

文章编号: 2095-2163(2022)09-0119-04

中图分类号: TP332

文献标志码: A

基于 RISC-V 架构的异常检测系统设计

张嘉, 李新增, 康鹏, 朱海云, 金婕

(上海工程技术大学 电子电气工程学院, 上海 201620)

摘要: 本文针对在金属的生产和制造过程中的金属图片表面异常, 设计了一种基于第五代精简指令集计算机的片上系统。运用结构精简、准确率高的生成对抗网络算法实现了对异常金属图片数据集的分类; 在 Xilinx Nexys4 DDR2 FPGA 开发板上通过设计的硬件片上系统生成比特流文件, 实现了整个端到端的识别系统, 并结合已有的数据测试集验证了识别系统的实用性。实验结果表明, 在异常金属图片数据集中, 输入图片分辨率为 64×64 , 在板上系统内实现了高达 99.9% 的识别准确率, 为工业生产制造与异常检测提供了质量和效率保证。

关键词: 第五代精简指令集计算机; 异常检测; 生成对抗网络; 片上系统

Design of anomaly detection system based on RISC-V architecture

ZHANG Jia, LI Xinzeng, KANG Peng, ZHU Haiyun, JIN Jie

(School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

[Abstract] Aiming at the abnormal surface of the metal pictures in the metal production and manufacturing process, a SOC based on the fifth-generation simplified instruction system RISC-V is designed. The detection and classification of anomaly metal images data sets are realized by using a simplified and highly accurate generative adversarial network algorithm. The whole end-to-end identification system is verified and implemented on Xilinx Nexys4 DDR2 FPGA development board. The experimental results show that in the abnormal metal images data set, the input images resolution is 64×64 , and the recognition accuracy rate achieved in the on-board system can reach 99.9%, which provides the guarantee of the quality and efficiency for industrial production, manufacturing and anomaly detection.

[Key words] Reduced Instruction Set Computer-Five (RISC-V); abnormal detection; generative adversarial networks; System On Chip (SOC)

0 引言

异常检测是发现与大部分对象不同的个体, 即发现离群点, 也称偏差检测, 异常对象是相对罕见的。异常检测主要应用于欺诈检测、入侵检测和医疗领域等。金属的工业化生产与制造过程中, 需要根据金属图片进行异常检测, 辨别高品质金属样本与低品质金属样本。

本文基于第五代精简指令集计算机 (Reduced Instruction Set Computer-Five, RISC-V) 设计了异常图片检测方案, 在检测算法上摒弃了传统的图像识别算法, 采用了人工智能应用领域中较为常见的生成对抗网络 (Generative Adversarial Networks, GAN)。通过测试实验, 验证了该方法的可行性以及实用性。

基金项目: 国家自然科学基金 (61801286)。

作者简介: 张嘉 (1997-), 男, 硕士研究生, 主要研究方向: FPGA、神经网络; 李新增 (1997-), 男, 硕士研究生, 主要研究方向: 无线通信、通信基带算法; 康鹏 (1995-), 男, 硕士研究生, 主要研究方向: 集成电路可测试性设计、CPU 设计; 朱海云 (1994-), 女, 硕士研究生, 主要研究方向: 无线通信、FBMC/OQAM 系统; 金婕 (1978-), 女, 博士, 副教授, 硕士生导师, 主要研究方向: 视频编解码、数字信号处理和 VLSI。

通讯作者: 金婕 Email: 02140001@sues.edu.cn

收稿日期: 2021-11-17

1 第五代精简指令集计算机

精简指令集计算机 (Reduced Instruction Set Computer, RISC) 具有指令长度固定、指令格式种类少、寻址方式种类少等特点。复杂指令的功能可由简单指令的组合来实现, 最具代表性的精简指令集计算机为 ARM 架构计算机。RISC 相较于以 x86 架构为代表的复杂指令集计算机 (Complex Instruction Set Computer, CISC), 具有指令更为精简、更易操作的特点。RISC-V 则是由加州大学伯克利分校所研发并开源的第五代精简指令集计算机, 是一款高质量、免许可证、开放的指令集, 也是一套由 RISC-V 基金会所通力维护的标准, 以达到代码开源、指令集开放的目的。采用 RISC-V 设计的芯片, 面积功耗比更低, 更为简洁, 面积功耗以及性能之比更加平

衡,在实际应用中,其性价比更高^[1]。

本设计采用了加州大学伯克利分校于2015年所开源的RISC-V裸核,在该裸核的基础上构建片上系统,即包括DDR2控制器、VGA显示器、SD卡读写器等外围设备,并生成了可下载于FPGA板上的比特流文件进行板上开发。

