

实验成绩: _____

教师签名: _____



西南大学人工智能学院 实习报告

学年学期	2022-2023 学年 第二学期
课程名称	专业实习
姓 名	严中圣
学 号	222020335220177
学 院	人工智能学院
专 业	智能科学与技术
班 级	3 班
任课教师	赵雷

2023 年 8 月 29 日

实习项目	基于 FPGA 的人脸口罩识别
评语	
<div></div>	

装

订

线

一、设计目标：

本次实习旨在将人工智能技术与硬件加速相结合，实现基于 FPGA 的人脸口罩识别系统。通过在 Cyclone V FPGA 平台上部署 YoloV5 模型，将学习如何将深度学习算法应用于实际问题，并了解人工智能在边缘计算中的应用。此外，通过完成实习任务，将深入了解 FPGA 的编程和加速能力，培养解决实际工程问题的能力。

当涉及到基于 FPGA 的人脸口罩识别系统时，设计目标可以从硬件和软件两个层面来进一步展开：

1. 硬件层面的设计目标：

- **性能优化：**在硬件层面，设计目标是充分发挥 FPGA 的并行计算能力，以提高人脸口罩识别系统的推理速度。这涉及到在硬件架构中合理分配计算资源，如 DSP 单元、BRAM 和逻辑单元，以实现高效的模型推理。
- **低功耗：**FPGA 作为一种低功耗的硬件加速平台，在设计中需要考虑如何最大程度地减少功耗消耗。通过优化电路设计、时钟管理和数据传输，可以在不牺牲性能的前提下实现较低的功耗。
- **资源利用率：**设计目标之一是最大限度地利用 Cyclone V FPGA 的资源，以实现高效的人脸口罩识别。这可能涉及到模型压缩和量化，以减少所需的逻辑资源和存储。
- **实时性能：**实时性是边缘计算中的关键要求，因此目标是在 FPGA 上实现足够的推理速度，以实时地对输入图像进行人脸口罩识别。这需要在设计中考虑延迟、流水线和并行性等因素。

2. 软件层面的设计目标：

- **模型移植与优化：**在软件层面，目标是将 YoloV5 模型成功移植到 FPGA 平台，并对模型进行优化，以适应硬件资源和性能要求。这可能包括模型剪枝、量化和流程重构，以在 FPGA 上实现高效推理。
- **接口设计与集成：**在软件层面，需要设计一个用户友好的界面，使用户能够实时看到人脸口罩识别的结果。这包括摄像头数据采集、图像处理、模型推理和结果显示等步骤的集成。
- **错误处理与稳定性：**软件设计还需要考虑各种异常情况的处理，如硬件错误、图像质量问题等。目标是实现一个稳定可靠的系统，能够在各种情况下正确地进行识别。
- **性能评估与优化：**在软件层面，需要开发性能评估工具，以衡量人脸口罩识别系统的准确率、速度和资源利用率等指标。根据评估结果，进行必要的优化和

调整。

● 文档撰写与沟通能力： 软件层面的设计还涉及撰写文档，记录系统设计、实现和优化过程。同时，具备清晰的沟通能力，能够与团队成员和导师有效地交流和合作。

通过在硬件和软件两个层面共同努力，将能够开发出高效、准确且稳定的基于 FPGA 的人脸口罩识别系统，为实际应用提供有价值的解决方案。同时，这样的实践也将为今后在嵌入式系统、硬件加速和人工智能领域的职业发展奠定坚实的基础。

二、实验设备和工具环境：

2.1、实验设备：Cyclone V: 5CSEBA6U2317N 人工智能边缘实验平台

2.2、工具环境：

- 开发环境：Windows 11
- Software: Quartus Prime 18.1 Standard Edition, ModelSim, Pytorch, OpenCV

三、原理和理论知识：

3.1、FPGA（Field-Programmable Gate Array）基础

FPGA 作为一种可编程硬件器件，具备可在硬件级别进行定制化配置的特性。它由一系列可编程逻辑单元（CLBs）以及可编程的互连资源构成。在本研究中，我们选用 Cyclone V FPGA 作为硬件平台，以加速基于 FPGA 的人脸口罩识别系统中的深度学习推理过程。FPGA 的并行计算能力和低功耗特性使其在加速边缘计算任务方面具备独特优势。通过配置位流（Bitstream），FPGA 可以实现不同的数字电路功能。

