

社会人类学与中国研究论文

|  |  |
| --- | --- |
| 论文题目： | 基于嵌入式FPGA平台的高能效深度学习框架与应用 |
| 学 院： | 信息科学技术学院 |
| 年 级： | 2012级 |
| 专 业： | 计算机科学与技术 |
| 姓 名： | 赵睿哲 |
| 学 号： | 1200012778 |
| 指导教师： | 梁云 |

年 月 日

# 

# 摘要

随着深度学习相关的理论与应用愈发成熟，计算机视觉研究领域在深度学习的基础上取得了突破性的进展：基于图像识别，分类，检索，以及基于动态视频流的的“即时定位与地图重建”（Simultaneous Localization And Mapping, 简称SLAM）算法等等，都采用深度学习来优化早期效率不高的启发式算法，并获得性能与效果上的长足进步。但是很多计算机视觉算法搭载于汽车、机器人乃至无人机上，这类设备对算法的实时性要求高，但计算资源却又很有限。因此，对深度学习的框架在嵌入式平台，尤其是具有高能效的FPGA平台上的部署与优化工作就成为至关重要的一步。

本研究采用赛灵思（Xilinx）公司推出的“全可编程”SoC平台ZYNQ进行开发，使用最新推出的SDSoC开发套件来优化开发效率。本研究将广泛应用于学术界和工业界的Caffe框架，通过交叉编译与硬件加速关键数学函数等方法部署于ZYNQ。本研究最后通过在Caffe上实现卷积神经网络（Convolutional Neural Network，简称CNN）在立体匹配（Stereo Matching）算法上的应用，来展示系统的高效率与低能耗特性。

关键词

关键词1；关键词2；关键词3

# **Abstract**

Abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract abstract.(英文摘要内容必须与中文摘要完全对应。英文摘要采用Times New Roman小四号字书写，毕业论文、毕业设计行与行之间、段落和层次标题以及各段落之间均为1.5倍行距。)

**Key words**

Key words;key words; key words(英文关键词内容必须与中文关键词完全对应。英文关键词采用Times New Roman小四号字书写，毕业论文、毕业设计行与行之间、段落和层次标题以及各段落之间均为1.5倍行距。关键词与关键词之间用“;”隔开)

目录

[摘要 I](#_Toc450580355)

[**Abstract** II](#_Toc450580356)

[前言 1](#_Toc450580357)

[第一章 基本原理 1](#_Toc450580358)

[1.1 深度学习 1](#_Toc450580359)

[1.1.1 深度学习算法 1](#_Toc450580360)

[1.1.2 卷积神经网络 2](#_Toc450580361)

[1.1.3 深度学习系统架构 2](#_Toc450580362)

[1.2 可编程逻辑门阵列（FPGA） 2](#_Toc450580363)

[1.2.1 架构概述 2](#_Toc450580364)

[1.2.2 高层次综合（HLS） 2](#_Toc450580365)

[1.3 片上系统（SoC） 2](#_Toc450580366)

[1.4 论文的主要工作 2](#_Toc450580367)

[第二章 总体设计 3](#_Toc450580368)

[2.1 深度学习框架Caffe 3](#_Toc450580369)

[2.1.1 软件架构 3](#_Toc450580370)

[2.1.2 GPU加速设计 3](#_Toc450580371)

[2.1.3 第三方库依赖 3](#_Toc450580372)

[2.1.4 使用方法 4](#_Toc450580373)

[2.2 ZYNQ平台 4](#_Toc450580374)

[2.2.1 系统架构 4](#_Toc450580375)

[2.2.2 ARM处理器 4](#_Toc450580376)

[2.2.3 FPGA硬件 4](#_Toc450580377)

[2.2.4 Zedboard开发板卡 4](#_Toc450580378)

[2.3 SDSoC开发工具 4](#_Toc450580379)

[2.3.1 基本原理与组成 4](#_Toc450580380)

[2.3.2 工作流程 4](#_Toc450580381)

