结果 |

torchscript：

|  |  |  |
| --- | --- | --- |
| 模型导出torch.jit： | | |
| load(file) | | 加载模型 |
| trace(model, args, strict) | | 创建模型 |
| 实例方法 | save(file) | 导出模型文件 |

openvino：

|  |  |
| --- | --- |
| mo命令： | |
| -w <input-model> | 指定模型文件 |
| -o <output-dir> | 指定输出目录 |
| --compress\_to\_fp16 | 指定半精度 |
| --input\_shape <shape> | 指定输入尺寸 |

**推理runtime：**

|  |  |  |
| --- | --- | --- |
| runtime.Core() | | 实例化内核 |
| 实例属性 | available\_devices | 可用设备 |
| 实例方法 | read\_model(xml) | 加载并返回模型 |
| compile\_model(model, device\_name) | 返回编译模型 |

|  |  |  |
| --- | --- | --- |
| runtime.CompiledModel(other) | | 实例化编译模型 |
| 实例方法 | input(idx) / output(idx) | 返回io结点 |
| \_\_call\_\_(\*arrays) | 前向传播 |

|  |  |  |
| --- | --- | --- |
| 结点： | | |
| 实例属性 | any\_name | 名称 |
| shape | 尺寸 |
| element\_type | 数据类型 |

coco：

from pycocotools.coco import COCO

|  |  |  |
| --- | --- | --- |
| COCO(annotation\_file) | | 初始化数据集 |
| 实例属性 | imgs | 图片信息 |
| anns | 标注实例 |
| cats | 类别信息 |
| 实例方法 | getAnnIds(imgIds) | 获取标注实例Id |

diffusers：

|  |  |  |
| --- | --- | --- |
| DiffusionPipeline | | 模型基类 |
| 类方法 | from\_pretrained(name\_or\_path, torch\_dtype) | 加载预训练模型 |
| numpy\_to\_pil(images) | 图像转换 |
| 实例属性 | scheduler | 调度器 |
| 实例方法 | to(dtype\_or\_device) | 转换类型或设备 |
| save\_pretrained(dir) | 保存到文件夹 |

|  |  |  |
| --- | --- | --- |
| **调度器：** | | |
| DDPMScheduler(num\_timesteps) | | 随机采样 (离线) |
| DDIMScheduler(num\_timesteps) | | 确定性采样 (实时) |
| 类方法 | from\_config(cfg) | 从配置创建 |
| 实例属性 | config | 配置 |
| 实例方法 | step(model\_out, timestep, sample) |  |

工具training\_utils：

|  |  |
| --- | --- |
| EMAModel | EMA模型 |

mediapipe：

模型solutions：

|  |  |  |
| --- | --- | --- |
| 检测器基类 | close() | 关闭检测器 |
| reset() | 重启检测器 |
| 关键点信息 | landmark | 返回关键点坐标 |

|  |  |  |  |
| --- | --- | --- | --- |
| 绘制： | | | |
| drawing\_utils | draw\_landmarks(image, landmark, connections,  landmark\_style, connection\_style) | OpenCV | 绘制关键点 |
| draw\_detection(image, detection) | 绘制检测框 |
| plot\_landmarks(landmark, connections) | Matplotlib | 绘制关键点 |
| drawing\_styles | get\_default\_hand\_landmarks\_style | 手掌 | 关键点样式 |
| get\_default\_hand\_connections\_style | 连接样式 |

|  |  |  |
| --- | --- | --- |
| 手掌hands： | | |
| HAND\_CONNECTIONS | | 关键点连接方式 |
| Hands(static\_image\_mode, max\_num\_hands,  model\_complexity, min\_detection\_confidence,  min\_tracking\_confidence) | | 实例化手掌关键点检测器 |
| 实例方法 | process(image) | 返回检测结果Output |
| Output  实例属性 | multi\_hand\_landmarks | 返回关键点绝对坐标列表 |
| multi\_hand\_world\_landmarks | 返回关键点相对坐标列表 |
| multi\_handedness | 返回手性检测结果 |
|  | | |

|  |  |  |
| --- | --- | --- |
| 人脸face\_detection： | | |
| get\_key\_point(Box, idx) | | 返回关键点坐标 |
| FaceDetection(min\_detection\_confidence, model\_selction) | | 实例化人脸关键点检测器 |
| 实例方法 | process(image).detections | 返回检测结果Box |

mindspore：

环境context：

|  |  |  |
| --- | --- | --- |
| context.setcontext(mode, devicd\_target="CPU") | | 配置运行环境 |
| mode：模式 | context.GRAPH\_MODE | 图模式 |
| context.PYNATIVE\_MODE |  |