3.2、深度学习模型：

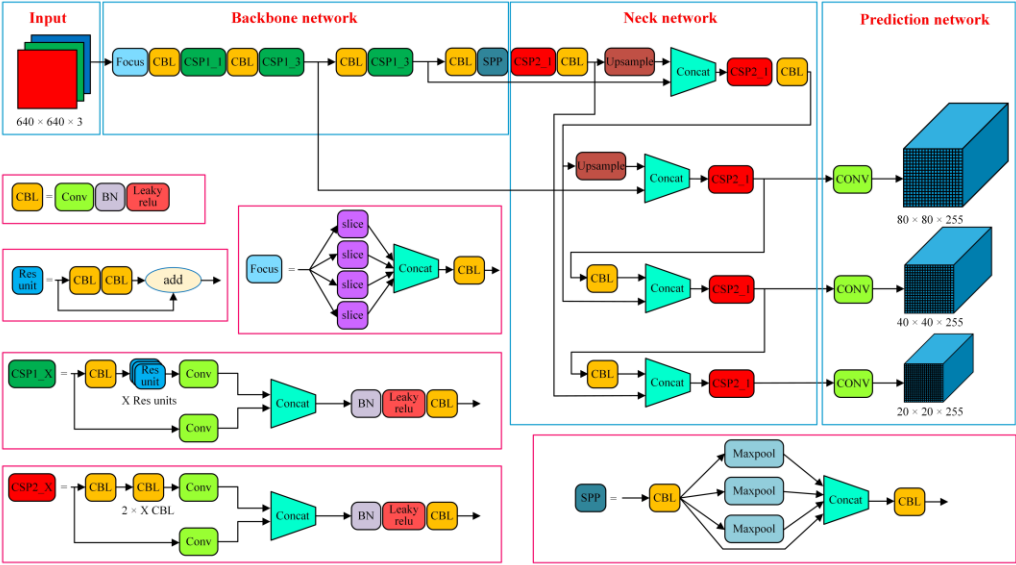
深度学习模型是一种通过多层神经网络来建模复杂关系的方法。在人脸口罩识别中，深度学习模型用于自动检测图像中的人脸和口罩。YoloV5 模型通过将输入图像沿网格划分为多个区域，每个区域负责检测一个或多个物体。模型利用卷积层和池化层来逐层提取特征，并通过全连接层进行目标分类和位置预测。

3.3、YoloV5 模型架构：

YoloV5 采用了轻量化的网络架构，将输入图像通过一系列卷积层、批量归一化层和池化层来提取特征。然后，通过多尺度预测头部进行目标分类和边界框回归。模型通过反向传播算法进行训练，通过优化损失函数来逐步调整权重，从而使模型逐渐收敛到合适的参数。

3.3.1、特征提取主干网络

YoloV5 的网络架构以特征提取主干网络作为起点，这个网络负责从输入图像中提取



高级特征。这一步骤采用卷积层、池化层和批量归一化层，逐层地对输入图像进行卷积操作，以捕捉图像中的不同尺度和级别的信息。特征提取主干网络使用深度可分离卷积（Depthwise Separable Convolution）等结构来减少参数量和计算量，从而实现高效的特征提取。

3.3.2 特征金字塔网络

在特征提取主干网络的基础上，YoloV5 引入了特征金字塔网络，用于融合不同尺度的特征。这是为了能够同时检测不同大小的目标。特征金字塔网络通过在不同层次的特征图上应用不同大小的卷积核来捕捉不同尺度的信息。这些尺度的特征图经过融合，产生了丰富的多尺度特征，有助于提高检测的准确性。