2 片上系统

本文的异常检测系统需要基于实际的硬件平台进行验证与测试,故仅仅通过中央处理器(Central Processing Unit, CPU)裸核进行系统设计不能满足整套测试验证硬件系统平台搭建的需求。故通过CPU裸核的选型,进行了片上系统(System-On-Chip, SOC)的设计。由CPU裸核通过AMBA AXI4总线协议,与MIG 7series、BRAM Controller相连,其中MIG 7series为DDR2控制器,BRAM Controller为块RAM控制器;PSOC是负责协调外设的控制器,目前挂在内核上的外设设有IIC接口、VGA接口、以太网接口和PS2等。片上系统总体设计如图1所示。对此拟展开研究分述如下。

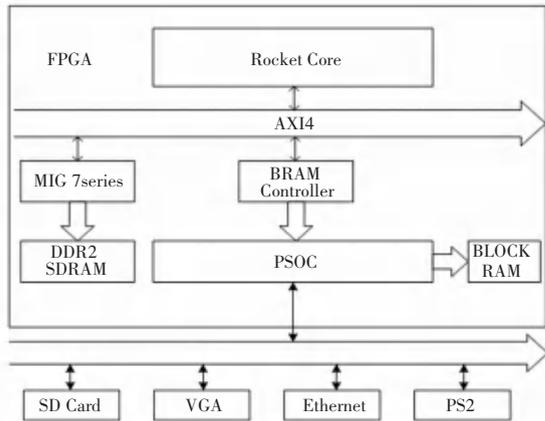


图1 片上系统设计

Fig. 1 System-On-Chip design

2.1 中央处理器

中央处理器Rocket是一款采用RISC-V指令系统的开源处理器,拥有五级流水线,且为单发射顺序执行的64位标量处理器。Rocket支持内存管理单元和分页虚拟内存,具有兼容IEEE 754-2008标准的浮点处理单元。

Rocket内部总线协议为TileLink协议,连接内部核内各个子模块与计算核心Rocket Tile。计算核心中集成了指令缓冲器(IBUF)、控制与状态寄存器(CSR)、中断单元(BPU)、算术逻辑单元(ALU)、乘法器单元(MULDIV)、浮点处理单元(FPU)等^[2]。

2.2 外设片上系统

外设片上系统(Periphery System-On-Chip, PSOC)内部集成了子模块SD卡读写器、VGA显示器、Ethernet、PS2键盘鼠标等模块的接口控制器。与Rocket核相连的BRAM Controller通过PSOC与块RAM、外围设备搭建数据通路。外围设备中,SD卡读写器采用SPI通信协议,VGA显示器接口通过HSYNC和VSYNC信号线控制行与帧时序进行VGA显示。

2.3 内存接口生成控制器

内存接口生成器(Memory Interface Generator, MIG) 7 series是Xilinx公司的7系列开发板所使用的内存控制器。MIG 7series控制器在软件VIVADO中创建,配置时钟参数以及引脚约束后,需单独编译此内存控制器IP核,编译成功则生成了可调用的MIG 7series控制器IP核工程,在SOC设计工程中可直接添加此IP核。

3 网络结构与异常检测

由于本文的应用设计需基于硬件系统,即部署于FPGA开发板内的硬件系统以及操作系统,在该系统内进行网络结构的设计与异常检测。文中将给出探讨论述如下。

3.1 网络结构设计

生成对抗神经网络设计了3组子网络。第一组子网络是一个自动编码器网络和一个解码器网络,第二组子网络是一个编码器网络,第三组子网络是一个判别器网络。

本文设计的具体网络结构如图2所示。生成对抗网络模型的生成器网络包括一个自动编码器网络、一个解码器网络和一个编码器网络。生成器网络分别通过自动编码器网络和解码器网络学习输入数据表示并重构输入图像。生成器网络首先读取输入图像 X ,并将其前向传播到第二个编码器网络中。第一个子网络中紧邻归一化卷积层和泄露线性修正激活函数,网络中的下采样过程通过压缩输入图片 X 为向量 Z 来实现,向量 Z 称为生成器的瓶颈特征,并假设 Z 为包含 X 的最好表征的最小维度。生成器网络中的解码器部分采用了深度卷积生成对抗网络(Deep Convolutional Generative Adversarial Networks, DCGAN)的体系结构,包括卷积变换层、ReLU激活函数、归一化层、以及tanh激活函数。

第二组子网络是编码器网络,进一步将第一组网络所重构的图像 T 进行下采样。本文所设计的

生成对抗网络模型中,第二组子网络中的编码器网络与第一组子网络中的自动编码器网络结构相同,但使用了不同的网络参数。编码器网络 E 将图像 T 下采样到 \hat{Z} ,且其特征表示为 $\hat{Z} = E(T)$ 。为了保持向量 Z 与向量 \hat{Z} 的一致性,向量 Z 与向量 \hat{Z} 的维度保持一致。