数据dataset：

|  |  |  |
| --- | --- | --- |
| ds.GeneratorDataset(seq, column\_names=seq) | | 返回seq转换的数据集GD |
| 实例方法 | get\_dataset\_size() | 返回数据集大小 |
| batch(batch\_size) | 将指定数量数据组合成一批 |
| repeat(repeat\_size) | 数据集数量倍增 |

网络nn：

|  |  |
| --- | --- |
| 基类： | |
| Cell() | 神经网络基类 |
| GraphKernel() | 图形内核基类 |
| CellList() | 功能如list |
| SequentialCell() | MindSpore有序列表 |

|  |  |  |
| --- | --- | --- |
| 卷积网络： | | |
| Conv1d(in\_channels, out\_channels, kernel\_size, stride, padding, groups) | | 一维卷积层 |
| Conv1dTranspose(in\_channels, out\_channels, kernel\_size, stride, padding, groups) | | 一维反卷积 |
| Conv2d(in\_channels, out\_channels, kernel\_size, stride, padding, groups) | | 二维卷积层 |
| Conv2dTranspose(in\_channels, out\_channels, kernel\_size, stride, padding, groups) | | 二维反卷积 |
| in\_channals：输入通道数 | kernel\_size：卷积核尺寸 |  |
| out\_channels：输出通道数 | stride：步长 |  |

openai：

|  |  |  |
| --- | --- | --- |
| 客户端： | | |
| OpenAI(api\_key, base\_url="https://api.deepseek.com") | | 实例化客户端 |
| 实例方法 | chat.completions.create(  model=["deepseek-chat", "deepseek-reasoner"][0],  messages=[  {"role": "system", "content": "You are a helpful assistant."},  {"role": "user", "content": "Write a haiku."}  ],  max\_tokens, temperature, stream=False  ) | 生成文本  (生成代码时温度0) |
| images.generate(  prompt="A cute baby sea otter",  n=2,  size="1024x1024"  ) | 生成图像 |
| embeddings.create(  model="text-embedding-3-large",  input="The food was delicious and the waiter..."  ) | 生成向量 |

|  |  |  |
| --- | --- | --- |
| types.chat.chat\_completion： | | |
| ChatCompletion | | 文本回复 |
| 实例属性 | choices | 选项列表 |
| Choice | | 选项 |
| 实例属性 | message.content | 消息 |

transformers：

|  |  |  |
| --- | --- | --- |
| GenerationMixin | | 模型基类 |
| Qwen2\_5\_VLForConditionalGeneration | | Qwen |
| 实例方法 | generate(\*\*kwargs) | 生成 |

|  |  |  |
| --- | --- | --- |
| PreTrainedModel | | 预训练模型 |
| 类方法 | from\_pretrained(name\_or\_path, torch\_dtype, device\_map) | 加载预训练模型 |
| 实例属性 | base\_model | nn.Module |
| 实例方法 | enable\_input\_require\_grads / disable\_input\_require\_grads() | 设置输入梯度 |

|  |  |  |
| --- | --- | --- |
| AutoProcessor | | 自动处理器 |
| 类方法 | from\_pretrained(name\_or\_path, torch\_dtype, device\_map) | 加载预训练模型 |
| 实例方法 | apply\_chat\_template(conversation) | 处理文本对话 |

|  |  |  |
| --- | --- | --- |
| Trainer(model, TrainingArguments,  train\_dataset, eval\_dataset, tokenizer) | | 训练器 |
| 实例方法 | train(resume\_from\_checkpoint, optuna.trial) | 开始 |

torch：

|  |  |
| --- | --- |
| pip安装： | |
| 1.9.0 | pip install torch==1.9.0+cu102 torchvision==0.10.0+cu102 -f https://download.pytorch.org/whl/torch\_stable.html |

flops, params = thop.profile(model, inputs, verbose=False)

# 测试浮点运算量、模型参数量

|  |  |
| --- | --- |
| 属性访问： | |
| x.shape | 尺寸 |
| x.size(i) | 指定维度信息 |
| x.numel() | 元素数 |
| x.dim() | 维度数 |
| x.item() | 标量 |
| x.data | 数据 |
| x.grad | 梯度 |

|  |  |
| --- | --- |
| 设备： | |
| device('cuda:0') | 实例化CPU设备 |
| x.to(device|type, non\_blocking=True) | 将实例在设备上运行/修改数据类型 |
| x.cpu() | 将实例在CPU上运行 |
| x.cuda() | 将实例在GPU上运行 |
| x.zero\_() | 梯度置0 |