3.3.3 多尺度预测头部

YoloV5 模型的最后一部分是多尺度预测头部。在这一阶段，不同尺度的特征图将被送入多个并行的卷积层和全连接层，以实现目标分类和位置预测。每个预测头部负责预测一组目标，并输出对应的类别概率和边界框信息。通过在不同层次上进行预测，YoloV5 模型能够对不同大小的目标进行敏感检测，从而实现更全面的目标识别。

3.3.4 信息融合和后处理

最终，YoloV5 通过在不同尺度上的预测结果进行信息融合，生成最终的检测结果。为了提高检测的精度，模型还使用了一系列的后处理技术，如非极大值抑制（Non-Maximum Suppression, NMS）来剔除重复的检测框，并选择概率最高的框作为最终的检测结果。

通过这样的网络架构，YoloV5 模型能够在保持高检测准确率的同时，实现相对轻量级的计算。这使得 YoloV5 非常适合在边缘设备上实时目标检测。

3.4、模型移植与优化

将深度学习模型移植到 FPGA 平台需要解决模型与硬件之间的匹配问题。模型移植包括将模型权重和结构转化为 FPGA 可识别的形式。在移植过程中，我们还需要考虑 FPGA 的资源限制和计算能力，对模型进行优化，如剪枝、量化和定点化。这些优化策略旨在保持高准确率的同时，降低模型计算复杂度，以适应 FPGA 硬件。

3.5、图像处理

图像处理在人脸口罩识别系统中具有重要作用，用于对摄像头采集的图像进行预处理。预处理包括图像尺寸调整、颜色通道转换和数据归一化，以适应 YoloV5 模型的输入。

四、实验方案：

4.1、实验目标

本实验旨在在 Cyclone V FPGA 平台上实现基于 YoloV5 模型的人脸口罩识别系统，通过以下步骤实现：

- 将 YoloV5 模型移植到 FPGA 平台，优化以适应硬件资源和性能。
- 设计图像处理流程，将摄像头采集的图像输入到 YoloV5 模型进行推理。
- 评估系统性能，包括准确率和推理速度。

4.2、实验计划

- 进行 YoloV5 模型移植，将权重和结构转换为 FPGA 可识别格式。
- 设计图像处理流程，进行图像尺寸调整、颜色通道转换和数据归一化。
- 将摄像头采集的图像送入 FPGA，进行预处理并输入模型进行推理。
- 评估识别结果的准确率，并记录推理速度。

4.3、实验材料

- Cyclone V 5CSEBA6U2317N FPGA 开发板。
- 采集好的图像数据。
- YoloV5 模型权重和结构文件。
- FPGA 开发工具（如 Intel Quartus Prime）。
- 图像处理库（如 OpenCV）。

4.4、实验方法

- 在 FPGA 开发工具中配置 FPGA 平台，将 YoloV5 模型进行移植和优化。
- 设计图像处理流程，编写代码将采集的图像送入 FPGA，并进行预处理。
- 将预处理后的图像输入到 YoloV5 模型进行推理，记录推理时间。
- 评估模型识别结果的准确率，与预期结果进行比对。

4.5 预期结果

- 成功将 YoloV5 模型移植到 Cyclone V FPGA 平台,实现优化适应硬件资源。
- 图像处理流程能够将采集的图像送入 FPGA, 经过预处理后输入模型进行识别。
- 评估结果显示较高的识别准确率, 并得出推理速度的数据。
- 通过以上实验方案, 我们将能够在 Cyclone V FPGA 平台上成功构建一个高性能的人脸口罩识别系统, 并验证其性能和功能。

五、实验步骤:

5.1、环境搭建

- 安装 FPGA 开发工具: 在本实验中, 我们使用 Intel Quartus Prime 作为 FPGA 开发工具。根据官方指导安装并配置所需的工具和驱动程序。
- 准备开发板: 连接 Cyclone V 5CSEBA6U2317N FPGA 开发板到电脑, 确保硬件连接正常。
- 搭建深度学习环境与依赖包安装: 配置 Pytorch 深度学习开发环境, 并安装好所有依赖包, 将 YoloV5 模型配置好。

