**反无人机大作业汇报文档**

**汇报人：杨超然**

**学号：1222906085**

1. **选题**

近年来关于无人机的话题非常热门。无人机具有体积小、成本低、机动性强等优势，而优势也可能造成危害。近来国内外多次发生无人机非法入侵事件，不仅对公民的个人隐私与生命财产安全造成了严重危害，而且对机场、军事基地、大型集会现场、核电站、政府机部门驻地等敏感区域的安防造成了极大威胁。反无人机算法可以自动、持续地保护重要区域的安全，探测和监管无人机攻击。因此我们决定研究反无人机算法。

反无人机指在野外发现、探测、识别和跟踪无人机（UAV）目标，同时根据RGB或热红外（IR）视频估计目标的跟踪状态。接着寻找反无人机的研究现状作为参考，北方电子设备研究所在研究反无人机项目，旨在推动无人机在野外发现、检测和跟踪的前沿发展。我们决定参考其项目进行研究。

1. **算法**

在算法方面，我们选择在图像识别方面著名的YOLO模型进行尝试。同时我们也在反无人机相关领域寻找使用YOLO模型的论文作为参考。最终在多角度的考虑之下我们选择了使用YOLOv11模型。

相比YOLO的其他版本，YOLOv11模型有具有一些特别的优势：对比Yolov5，Yolov8的Backbone将C3模块替换成了C2f模块，实现了进一步轻量化，同时沿用Yolov5中的SPPF模块。Head部分将耦合头换成了目前主流的解耦头结构，将检测和分类头分离。对比Yolov8，Yolov11中的Backbone将C2f模块变为C3K2模块。Backbone中的最后一层（SPPF层）后增加了C2PSA模块。Head解耦头中的分类检测头增加了DWConv。

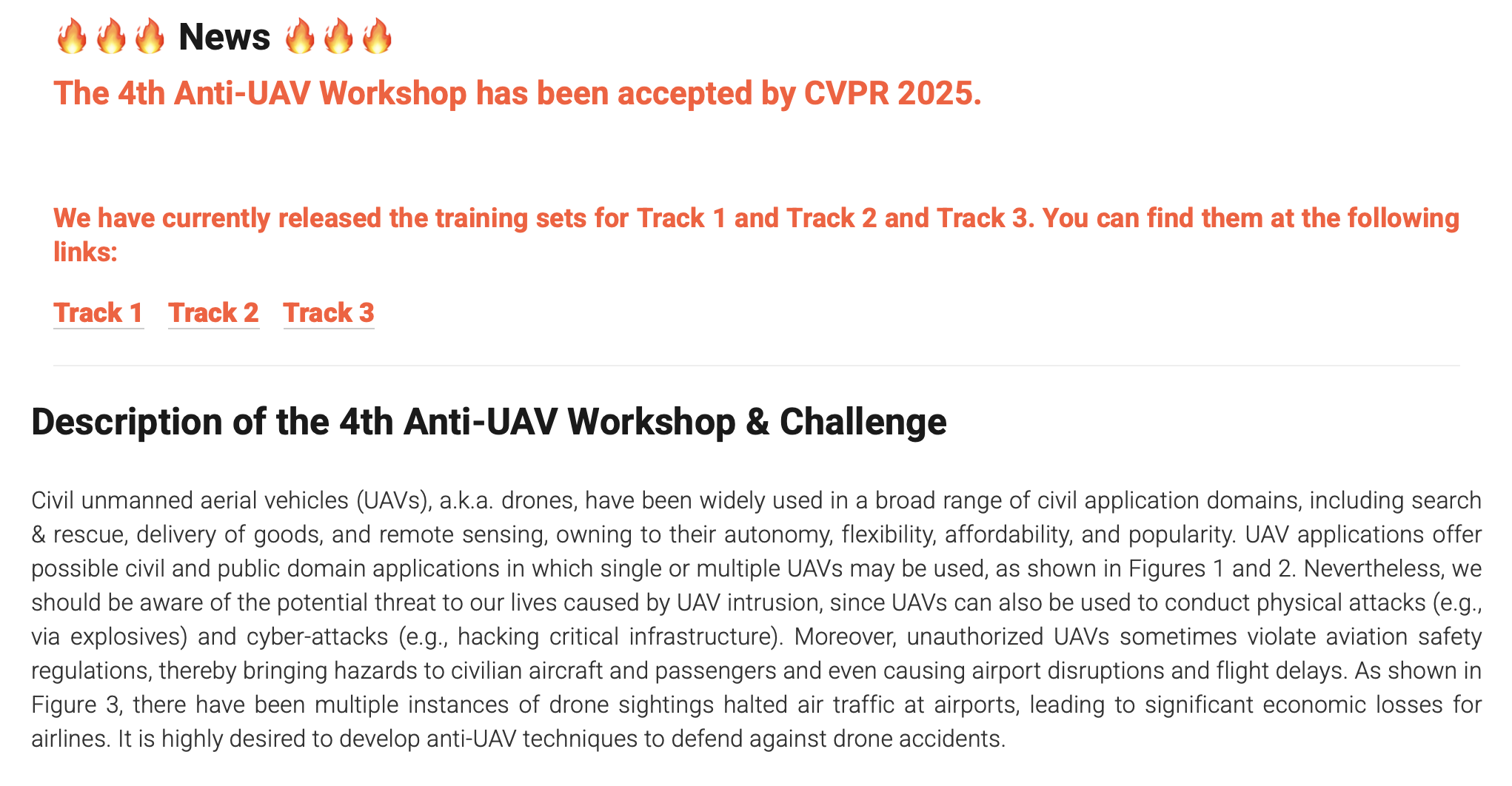
C3模块即上述CSP结构。C3k允许设置不同的卷积核大小，以便更灵活地应对各种不同的特征提取需求。C2f通过减少卷积层的数量和采用更高效的特征合并策略来提高速度。C3K2结合了C2f的速度优势和C3k的灵活性，当C3k参数设置为True时使用C3k层，否则使用标准的瓶颈层，与 C2f 类似。