|  |  |  |
| --- | --- | --- |
| 张量转换： | | |
| tensor(x, dtype=torch.float32) | 返回序列对应的tensor | |
| from\_numpy(x) |
| x.numpy() | 返回tensor对应的array | |
| x.tolist() | 返回tensor对应的列表 | |
| x.view(\*shape) / x.resize\_as(x) | 返回尺寸转换后的tensor | |
| x.float() | 数据类型转为 | float32 |
| x.byte() | uint8 |

|  |  |
| --- | --- |
| 张量操作： | |
| x.clone() | 深拷贝张量所有属性 |
| x.detach() | 张量脱离计算图 |
| x.contiguous() | 使张量内存连续 |

|  |  |  |
| --- | --- | --- |
| 张量维度变换： | | |
| x.flatten(start\_dim, end\_dim) | 张量展平 | |
| x.unflatten(dim, sizes) | 张量展开 | |
| x.narrow(dim, start, length) | 截取矩阵块 | |
| x.flip(dims) | 按维度翻转 | |
| x.unsqueeze(dim) | 添加新维度 | |
| x.squeeze(dim) | 删除冗余维度 | |
| x.repeat(\*shape\_blow) | 各维度数据复制扩增 | |
| x.transpose(dim1, dim2) | 交换对应维度 | |
| x.permute(\*dims) | 重排维度 | |
| cat(xs, dim) | 在原维度 | 合并tensor |
| stack(xs, dim) | 创建维度 |
| x.split(lengths, dim) | 按长度 | 分割张量 |
| x.chunk(chunks, dim) | 按块 |

|  |  |
| --- | --- |
| 张量创建： | |
| zeros / empty / ones(shape) | 默认值**填充** |
| eye(shape) | 单位阵 |
| full(shape, element) | 指定值**填充** |
| linspace(min, max, n) / arange(min, max, h) | 等差**数列** |
| meshgrid(xs) | 返回多维坐标网格点矩阵 |

|  |  |  |
| --- | --- | --- |
| 随机： | | |
| manual\_seed(idx) | 选定指定的随机数初态 | |
| rand(\*size) | 返回均匀分布 |  |
| randint(low, high, size) |  |
| x.uniform\_(low, high) |
| randn(\*size) | 返回正态分布 |  |
| normal(mu, sigma, size) |  |
| bernoulli(x, p) | 返回伯努利分布 | |
| randperm(n) | [0, n) 整数，顺序打散 | |

|  |  |
| --- | --- |
| 张量运算： | |
| x.add\_ / x.mul\_ / x.round\_() | 就地运算 |
| clamp(x, min, max) | 元素限制 |
| sign / abs / sqrt / pow(x) | 基本 |
| all / any(x) | 非零**判断** |
| mean / sum / prod / norm() | 均值 / 和 / 积 / 二范数 |
| max / min / argmax / argmin(x) | 最值 |
| maximum / minimum(x1, x2) | 逐元素最值 |
| sin / cos / tan / asin / acos / atan(x) | **三角函数** |
| exp / log / log2 / log10(x) | **指数函数** |
| round / ceil / floor(x) | 取整 |
| cumsum / diff(x) | 累加 / 差分 |
| std / var(x) | 统计 |
| erf(x) |  |
| cosine\_similarity(x1, x2, dim) | 余弦相似度 |

|  |  |
| --- | --- |
| 功能： | |
| nonzero(x) | 筛选出非零元素 |
| scatter(x, dim, index, value) | 按位置填充值 |
| equal(x1, x2) | 判断相等 |
| where(bool, true\_out, false\_out) | 返回布尔值对应tensor |
| x.gather(dim, index) | 元素索引 |
| x.unique(sorted, dim) | 返回不重复元素的张量 |
| sort(x, dim, descending) | 返回排序结果 |
| x.backward(retain\_graph=False) | 反向传播梯度 |

|  |  |
| --- | --- |
| 通用： | |
| save(data, file) | 将data存储到file |
| load(file, map\_location) | 读取file并返回存储的数据 |
| no\_grad() | 零梯度修饰器 (\_\_wrapped\_\_原函数) |
| set\_printoptions(precision, sci\_mode) | 设置输出格式 |

梯度autograd：

|  |  |
| --- | --- |
| set\_detect\_anomaly(True) | 设置梯度异常检测器 |

|  |  |  |
| --- | --- | --- |
| ctx | | 上下文管理器 |
| 实例属性 | saved\_tensors | 获取用于反向传播的张量 |
| 实例方法 | save\_for\_backward(\*args) | 保存用于反向传播的张量 |

|  |  |  |
| --- | --- | --- |
| Function() | | 实例化传播函数 |
| 静态方法  (重写) | forward(ctx, \*args) | 前向传播 |
| backward(ctx, \*grad\_outputs) | 反向传播 |
| 类方法 | apply(x) | 调用 |