[2.4 综合设计方案 5](#_Toc450580382)

[第三章 主体工作 5](#_Toc450580383)

[3.1 系统架构设计 5](#_Toc450580384)

[3.2 FPGA加速器设计 5](#_Toc450580385)

[3.2.1 GEMM算法优化实现 5](#_Toc450580386)

[3.2.2 GEMV算法优化实现 6](#_Toc450580387)

[3.3 ARM处理系统与数据通路设计 6](#_Toc450580388)

[3.3.1 ARM处理系统设计 6](#_Toc450580389)

[3.3.2 数据通路设计 6](#_Toc450580390)

[3.4 应用：基于卷积神经网络的立体匹配算法实现 6](#_Toc450580391)

[3.4.1 立体匹配算法 6](#_Toc450580392)

[3.4.2 神经网络设计 6](#_Toc450580393)

[3.4.3 前处理与后处理 6](#_Toc450580394)

[第四章 测试结果 7](#_Toc450580395)

[4.1 FPGA加速器测试 7](#_Toc450580396)

[4.1.1 GEMM性能测试 7](#_Toc450580397)

[4.1.2 GEMV性能测试 7](#_Toc450580398)

[4.1.3 综合性能测试 7](#_Toc450580399)

[4.1.4 资源利用 7](#_Toc450580400)

[4.2 Caffe整体性能测试 8](#_Toc450580401)

[4.2.1 单元测试 8](#_Toc450580402)

[4.2.2 样例应用测试 8](#_Toc450580403)

[4.3 立体匹配算法效率测试 8](#_Toc450580404)

[第五章 总结与展望 8](#_Toc450580405)

[5.1 主要工作总结 9](#_Toc450580406)

[5.2 未来优化方向 9](#_Toc450580407)

[结论 10](#_Toc450580408)

[参考文献 11](#_Toc450580409)

[附录一 12](#_Toc450580410)

[致谢 13](#_Toc450580411)

（备注：目录按2～3级标题编写“目录”二字使用黑体小二号字居中书写，隔行书写目录内容，“摘要、Abstract、正文的一级标题、结论、参考文献、附录目录、致谢”采用黑体小四号字书写，正文其他层次标题均采用宋体小四号字书写；“摘要、Abstract”与“正文”之间隔一行；正文中的二级标题、三级标题相对于上一级标题均缩进二个空格书写；目录行与行之间均为1.5倍行距；目录内容多者，正文中的二、三级标题可使用宋体五号字书写，行与行之间可采用单倍行距。目录修改时单击目录点右键选择更新域，选择更新整个目录。之后，在“Abstract”的页码后回车加一空行。）

# 前言

正文采用宋体（英语用Times New Roman）小四号字，毕业论文、毕业设计行与行之间、段落和层次标题以及各段落之间均为1.5倍行距。[[1]](#footnote-1)

毕业论文的前言应综合评述前人工作，说明论文工作的选题目的、背景和意义、国内外文献综述以及论文所要研究的主要内容，对所研究问题的认识，以及提出问题等。前言只是文章的开头，可不写章号，也可不出现“前言”二字。

毕业设计的前言部分应说明设计的目的、意义、范围及应达到的技术要求；简述课题在国内外的发展概况及存在的问题；阐明设计的指导思想；阐述设计应解决的主要问题。[[2]](#footnote-2)