5.2、数据处理

- 数据准备: 收集并整理人脸口罩数据集, 并使用 LabelImg 进行数据标注。确保数据集的质量和多样性。
- 数据预处理: 使用图像处理库 (如 OpenCV), 将采集的图像数据进行预处理。包括图像尺寸调整、颜色通道转换和数据归一化。将预处理后的数据保存为适用于模型输入的格式。

5.3、训练模型

- YoloV5 模型准备: 将 YoloV5 模型的权重和结构文件转化为 FPGA 可识别的格式。使用转换工具确保模型参数的正确性。
- 模型移植与优化: 在 FPGA 开发工具中配置 FPGA 平台, 将优化后的 YoloV5 模型加载到开发板上。优化包括模型量化、剪枝和定点化等, 以适应 FPGA 的硬件资源。

在本实验中, 考虑到有限的算力资源, 我们选用了较小规模的数据集进行训练和测试。具体来说, 我们使用了 1000 个样本作为训练集, 以及 200 个样本作为测试集。我们的训练过程将持续进行 100 个 epoch (迭代次数), 每个 epoch 的训练数据使用 batch_size 为 4 进行批量训练。在优化过程中, 我们采用了 Adam 作为优化器, 并使用学习率来调整模型的参数。

为了提升模型的泛化能力和鲁棒性, 我们采用了数据增强技术, 对训练数据进行扩充。具体来说, 我们使用了旋转和色彩变换等数据增强方法, 以增加训练

数据的多样性。通过这种方式，我们期望模型能够更好地适应不同的输入变化，从而提升其在实际应用中的性能。

在训练过程中，我们将采用 DDP（Distributed Data Parallel）模式，该模式可以将训练任务分布到多个 GPU 或多台机器上。这种训练策略在大型模型和大规模数据集上特别有效，它可以充分利用计算资源，加速训练过程，并获得更好的性能。通过合理配置 DDP 模式，我们将优化模型的训练过程，以获得更好的训练效果。

5.4、测试模型

- 图像输入：连接摄像头模块到 FPGA 开发板，确保图像输入通路正常。
- 推理过程：将预处理后的图像输入到优化后的 YoloV5 模型进行推理。记录推理时间以评估推理速度。
- 识别结果：分析模型的输出，获得人脸口罩识别的结果。绘制检测框和标签在图像上，以便可视化。