函数functional：

import torch.autograd.functional as aF

|  |  |
| --- | --- |
| jacobian(func, inputs) | 生成函数的雅可比矩阵 |

显卡cuda：

|  |  |
| --- | --- |
| 显卡操作： | |
| empty\_cache() | 释放显卡缓存 |
| synchronize() | 等待所有协程完成 |

|  |  |
| --- | --- |
| 信息获取： | |
| is\_available() | 显卡是否可用 |
| device\_count() | 显卡数量 |
| memory\_reserved(i) | 返回当前的GPU占用量 |

|  |  |  |
| --- | --- | --- |
| 显卡实例： | | |
| get\_device\_properties(i) | | 返回显卡实例 |
| 实例属性 | name | 显卡名称 |
| total\_memory | 显存大小 |

|  |  |  |
| --- | --- | --- |
| 事件： | | |
| Event(enable\_timing=True) | | 返回事件 |
| 实例方法 | record() | 更新记录点 |
| elapsed\_time(event) | 返回时间差 |

混合amp：

|  |  |
| --- | --- |
| autocast(enabled) | 实例化混合精度训练装饰器 |

|  |  |  |
| --- | --- | --- |
| GradScaler() | | 实例化梯度缩放器 |
| 实例方法 | scale(loss) | 缩放损失值 |
| step(optim) | 装饰优化器，检查非有限梯度 |
| update() | 启动迭代 |

分布distributed：

|  |  |  |
| --- | --- | --- |
| 环境变量： | | |
| CUDA\_VISIBLE\_DEVICES | cuda可见设备 | '0,1,2'：多 cuda |
| ''：GPU |
| RANK | 进程在所有进程中的级别 | |
| LOCAL\_RANK | 进程在本地进程中的级别 | |
| WORLD\_SIZE | 进程组中的进程数量 | |

|  |  |
| --- | --- |
| 分布式训练： | |
| init\_process\_group('nccl', 'env://', world\_size, rank) | 初始化进程池 |
| destroy\_process\_group() | 释放进程池 |
| barrier() | 等待所有进程到达同步点 |

from torch.utils.data.distributed import DistributedSampler

|  |  |  |
| --- | --- | --- |
| DistributedSampler(dataset, num\_replicas, rank) | | 实例化分布式数据集 |
| 实例方法 | set\_epoch(epoch) | 设置训练轮次 |

from torch.nn.parallel import DataParallel

|  |  |  |
| --- | --- | --- |
| DataParallel(module, device\_ids, output\_device) | | 实例化数据并行器 |
| 实例属性 | module | 网络模型 |
| 实例方法 | train(mode) | 运算模式切换 |

激活gradcam：

import pytorch\_grad\_cam as gradcam

|  |  |  |
| --- | --- | --- |
| EigenGradCAM(model, target\_layers, use\_cuda) | | 返回Grad-CAM绘制器 |
| 实例方法 | module | 网络模型 |

通用utils：

|  |  |
| --- | --- |
| 目标model\_targets： | |
| ClassifierOutputTarget(category).\_\_call\_\_(output) | 返回对应类别得分 |

进程multiprocessing：

|  |  |
| --- | --- |
| set\_sharing\_strategy("file\_system") | 设置共享策略 (默认文件描述符) |
| spawn(func, args, daemon, start\_method) | 启动进程 (指定启动方法) |

网络nn：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 通用： | | | | |
| Parameter(x, requires\_grad) | | 实例化模型参数 | | |
| Sequential(\*modules) /  ModuleList(list) / ModuleDict(dict\_or\_items) | | 实例化网络模块容器 | | |
| Module() | | 网络模块基类 | | |
| 实例属性 | parameters() | 参数生成器 | | |
| named\_parameters | 参数字典生成器 | | |
| \_modules | 子模块字典 | | |
| modules() | 子模块生成器 | | |
| state\_dict() | 模型参数生成器 | | |
| 实例方法  (重写) | forward(\*args) | 前向传播方法 | | |
| extra\_repr() | 额外信息 | | |
| 实例方法 | load\_state\_dict(dict, strict) | 加载模型参数 | | |
| apply(fn: Callable[module]) | 递归地将函数应用于模块 | | |
| train() / eval() | 运算模式切换 | | |
| half() / float() / fuse() / type(dtype) | 数据类型切换 | | |
| add\_module(name, module) | 添加模块到实例属性 | | |
| register\_buffer(name, tensor, persistent) | 注册网络成员变量 (可不计入state\_dict) | | |
| register\_forward\_hook  (hook: Callable[module, x, output]) | 注册挂钩  返回handle | .remove()  释放 | 前传挂钩 |
| register\_full\_backward\_hook  (hook: Callable[module, gx, gy]) | 反传挂钩 |