第三组子网络是判别器网络,负责分辨输入 X 和输出 T 的真假。该子网络是 DCGAN 中引入的标准判别器网络。

图片数据输入网络之前,已将异常图片修改为网络特定的输入尺寸,网络输出再经过 *Softmax* 层输出为真类和假类,输出层节点数为 2,即为二分类。

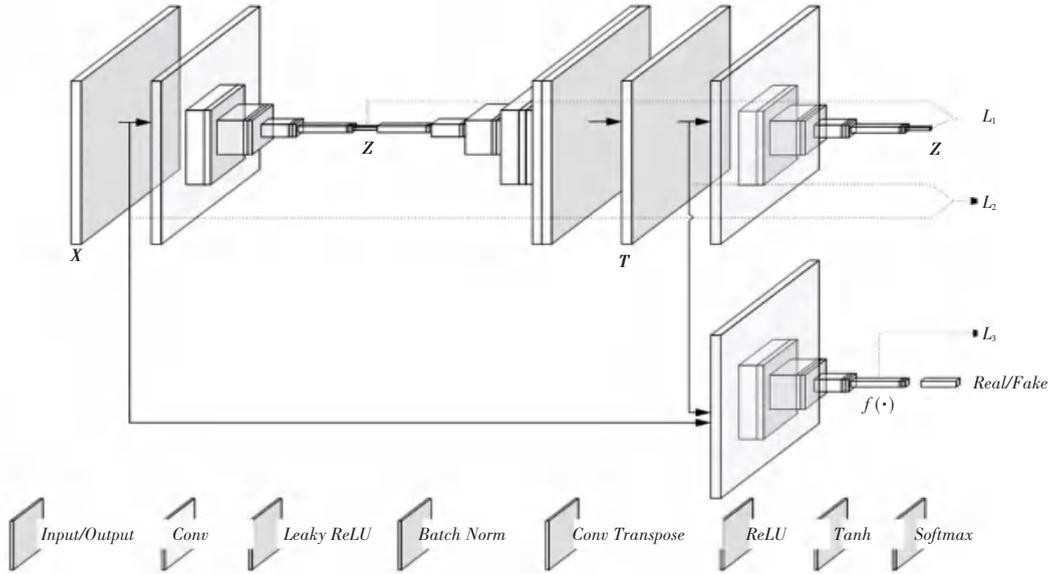


图 2 网络结构

Fig. 2 Network structure

3.2 金属检测

在金属铝材的生产制造过程中,金属表面会因为各方面因素而产生差异性,出现裂纹、划伤、起皮等不规则纹理,这些瑕疵会引起金属应用时的不确定性,从而增加了工业生产及测试的工作量。实验中读取各个不同特征的金属图片进行训练,包括斑块、开裂、划痕等。通过生成对抗网络对缺陷图片进行特征捕捉,将具有异常缺陷的图片特征与表面光滑、即完好的图片特征一并导入网络中,从而生成权重文件以供测试识别过程中对真类图片及假类图片进行分类。

异常检测算法的识别步骤:首先,导入深度学习框架,创建并设定神经网络参数;其次,导入数据集并开始训练,训练完成时保持模型权重数据;最后,导入训练完成的模型并进行图片的识别。

4 实现

4.1 系统平台

本文提出的异常图片检测与识别系统是基于 Xilinx 公司的 XC7A100T FPGA 芯片构成的开发板

Nexys4 DDR2, SOC 使用 VIVADO 软件设计并实现。

4.2 数据准备

由于所需的钢材图片数据集为二次修改后生成的图片,故在网络训练前,需要对钢材图片数据集进行微调,包括修改图片名称、图片大小及图片格式等。

修改后的图片为 32×32 或 64×64 的 PNG 图片,本实验所使用的训练集和测试集为 64×64 的 PNG 图片。数据集中的每一类的标签即为该类所在文件夹的文件名。在网络训练时,可读取该文件夹名作为对应的标签。本文将取另外一部分同源图片作为网络的测试验证集。

4.3 网络训练

基于 Python-OpenCV 以及 Pytorch 深度学习框架环境进行实验。

由于本文采用的训练方法是反向传播法 (back-propagation),权重梯度更新参数设置过大会引起过拟合效应,并且动量参数应与权重梯度更新参数相近,故设定该反向传播中所需要的权重梯度更新参数为 0.1,动量参数也为 0.1,训练停止条件为