C2PSA用于增强特征提取，结合了CSP结构和PSA注意力机制。特征图被分为两部分，一部分直接传递，另一部分通过PSA注意力模块处理，最终拼接融合。PSA模块通过引入不同大小的卷积核来提取多尺度的空间信息，同时结合 Squeeze-and-Excitation 模块对特征通道进行加权，从而增强网络对不同尺度目标的注意力聚焦。

DWConv（深度卷积）是一种高效卷积操作，主要用于减少计算复杂度和参数量。在标准卷积操作中，对于一个输入张量，卷积核与输入张量在每个位置都会计算与所有输入通道的点积，每个输出通道是所有输入通道的加权求和，导致计算量比较大。深度卷积将输入的每个通道单独处理，即每个通道都有自己的卷积核进行卷积，不与其他通道进行交互。

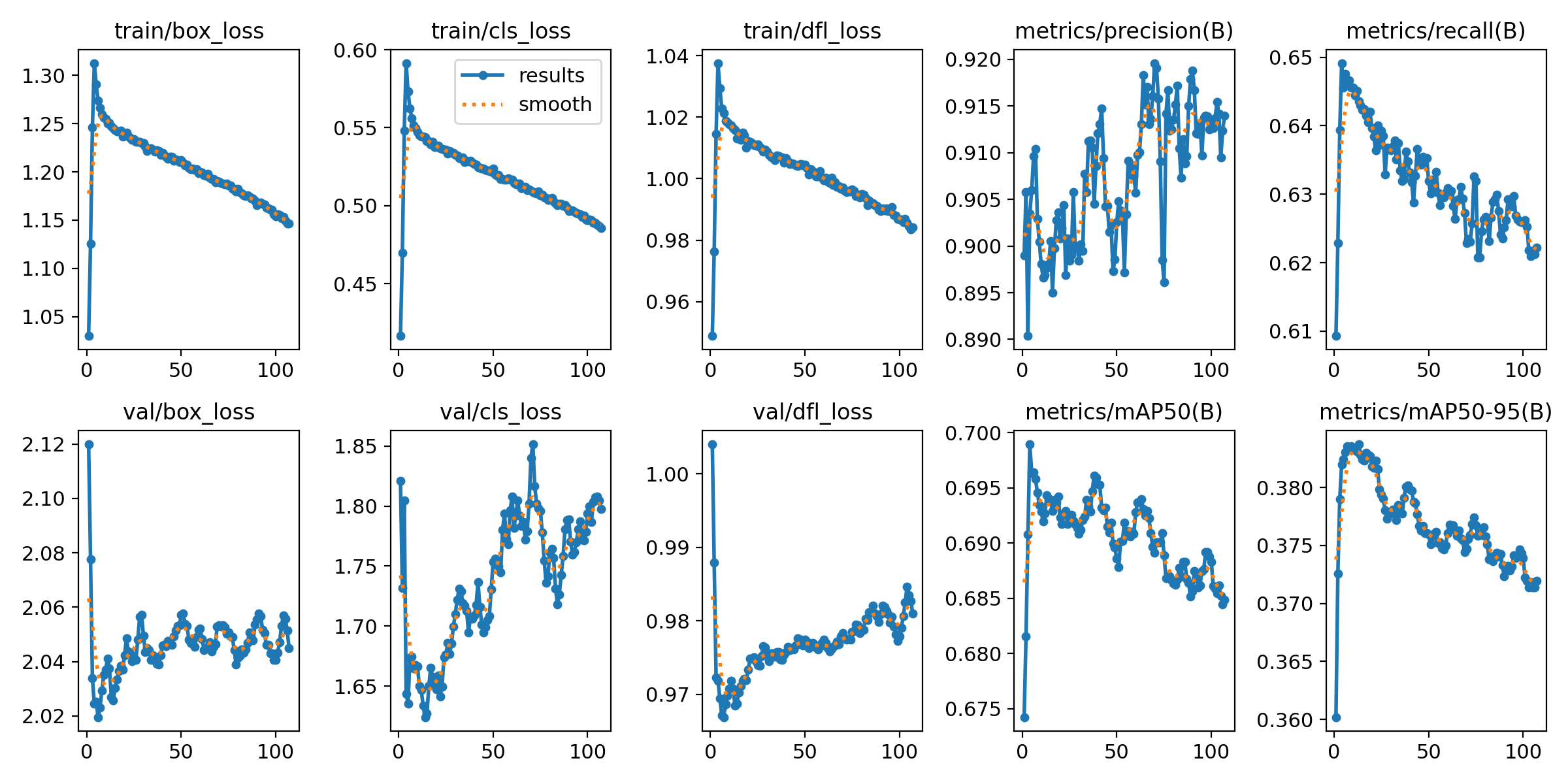