|  |  |
| --- | --- |
| 基本模块 shape = [B, D]： | |
| Dropout(proba, inplace) | 随机置零输入值 |
| Identity() | 恒等映射 |
| Linear(in\_features, out\_features) | 线性变换 |

|  |  |
| --- | --- |
| timm.models.layers： | |
| to\_2tuple(x) | 转化成二元组 |
| DropPath(drop\_prob) | 随机置零样本 |
| DropBlock2d(drop\_prob, block\_size, batchwise) | 随机置零连续区域 |

|  |  |  |
| --- | --- | --- |
| CNN模块 shape = [B, C, H, W]： | | |
| Conv1d(in\_ch, out\_ch, kernel\_size, stride, padding, dilation) | | 卷积 |
| Conv2d(in\_ch, out\_ch, kernel\_size, stride, padding, dilation) | |
| ConvTranspose2d(in\_ch, out\_ch, kernel\_size, stride, padding) | | 反卷积 |
| Upsample(scale\_factor, mode='nearest') | | 上采样 |
| AdaptiveAvgPool2d(ouput\_size) | | 平均池化 |
| AvgPool2d(kernel\_size, stride, padding) | |
| MaxPool2d(kernel\_size, stride, padding) | | 最大池化 |
| BatchNorm2d(channels, eps, momentum) | | 批标准化 |
| GroupNorm(groups, channels) | | 组标准化 |
| in\_channels：输入通道数 | kernel\_size：卷积核尺寸 |  |
| out\_channels：卷积核通道数 | stride：步长 |  |
| padding：填充规模 | dilation：膨胀尺寸 |  |



|  |  |
| --- | --- |
| Transformer： | |
| MultiheadAttention(embed\_dim, num\_heads,  dropout, bias=False) | 多头注意力网络 (dropout在softmax之后) |
| head\_dim = embed\_dim / num\_heads |
| 单头注意力： |
| TransformerEncoderLayer(d\_model, nhead,  dim\_feedforward, dropout, activation) | 编码器 |
| TransformerDecoderLayer(d\_model, nhead,  dim\_feedforward, dropout, activation) | 解码器 (比编码器多了对memory的MSA) |
| LayerNorm(normalized\_shape) | 层标准化 |
| Transformer(d\_model, nhead,  num\_encoder\_layers, num\_decoder\_layers,  dim\_feedforward, dropout, activation) | 编码器+解码器 |

|  |  |
| --- | --- |
| NLP模块 shape = [L, B, D]： | |
| Embedding(num\_embeddings, embedding\_dim, max\_norm) | 数据编码 |
| RNN(input\_size, hidden\_size, num\_layers, batch\_first) | 循环网络 |
| RNNCell((input\_size, hidden\_size) | 循环网络单元 |
| LSTM(input\_size, hidden\_size, num\_layers, batch\_first) | 循环网络 |

函数functional：

import torch.nn.functional as F

|  |  |
| --- | --- |
| 功能函数： | |
| one\_hot(x, num\_classes) | 返回one hot矩阵 |
| normalize(x, p, dim) | 向量规范化 |
| grid\_sample(input, grid, align\_corners) | 以[-1, 1]值做网格采样 |
| unfold(x, kernel\_size, dilation=1, padding=0, stride=1) | 滑动窗口取值 [B, C k^2, N] |
| scaled\_dot\_product\_attention(query, key, value,  attn\_mask, dropout\_p) | 多头注意力 out[B, ..., L, Ev]  q[B, ..., L, E], k[B, ..., S, E], v[B, ..., S, Ev] |

Logit变换：

|  |  |  |
| --- | --- | --- |
|  | 激活函数 | 值域 |
| x.sigmoid(x) |  | [0, 1] |
| tanh(x) |  | [-1, 1] |
| hardtanh(x, min, max) |  | [min, max] |
| softplus(x, beta) |  | [0, ∞] |
| leaky\_relu(x, negetive\_slope) |  | [-∞, ∞] |
| mish(x) |  |  |
| relu(x) |  | [0, ∞] |
| silu(x) |  | [0, ∞] |
| gelu(x) |  |  |

|  |  |
| --- | --- |
| 损失函数： | |
| softmax(x, dim) | 相对概率： |
| log\_softmax(x, dim) | 相对概率： |
| mse\_loss(x, y, reduction='sum') | 均方差： |
| nll\_loss(x, y) | 负对数似然： |
| cross\_entropy(x, y, weight) | 交叉熵：softmax → nll\_loss |
| binary\_cross\_entropy(x, y) | 二元交叉熵： |
| binary\_cross\_entropy\_with\_logits(x, y) | 二元交叉熵： |

|  |  |  |
| --- | --- | --- |
| 图像函数： | | |
| interpolate(input, scale\_factor, mode="nearest") | | 向上/向下采样指定倍数 |
| adaptive\_avg\_pool2d(input, shape) | | 自适应平均池化，得到目标尺寸图像 |
| 形参 | scale\_factor | 比例因子 |

