

基于 Rocket-Chip 开源处理器的 CNN 加速模块的设计及实现

杨维科, 贺光辉, 景乃锋

(上海交通大学 电子信息与电气工程学院, 上海 200240)

摘要: 基于 RISC-V 开源指令集及 Rocket-Chip 开源处理器, 提出了一种基于 Eyeriss 结构的卷积神经网络加速模块, 并与处理器连接形成完整的系统. 该加速器结构通过进行横向 (卷积核权值), 纵向 (输出计算结果) 以及斜向 (输入图像) 的数据重用, 大大减少了卷积层的时间消耗. 此外, 利用加州大学伯克利分校开发的 Chisel3 语言, 一方面生成该系统的 C 语言模拟器进行调试, 另一方面生成 Verilog 进行综合和布局布线. 通过对 LeNet-5 手写数字识别网络及 MNIST 数据集进行测试, 准确度、速度等都达到了令人满意的结果.

关键词: 开源处理器; RISC-V; 卷积神经网络; Eyeriss

Design and Implementation of CNN Acceleration Module Based on Rocket-Chip Open Source Processor

YANG Wei-ke, HE Guang-hui, JING Nai-feng

(School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China)

Abstract: Based on the RISC-V open source instruction set and Rocket-Chip open source processor, we propose a convolution neural network acceleration module based on Eyeriss structure and form a complete system with the processor connection. The accelerator structure greatly reduces the time consumption of the convolution layer by performing data reuse in the lateral (convolution core weight), longitudinal (output calculation result) and oblique (input image). In addition, the use of the UCB developed Chisel3 language, on the one hand to generate the system C language simulator for debugging, on the other hand generate Verilog for integrated and layout and routing. Through the LeNet-5 handwritten digital identification network and MNIST data set, we achieved satisfactory results on speed and accuracy.

Key words: open source processor; RISC-V; convolutional neural network; Eyeriss

作者简介

杨维科 男, (1993-), 硕士研究生. 研究方向为可重构处理器的系统架构. E-mail: wakefielddyoung@sjtu.edu.cn.

贺光辉 男, (1980-), 博士, 副研究员. 研究方向为通信系统、VLSI 等.

景乃锋 男, (1982-), 博士, 副研究员. 研究方向为计算机体系结构.