最大迭代次数或阈值,满足其一则停止本次训练。基于算法鲁棒性考虑,最大迭代次数或阈值依次设定为 500 及 0.01,设定了整体训练集的训练次数 *epoch* 为 1 200,即整个数据集训练 1 200 次后,就会保存已训练完成的模型权重数据参数以供识别过程中导入该模型。

4.4 实验结果

本文所提出的基于 FPGA 的 RISC-V 架构开发的识别应用系统有效地实现了异常图片的检测与识别,异常图片检测准确率可达到 99.9%。本实验中的假类图片与真类图片测试中可清晰比出真假图片,如图 3、图 4 所示。

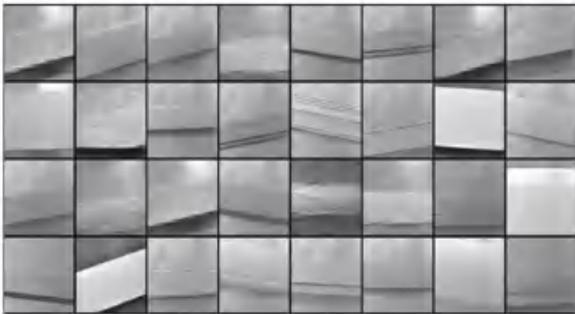


图 3 假类图片

Fig. 3 Negative samples

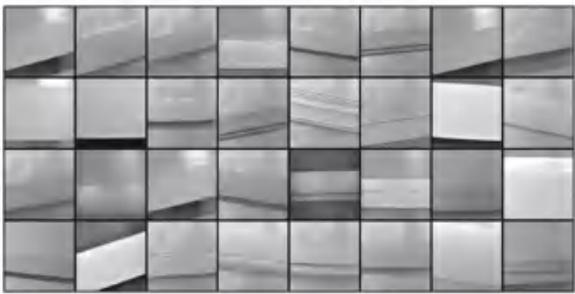


图 4 真类图片

Fig. 4 Positive samples

与采用无监督、预训练的卷积神经网络以及卷积自编码器网络的结果对比,采用生成对抗网络算法具有一定的优越性,算法比较结果见表 1。

表 1 算法比较

Tab. 1 Algorithms comparison

算法	准确率/%	平台
无监督预训练卷积神经网络 ^[3]	55.5	CPU
卷积自编码器网络 ^[4]	98.8	CPU
生成对抗网络(The proposed)	99.9	FPGA

5 结束语

本文通过 FPGA 平台,搭建了基于 RISC-V 架构的 SOC,并在 SOC 上完成了异常图片检测与识别系统的设计;在 SOC 中运用高准确率以及结构精简的生成对抗网络,对异常金属图片进行检测,实现了高达 99.9%识别准确率,为工业生产制造与异常检测提供了有效帮助,为基于 RISC-V 的微系统开发和芯片设计奠定了基础。

参考文献

- [1] ANDREW W, KRSTE A. The RISC-V Instruction Set Manual Volume I: Unprivileged ISA[EB/OL]. [2019-12-13]. <https://riscv.org/technical/specifications/>.
- [2] KRSTE A, RIMAS A, JONATHAN B, et al. The rocket chip generator[EB/OL]. [2016-04-15]. <http://ww-w2.eecs.berkeley.edu/Pubs/TechRpts/2016/EECS-2016-17.pdf>.
- [3] SOUKUP D, HUBER-MÖRK R. Convolutional neural networks for steel surface defect detection from photometric stereo images[M]//BEBIS G, et al. International Symposium on Visual Computing(ISVC 2014), Part I, LNCS 8887. Cham: Springer, 2014: 668-677.
- [4] HEGER J, DESAI G, ABDINE M Z E. Anomaly detection in formed sheet metals using convolutional autoencoders[J]. Procedia CIRP, 2020, 93: 1281-1285.

(上接第 118 页)

参考文献

- [1] 饶文姬,赵良举,刘朝,等. 利用 LNG 冷能与工业余热的有机朗肯循环研究[J]. 工程热物理学报,2014,35(02): 213-217.
- [2] 冯驯,徐建,王墨南,等. 有机朗肯循环系统回收低温余热的优势[J]. 节能技术,2010,28(05): 387-391.
- [3] 张任平,刘江. 基于热力学第二定律的有机朗肯循环效能分析

[J]. 低温与超导,2018,46(11):61-65.

- [4] 王辉涛,王华,葛众. 中低温地热发电有机朗肯循环工质的选择[J]. 昆明理工大学学报(自然科学版),2012,37(01):41-46.
- [5] 刘美丽,燕旭,杨凯,等. 低温余热回收 ORC 系统工质的筛选[J]. 制冷与空调,2017,17(01):44-49,43.