初始化init：

|  |  |
| --- | --- |
| trunc\_normal\_(tensor, mean, std) | 截断正态分布，初始化weight |
| constant\_(tensor, val) | 填充定值，初始化bias |

剪枝utils.prune：

|  |  |
| --- | --- |
| global\_unstructured(param, method=L1Unstructured, \*\*kwargs) | 全局非结构化剪枝 |
| l1\_unstructured(module, name, amount) | 非结构化剪枝 |
| ln\_structured(module, name, amount, n, dim) | 结构化剪枝 |

优化optim：

|  |  |  |
| --- | --- | --- |
| Optimizer(params, defaults) | | 优化器基类，重写 step |
| [SGD](https://pytorch.org/docs/stable/generated/torch.optim.SGD.html)(params, lr, momentum, weight\_decay) | | 实例化SGD算法优化器 |
| [Adam](https://pytorch.org/docs/stable/generated/torch.optim.Adam.html)(params, lr, betas, eps, weight\_decay) | | 实例化Adam算法优化器 (在梯度上权重衰减) |
| [AdamW](https://pytorch.org/docs/stable/generated/torch.optim.AdamW.html)(params, lr, betas, eps, weight\_decay) | | 实例化AdamW算法优化器 (直接权重衰减) |
| 实例方法 | add\_param\_group(dict) | 添加参数组 |
| step() | 使用优化器优化变量 |
| zero\_grad() | 梯度信息置零 |
| lr：学习率 | momentum：动量 (0.78) | weight\_decay：L2 范数正则项系数 |

调度lr\_scheduler：

|  |  |  |
| --- | --- | --- |
| ReduceLROnPlateau(optimizer, factor, patience, threshold, min\_lr) | | 实例化损失监听器 |
| 实例方法 | step(loss) | 若loss稳定则更新学习率 |

|  |  |  |  |
| --- | --- | --- | --- |
| StepLR(optimizer, step\_size, gamma) | | 实例化按步更新器 | step\_size：优化次数 |
| gamma：学习率缩放 |
| LambdaLR(optimizer, lr\_lambda, last\_epoch) | | 实例化自定义更新器 | |
| 实例方法 | step() | 满足条件则更新学习率 | |

通用utils：

关卡checkpoint：

|  |  |
| --- | --- |
| checkpoint(func, \*args) | 执行运算，不保留中间激活值，节约内存空间 |

数据data：

|  |  |
| --- | --- |
| Dataset() | 数据集基类，重写\_\_init\_\_、\_\_getitem\_\_、\_\_len\_\_ |

|  |  |  |
| --- | --- | --- |
| DataLoader(dataset, batch\_size, shuffle, num\_workers, drop\_last) | | 返回分批加载器 |
| batch\_size：批大小 | shuffle：是否打乱 |  |
| num\_workers：线程数 | drop\_last：是否舍去余数 |  |

|  |  |
| --- | --- |
| random\_split(DS, [\*int]) | 随机拆分数据集，返回数据集元组 |

张量板tensorboard：

|  |  |  |
| --- | --- | --- |
| SummaryWriter(log\_dir) | | 实例化SW编辑器 |
| 实例方法 | add\_image(str, img, step, dataformats="HWC") | 添加图像 |
| add\_scalar(str, y, x) | 添加标量 |
| close() | 关闭编辑器 |

视觉torchvision：

模型models：

ResNet, VGG, ShuffleNet, DenseNet, MnasNet, MobileNet, SqueezeNet

数据集datasets：

|  |  |  |
| --- | --- | --- |
| CIFAR10(root, train, transform, download) | | 实例化CIFAR数据集 |
| 形参 | train | 是否为训练集 |
| transform | 转换器实例 |
| download | 是否下载 |
| 实例属性 | classes | 种类列表 |
| 实例方法 | \_\_getitem\_\_(item) | 返回 [img, target] |

运算ops：

|  |  |
| --- | --- |
| box\_iou(xyxy1, xyxy2) | 边界框IoU |
| nms(boxes, scores, iou\_threshold) | 不区分类别非极大抑制 |
| batched\_nms(boxes, scores, idxs, iou\_threshold) | 分类别非极大抑制 |