# 第一章 基本原理

正文采用宋体（英语用Times New Roman）小四号字，毕业论文、毕业设计行与行之间、段落和层次标题以及各段落之间均为1.5倍行距。

## 1.1 深度学习

正文采用宋体（英语用Times New Roman）小四号字，毕业论文、毕业设计行与行之间、段落和层次标题以及各段落之间均为1.5倍行距。

具体内容 具体内容 具体内容 具体内容 具体内容 具体内容 具体内容[[[3]](#endnote-1)]。

### 1.1.1 深度学习算法

正文采用宋体（英语用Times New Roman）小四号字，毕业论文、毕业设计行与行之间、段落和层次标题以及各段落之间均为1.5倍行距。

### 1.1.2 卷积神经网络

正文

### 1.1.3 深度学习系统架构

正文

## 1.2 可编程逻辑门阵列（FPGA）

正文

### 1.2.1 架构概述

正文

### 1.2.2 高层次综合（HLS）

正文

## 1.3 片上系统（SoC）

正文

## 1.4 论文的主要工作

正文

# 第二章 总体设计

正文

## 2.1 深度学习框架Caffe

正文

### 2.1.1 软件架构

正文

### 2.1.2 GPU加速设计

正文

### 2.1.3 第三方库依赖

正文

### 2.1.4 使用方法

正文

## 2.2 ZYNQ平台

正文

### 2.2.1 系统架构

正文

### 2.2.2 ARM处理器

正文

### 2.2.3 FPGA硬件

正文

### 2.2.4 Zedboard开发板卡

正文

## 2.3 SDSoC开发工具

正文

### 2.3.1 基本原理与组成

正文

### 2.3.2 工作流程

正文

## 2.4 综合设计方案

正文

# 第三章 主体工作

正文

## 3.1 系统架构设计

正文

## 3.2 FPGA加速器设计

正文

### 3.2.1 GEMM算法优化实现

正文

### 3.2.2 GEMV算法优化实现

正文

## 3.3 ARM处理系统与数据通路设计

正文

### 3.3.1 ARM处理系统设计

正文

### 3.3.2 数据通路设计

正文

## 3.4 应用：基于卷积神经网络的立体匹配算法实现

正文

### 3.4.1 立体匹配算法

正文

### 3.4.2 神经网络设计

正文

### 3.4.3 前处理与后处理

正文

# 第四章 测试结果

正文

## 4.1 FPGA加速器测试

正文

### 4.1.1 GEMM性能测试

正文

### 4.1.2 GEMV性能测试

正文

### 4.1.3 综合性能测试

正文

### 4.1.4 资源利用

正文

## 4.2 Caffe整体性能测试

正文

### 4.2.1 单元测试

正文

### 4.2.2 样例应用测试

正文

## 4.3 立体匹配算法效率测试

正文

# 第五章 总结与展望

正文

## 5.1 主要工作总结

正文

## 5.2 未来优化方向

正文

# 结论

毕业论文的结论是对整个论文主要成果的归纳，应突出论文的创新点，以简练的文字对论文的主要工作进行评价。若不可能得出应有的结论，则需进行必要的讨论。可以在结论或讨论中提出建议、研究设想及尚待解决的问题等。

毕业设计的结论是概括说明设计的情况和价值，分析其优点和特色，有何创新，性能达到何水平，并应指出其中存在的问题和今后改进的方向。

结论要单独成页，“结论”二字采用黑体小二号字居中书写，隔行采用宋体（英语用Times New Roman）小四号字书写具体内容，行与行之间、各段落之间均为1.5倍行距。

# 参考文献

# 附录一

附录序号采用“附录一”、“附录二”、“附录1”、“附录2” 、“附录A”、“附录B”等，用四号黑体字左起顶格排写，其后不加标点符号，空一行书写附录内容。附录内容采用宋体（英语用Times New Roman）小四号字书写。行与行之间、各段落之间均为1.5倍行距。

附录可有可无，当无附录内容时，此页可删除；当附录唯一时，只写“附录”即可。

# 致谢

致谢要单独成页，“致谢”二字采用黑体小二号字居中书写，隔行采用宋体（英语用Times New Roman）小四号字书写具体内容。行与行之间、各段落之间均为1.5倍行距。

1. 依次书写作者名、资料的篇名、发表的刊物名、出版年份和期号。 [↑](#footnote-ref-1)
2. 如果是著作则应写明出版单位和出版年份，见《黑龙江大学本科生毕业论文（设计）撰写规范》。 [↑](#footnote-ref-2)
3. [] 主要责任者.文献题名[J].刊名.出版年份,卷号(期号):起止页码． [↑](#endnote-ref-1)