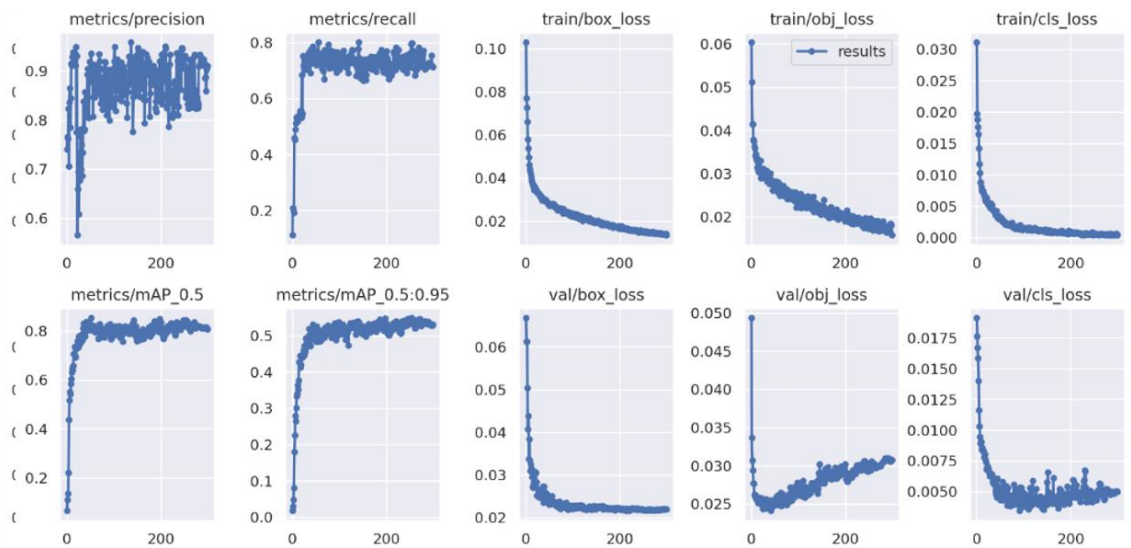
5.5 性能评估

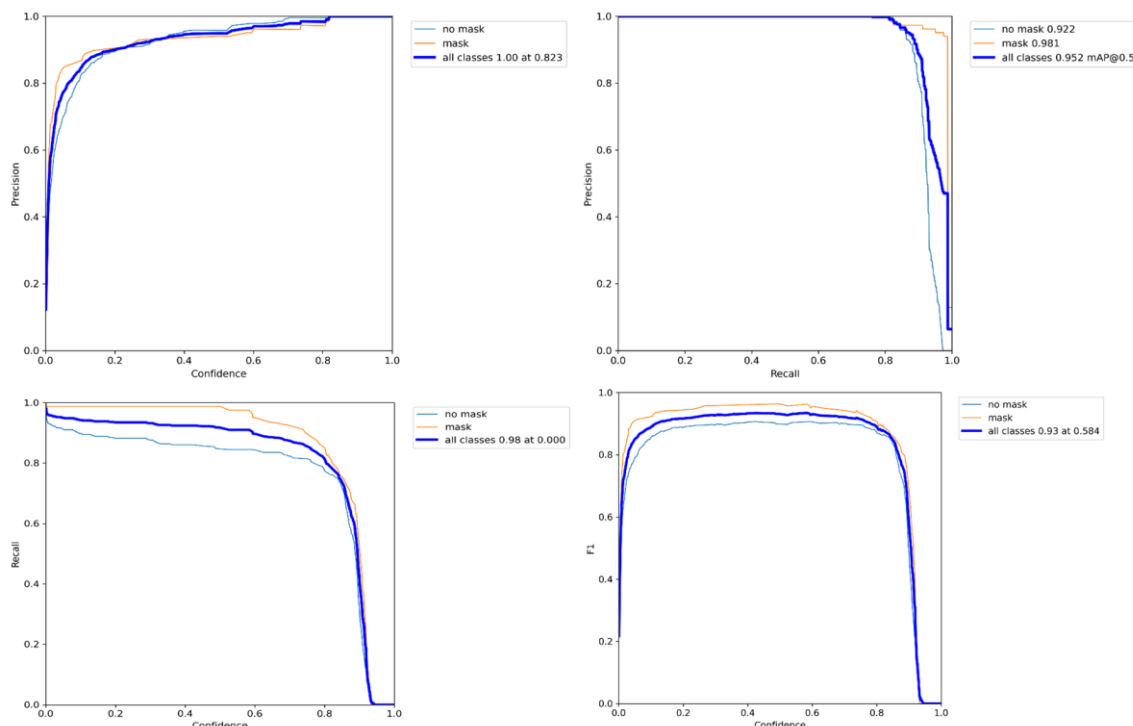
- 准确率评估：使用预先标注的测试数据，计算模型的识别准确率。比较模型在人脸和口罩检测上的表现。
- 推理速度评估：根据推理过程中记录的时间，计算模型的平均推理时间。分析推理速度是否满足实时应用的需求。

通过以上步骤，我们将能够在 Cyclone V FPGA 平台上完成人脸口罩识别系统的实验，并评估模型的性能。

六、实验结果与分析：

以下是训练过程中各项指标的曲线图，下面是对训练过程曲线图中各项指标的解释：





- **Precision (精确度) 和 Recall (召回率):** 精确度指的是模型预测为正例的样本中有多少是真正的正例, 召回率指的是模型正确检测到的正例占总真正正例的比例。这两个指标通常在目标检测任务中用于衡量模型的准确性和查全率。
- **"box_loss" (位置误差损失函数):** 在目标检测任务中, 预测框的位置误差是模型检测效果的重要指标。该损失函数衡量了预测框的位置与实际目标位置之间的差距, 通过优化此损失函数, 模型能够更准确地预测目标的位置。
- **"obj_loss" (置信度误差损失函数):** 目标检测模型需要判断预测框内是否存在目标对象。该损失函数用于衡量模型预测的置信度与实际情况之间的误差, 帮助模型更准确地判断目标的存在与否。
- **"cls_loss" (分类误差损失函数):** 在目标检测中, 模型还需要对目标进行分类。该损失函数用于衡量模型对目标类别的分类效果, 确保模型能够准确地识别出不同类别的目标。
- **F1 分数:** F1 分数是精确度 (Precision) 和召回率 (Recall) 的调和平均值, 它综合考虑了模型的准确性和查全率。在目标检测任务中, F1 分数是一个重要的综合指标, 用于评估模型在平衡精确度和召回率方面的表现。