转换transforms：

|  |  |  |  |
| --- | --- | --- | --- |
| Compose(trans\_list) | 实例化变换器实例的组合变换器 | | |
| ToTensor() | 实例化tensor转化器，并归一化 | | |
| Normalize(mean, std) | 实例化归一化器： | | |
| Resize([height, width]) | 实例化尺寸转换器 | | |
| RandomApply(transforms, p) | 实例化随机应用器 | | |
| RandomHorizontalFlip(proba) | 实例化随机镜像器 | 水平方向 | proba：概率 |
| RandomVerticalFlip(proba) | 竖直方向 |
| RandomRotation([min, max], fill\_color) | 实例化随机旋转器 | | |
| RandomCrop(shape) | 实例化随机裁剪器 | | |
| RandomAffine(rotate, translate, scale, fill\_color) | 实例化随机仿射变换器 | | |
| ColorJitter(brightness, contrast, saturation, hue) | 实例化色彩变化器 | | |
| Grayscale(num\_out\_channels) | 实例化灰度转换器 | | |
| GaussianBlur(kernel\_size, sigma) | 实例化高斯模糊器 | | |

拓展：

diff-gaussian-rasterization：

<https://github.com/graphdeco-inria/diff-gaussian-rasterization/tree/3dgs_accel>

|  |  |  |
| --- | --- | --- |
| **GaussianRasterizationSettings** | | 渲染配置 **(结构体)** |
| **实例属性** | **tanfovx / tanfovy** | **tan(0.5 \* FoV)** |
| **campos** | **相机位置** |
| **debug** | **???** |
| **pixel\_weights** | **像素权重** |

|  |  |  |
| --- | --- | --- |
| **GaussianRasterizer(setting)** | | 光栅化器 **(nn.Module)** |
| **实例方法** | **forward(**  **means3D,**  **means2D, # grad：在视角空间中的位移**  **opacities,**  **dc, shs, # C的一阶、其它参数**  **colors\_precomp, # None时由C计算**  **scales, rotations,**  **cov3D\_precomp, # None时由S、R计算**  **) -> Tuple[torch.Tensor] # render, radii, \*arg** | 光栅化 **(radii: 3DGS在图像的投影半径)** |

fused\_ssim：

<https://github.com/rahul-goel/fused-ssim>

|  |  |
| --- | --- |
| fused\_ssim(img0, img1, padding="same", train=True) | 结构相似性 |

mmcv：

|  |  |
| --- | --- |
| to\_2tuple(arg) / to\_ntuple(n)(arg) | 返回参数的n元组 |

|  |  |  |
| --- | --- | --- |
| 图像： | | |
| IO | imread(img\_or\_path, flag) / imfrombytes(content) | 读取图像 |
| imwrite(img, file\_path) | 保存图像 |
| imshow(img, wait\_time) | 展示图像 |
| 变换 | bgr2gray / bgr2hsv(img) | 颜色空间转换 |
| frames2video(frame\_dir, video\_file, fps) | 图像 → 视频 |

|  |  |  |
| --- | --- | --- |
| 光流： | | |
| IO | flowread(flow\_or\_path, quantize, concat\_axis) | 读取光流 |
| flowwrite(flow, filename, quantize, concat\_axis) | 保存光流 |
| flowshow(flow, wait\_time) | 展示光流 |
| 变换 | flow2rgb(flow) | 颜色空间转换 |

|  |  |  |  |
| --- | --- | --- | --- |
| 视频： | | | |
| 编辑 | **cut\_video(in\_file, out\_file, start, end, vcodec='h264')** | | **裁剪** |
| **resize\_video(in\_file, out\_file, size, ratio, keep\_ar)** | | **尺寸**变换 |
| **concat\_video(video\_list, out\_file, vcodec)** | | **拼接** |
| IO | VideoReader(filename) | | 实例化视频 (可切片) |
| 实例属性 | width / height / resolution | 分辨率 |
| fps | FPS |
| 实例方法 | read() | 读取图像 |

|  |  |
| --- | --- |
| 变换： | |
| imresize(img, size, return\_scale) /  imresize\_like(img, dst\_img, return\_scale) | 尺寸硬变换 |
| imresize\_to\_multiple(img, divisor, size, keep\_ratio, return\_scale) /  impad\_to\_multiple(img, divisor, pad\_val) | 尺寸软变换 |
| imflip / imflip\_(img, direction) | 图像翻转 |
| imrotate(img, angle, center, scale, border\_value) | 图像旋转 |
| imcrop(img, bboxes, scale) | 图像裁剪 |

