

2.1.2 Classification

利用提取的特征图计算类别，YOLOv5 的检测结果通过 3 个不同的尺度，分别来检测小、中、大目标。通过 GloU 算法：先计算两个框的最小闭包区域面积（通俗理解：同时包含了预测框和真实框的最小框的面积），计算出 IoU，再计算闭包区域中不属于两个框的区域占闭包区域的比重，最后用 IoU 减去这个比重得到 GloU，以及 NMS 算法得到最终预测框，计算损失值，每个 EPOCH 结束后，将验证集传入网络，验证模型的可靠性，计算损失值、准确率。

2.2 数据采集方式

当前试点的区域为管道连接处的珐琅区域，所以在数据采集上主要集中于该区域采集符合清晰、斜角小于 30 度的视频数据用于数据训练。数据采集上：

(1) 正样本的采集：由于当前试点的油库很难采集到滴漏或者喷油情况下的数据，在正样本采集上，采用模拟滴漏、喷油的方法，在管道连接处进行模拟作业，仿制场景相似的数据用于数据训练，对数据规模上要求不低于 2000 张有效不重复的正样本数据（平均单一场景大致一个厂区 2000 张左右），样本数量应当分布均匀的采集各个不同管道连接处珐琅的模拟滴漏、喷油情况的数据。

(2) 负样本的采集：在各个不同管道连接处珐琅处采集不同时段、不同天气、不同角度、不同距离的正常状态下的连接处数据，用于训练识别管道连接处检测算法以及正常状态下的管道连接处。

标注内容包括两个目标：漏液区域和漏液发生在珐琅区域，需对珐琅进行目标检测标注。



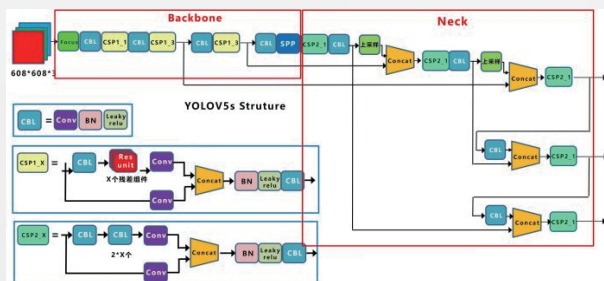
2.3 算法训练策略

2.3.1 Mosaic 数据增强策略

在数据喂给神经网络之前，打乱数据集顺序，再采用 Mosaic 数据增强策略，随机缩放、随机裁减、随机排布的方式进行拼接等，丰富了检测物体的背景和小目标，并且在计算 Batch Normalization 的时候一次会计算 4 张图片的数据，使得 mini-batch 大小不需要很大。然后对图像进行归一化处理，采用自适应图片缩放算法，将图片长边缩放到统一尺寸 416 像素。其中，短边以相同的比例缩放，不足 416 的部分对 64 取余数，然后用 0 对图像短边填充余数个长度，并且同时对标签位置信息做相同的变换，再送进网络。这样训练时，每个 epoch 接受的数据都不相同，可以增加模型的鲁棒性。

2.3.2 CSP 结构

Yolov5 中设计了两种 CSP 结构，以 Yolov5s 网络为例，CSP1_X 结构应用于 Backbone 主干网络，另一种 CSP2_X 结构则应用于 Neck 中。



CSPNet(Cross Stage Partial Network): 跨阶段局部网络，CSPNet 和 PRN 都是一个思想，将 feature map 拆成两个部分，一部分进行卷积操作，另一部分和上一部分卷积操作的结果进行 concat。增强了 CNN 的学习能力，能够在轻量化的同时保持准确性。降低计算瓶颈和内存成本。CSPNet 通过将梯度的变化从头到尾地集成到特征图中，在减少了计算量的同时可以保证准确率。

2.3.3 Neck

Yolov5 的 Neck 部分采用了 PANet 结构，Neck 主要用于生成特征金字塔。特征金字塔会增强模型对于不同缩放尺度对象的检测，从而能够识别不同大小和尺度的同一个物体。