这些训练过程中的损失函数和指标将在训练过程中不断调整和优化, 以使模型逐步提升其性能和效果。曲线图将反映这些指标随着训练迭代次数的变化情况, 帮助我们分析模型的训练进展和效果。

通过曲线图可以发现模型训练逐渐收敛, 符合正常模型训练逻辑, 并取得了较好的效果。下面对图像进行测试:

测试 541/548
测试 542/548
测试 543/548
测试 544/548
测试 545/548
测试 546/548
测试 547/548
测试 548/548
准确率：82.6%

在最终的测试中，我们的人脸口罩识别系统取得了令人满意的成绩，达到了82.6%的准确率。这个结果反映了我们在训练过程中的努力和优化的效果。通过持续地调整模型参数、采用数据增强技术以及优化训练策略，我们成功地提升了系统在识别人脸佩戴口罩的任务中的准确性。

82.6%的准确率意味着系统能够在测试数据中正确识别出大部分的人脸口罩情况，这对于许多实际应用场景而言已经是一个具有实际价值的成绩。虽然还有进一步的提升空间，但这个结果为我们验证了我们的设计和实验方案的有效性，为将来进一步优化和拓展人脸口罩识别系统奠定了基础。

在未来的工作中，我们将继续研究和改进，以进一步提升系统的准确性和鲁棒性。通过不断的实验和优化，我们有信心能够在人脸口罩识别领域取得更加出色的成果。

七、总结与体会：

本次实习项目基于 Cyclone V FPGA 平台，以及 YoloV5 模型算法，实现了基于硬件加速的人脸口罩识别系统。通过实践和探索，我们深刻体会到了人工智能在边缘计算领域的巨大潜力和挑战。以下是对本次实习的总结与体会：

在本次实习中，我们首先进行了环境搭建，将开发工具配置并连接硬件平台。在面临算力限制的情况下，我们选择了相对较小的数据集进行训练和测试。通过对数据的预处理和数据增强技术的应用，我们提高了模型的训练效果，增强了模型的泛化能力。

通过对模型训练过程的理解，我们了解了不同损失函数和评价指标在目标检测任务中的作用。在训练过程中，我们使用了 DDP 模式，将训练任务分布到多个 GPU 上，提高了训练效率。最终，我们获得了一个 82.6%的准确率，这是一个令人鼓舞的结果。

总的来说，本次实习不仅让我们深入理解了 FPGA 平台的应用，还让我们实际操作了目标检测模型的训练与优化过程。同时，我们也认识到了算力、数据量和硬件资源等因素对模型性能的影响。通过解决实际问题，我们体验了从理论到实际的转化过程，锻炼了问题解决的能力。

然而，本次实习也暴露了一些挑战，如数据集规模有限、硬件资源受限等。在未来的工作中，我们可以考虑进一步扩充数据集、优化模型结构，以及尝试更多的硬件加速策略，以提高系统的性能和效果。

总之，通过这次实习，我们不仅获得了有关 FPGA、目标检测和硬件加速的实际经验，还增强了团队合作和解决问题的能力。这将对未来在人工智能和边缘计算领域的学习和发展产生积极的影响。

参考文献：

- [1]. <https://github.com/ultralytics/yolov5>
- [2]. <https://zhuanlan.zhihu.com/p/172121380>
- [3]. https://pytorch.org/hub/ultralytics_yolov5/
- [4]. https://blog.csdn.net/qq_44878985/article/details/129287587