1. **数据集与部署**

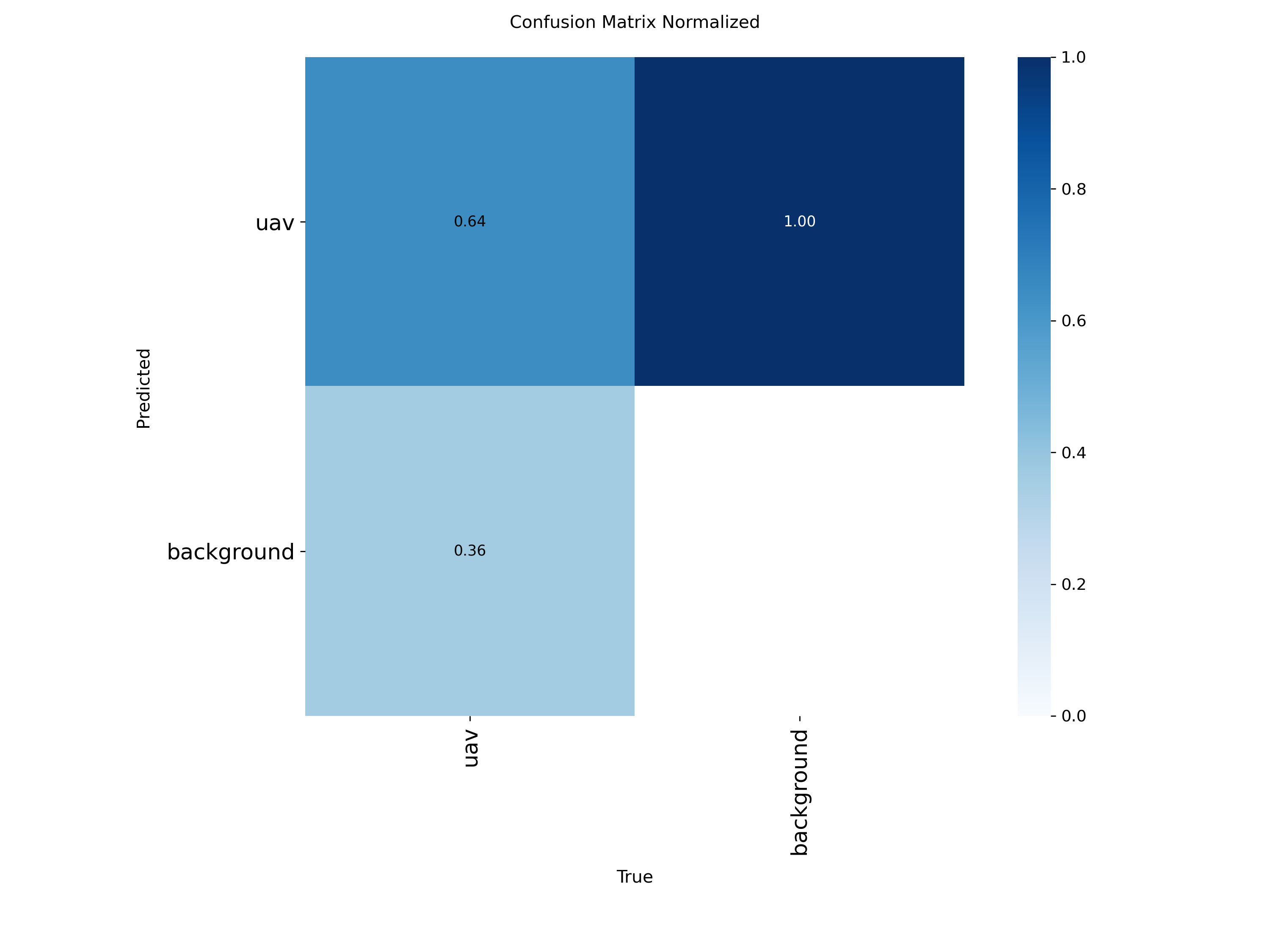
我们选择了一个无人机挑战赛项目的公开数据集，作为训练用数据集。其中已经配有预分割的验证集、约20000条训练样本、且格式为RGBD图像。我们使用conda环境下的modelscope开源平台下载了该数据集，但发现数据集本身仍存在以下问题：文件夹内层数超过一层；数据集格式不是coco格式的JSON脚本，缺少了长宽信息；本身并不是yolo格式，需要转换成yolo格式。据此，我们编写了3个python脚本以解决上述问题：addwh.py(用于补全长宽信息)、coco2yolo.py(用于将coco格式转换为yolo格式)、flatten\_img.py（将多层嵌套的文件平铺到一个单一目录中）。



经过python脚本的处理后，我们终于获得了一个可以用于训练的数据集。接下来考虑部署的问题。我们所使用的设备就是一台普通的装载3070ti的笔记本电脑，按照教程安装nvidia docker utils以及docker pull yolo11的最新镜像。因受学校断电限制，我们选择使用内网穿透，远程控制放在家中的笔记本进行训练。

1. **成果分析**





从图表中可以看出训练损失随着训练轮次的增加逐渐减小，精度和召回率随着训练轮次的增加不断提高，说明模型的准确性和完整性都在增强；同时在验证集上，这些损失同样呈现出逐渐缩小的趋势，说明模型在没有见过的数据集上也有较为一定的泛化能力。

但从混淆矩阵中的数据可以看到，模型的识别准确率仅为64%，对背景场景的识别正确率为100%，仍然还有很大的改进空间。