|  |  |
| --- | --- |
| 边界框： | |
| bbox\_scaling(bboxes, scale, clip\_shape) | 边界框缩放 |
| imshow\_bboxes(img, bboxes, thickness, show, wait\_time, out\_file) | 绘制无标签边界框 |
| imshow\_det\_bboxes(img, bboxes, labels, class\_names, score\_thr,  thickness, show, wait\_time, out\_file) | 绘制有标签边界框 |

|  |  |  |
| --- | --- | --- |
| 注册表： | | |
| Registry(name) | | 实例化注册表 |
| 实例方法 | register\_module(name, force) | 返回注册器 |
| get(key) | 返回注册表记录 |

实例可使用字典表示，type特指类名，其它key表关键字参数

|  |  |  |
| --- | --- | --- |
| 配置： | | |
| Config(cfg\_dict) | | 实例化配置 |
| 类方法 | fromfile(file) | 文件 → 配置工具 |
| 实例方法 | dump(file) | 配置工具 → 文件 |

部署tensorrt：

|  |  |
| --- | --- |
| onnx2trt(onnx\_model, opt\_shape\_dict, fp16\_mode) | ONNX → TensorRT |

mmengine：

Config = mmcv.Config

评估evaluator：

|  |  |  |
| --- | --- | --- |
| 度量BaseMetric： | | |
| 实例方法  (重写) | process(self, data\_batch, data\_samples) | 处理过程数据，并存储在self.results |
| compute\_metrics(self, results) -> dict | 根据过程数据，计算并返回性能度量 |

模型model：

|  |  |  |
| --- | --- | --- |
| 模型BaseModel： | | |
| 实例方法  (重写) | \_\_init\_\_(self, \*args, \*\*kwargs) | 设置实例属性 |
| forward(\*\*kwargs, mode) | mode='loss'：dict  mode='predict'：any |

执行器runner：

|  |  |  |
| --- | --- | --- |
| 训练循环： | | |
| EpochBasedTrainLoop | | 基于epoch |
| IterBasedTrainLoop | | 基于iter |
| 形参 | max\_epochs / max\_iters: int | 训练总轮次 |
| val\_begin: int | 开始验证的时间 |
| val\_interval: int | 验证轮次的间隔 |

|  |  |  |
| --- | --- | --- |
| 执行器Runner： | | |
| 形参 | model: nn.Module | 模型对象 |
| workdir: str | 工作目录，保存参数、日志 |
| load\_from: str | 检查点文件 |
| train\_dataloader /  val\_dataloader: Dataloader | 数据集加载器 |
| train\_cfg: dict | 'by\_epoch' 指定训练TrainLoop |
| val\_evaluator: Evaluator / dict / list | 验证评价指标 |
| optim\_wrapper: dict | 'optimizer' 配置迭代器 |
| param\_scheduler: dict | 学习率步进包装 |
| auto\_scale\_lr: dict | 学习率缩放  {'base\_batch\_size': None, 'enable': True} |
| resume: bool | 还原训练进程 |
| 类方法 | from\_cfg(cfg) | 通过配置文件创建 |
| 实例方法 | train() -> nn.Module | 启动训练进程 |
| val() -> dict | 启动验证进程，返回度量 |

钩子hooks：

|  |  |
| --- | --- |
| EMAHook(begin\_iter, begin\_epoch) | 实例化EMA模型 |

mmdet：

评估evaluation：

|  |  |
| --- | --- |
| eval\_map(preds: [b, nc, n, 5], gts, scale\_ranges, iou\_thr) | 计算逐尺度mAP  AP：P - score 曲线下的面积  mAP：按类别平均的AP |
| eval\_recalls(gts, preds, iou\_thrs) | 计算逐置信度、逐IoU阈值Recall |

模型models：

|  |  |
| --- | --- |
| multiclass\_nms(multi\_bboxes, multi\_scores, score\_thr, nms\_cfg, max\_num) | 多类别边界框NMS |

mmgen：

|  |  |
| --- | --- |
| init\_model(config, checkpoint, device) | 建立模型 |
| sample\_unconditional\_model(model, num\_samples, num\_batches) | 无条件采样 |

mmpretrain：

|  |  |  |
| --- | --- | --- |
| ModelHub | | 模型仓库 |
| 类属性 | \_model\_dict | 所有模型信息 (需先调用get) |
| 类方法 | get(model\_name) | 获取模型信息 |

|  |  |
| --- | --- |
| get\_model(model, pretrained, device) | 建立模型 |
| inference\_model(model, \*args, \*\*kwargs) | 模型自适应推理 |