

2. 答辩演讲

前两页 PPT-首页+目录

- **第一页**：各位老师，大家早上好。我是计科专业的薄劲阳。今天我的答辩题目是轻量化目标检测系统的设计与实现。
- **第二页**：今天我主要从以下五个方面开展答辩。

第三页 PPT-研究背景及意义

- 首先说一下研究背景及意义，目前目标检测是计算机视觉中一个相当重要的方向。传统的目标检测算法通常要求较大的模型大小和高计算复杂度，这会对计算和存储资源造成较大的负担。轻量化目标检测算法通过优化网络结构、减少参数数量以及采用高效的计算方式，可以大幅降低算法对资源的需求。

第四页-第七页 PPT-YOLO 基本原理及网络架构

- **第四页**：YOLO 的基本原理如右图所示，主要可以分为三个阶段。
 - 第一个阶段是将图片进行处理，输入到网络模型中。
 - 第二阶段是网络模型进行推理，得到一个 tensor 张量。这个张量包含了物体的可能位置以及各个类别的概率。
 - 第三阶段是对第二阶段的 tensor 进行非极大值 NMS 处理，去除多余的候选框，最终确定图像中物体的具体位置和类型。
- 我的毕设实验，主要是针对 YOLOv5、YOLOv5-lite、YOLOv7 三个模型进行实验的。YOLO 目标检测模型的网络结构，主要分为 Backbone、Neck、Head

结构。Backbone 结构是网络的核心部分，主要负责从输入图像中提取关键特征。Neck 结构主要负责输入特征的融合以及加强。Head 结构负责去除多余的预测框、物体类别预测。我们主要的研究对象是网络的 Backbone 结构。

- **第五页**：YOLOv5 的 Backbone 结构，由 Focus、CBL 以及 CSP 组成。Focus 模块负责对图像进行 Slice 切片操作，将高维图像转换为低维图像（会导致增加通道数）。CBL 模块是由 Conv 卷积模块、BatchNorm 归一化处理模块、Leaky ReLU 激活函数模块组成。CSP 由 CBL、Res Unit、Conv 构成，主要负责 XXX。
- **第六页**：YOLOv5-lite 的 Backbone 结构由 CBL、ShuffleNet 结构构成。相比于 YOLOv5，去除了 Focus 层。另外在提取特征模块，使用 ShuffleNet 替代 CSP 模块。相比于 CSP 模块，ShuffleNet 减少了缓存的使用。
- **第七页**：YOLOv7 的 Backbone 结构由 CBS、ELAN、MP1 模块组成。CBS 模块由 Conv 卷积层、BatchNorm 归一化层、SiLU 激活函数。ELAN 则是由 5 个连续的 CBS 模块组成。MP1 是负责最大池化的模块。模块中存在大量并行的网络结构。

第八-第十三页 PPT-模型对比实验及结论分析

- **第八页**：我主要将训练好的模型分别在 PC 端和树莓派端运行，比较不同模型的轻量化效果。PC 和树莓派的配置环境如下表所示。然后训练模型采用的是 Fruits detection 数据集，他总共有五种类别。测试数据选择了一段时长为 58s 的视频，用来进行模型的比对实验。
- **第九页**：模型轻量化的效果，我主要使用了 FPS、参数量以及模型占用内存三个指标反应。FPS 代表了模型一秒内能够推理完成的图像张数。其从图片

处理处开始计时，到推理结束停止。具体计算公式如下所示。参数量反应了网络结构的复杂程度。参数量越多，代表模型越复杂。模型占用内存则反应了模型所占用的存储空间大小。

- **第十页**: 具体的实验结果如下表所示。FPS 层面, YOLOv5、YOLOv5-Lite、YOLOv7 在 PC 端的运行效果要远高于在树莓派的运行效果。在 PC 端, YOLOv5-lite 的运行效果最好, YOLOv5 次之, YOLOv7 最差。在树莓派端, YOLOv5-lite 的效果最好, YOLOv5 与 YOLOv7 相近, 都很差。
- **第十一页**: 从参数量和模型占用内存大小来看, YOLOv7 的参数量和模型占用内存最大, 这也说明了 YOLOv7 的网络最复杂。而 YOLOv5-lite 则在这两方面都是最小的, 也正确反映了为什么其在树莓派上的效果好的原因。
- **第十二页**: 针对实验结论, 我阅读了 Shufflenet v2 论文。这篇论文指出了轻量化的四个原则。其从内存访问代价和 GPU 并行性方向分析了网络应该如何设计才能减少存储器开销。概括下来就是要减少存储器的使用, 增加网络的并行度。具体四个准则如下所示。
- 第十三页: 读 PPT。

第十四页-第十五页-成果展示

- 第十四页: 另外我对模型搭建了一个简单的 QT 平台。介绍 PPT
- 第十五页: 读 PPT
- 第十六页: 另外, 我目前还存在一些小问题。读 PPT