

文章编号: 1674-8190(2024)02-188-07

机载超轻量化卷积神经网络加速器设计

石添介, 刘飞阳, 张晓

(航空工业西安航空计算技术研究所 预先研究部, 西安 710068)

摘要: 卷积神经网络庞大的权重参数和复杂的网络层结构, 使其计算复杂度过高, 所需的计算资源和存储资源也随着网络层数的增加而快速增长, 难以在资源和功耗有严格要求的机载嵌入式计算系统中部署, 制约了机载嵌入式计算系统朝着高智能化发展。针对资源受限的机载嵌入式计算系统对超轻量化智能计算的需求, 提出一套全流程的卷积神经网络模型优化加速方法, 在对算法模型进行超轻量化处理后, 通过组合加速算子搭建卷积神经网络加速器, 并基于 FPGA 开展网络模型推理过程的功能验证。结果证明: 本文搭建的加速器能够显著降低硬件资源占用率, 获得良好的算法加速比, 对机载嵌入式智能计算系统设计具有重要意义。

关键词: 嵌入式计算系统; 卷积神经网络; 轻量化; 硬件加速器; FPGA 验证

中图分类号: V247; TP183; TP368.1

文献标识码: A

DOI: 10.16615/j.cnki.1674-8190.2024.02.21

Design of airborne ultra-lightweight convolutional neural network accelerator

SHI Tianjie, LIU Feiyang, ZHANG Xiao

(Pre-research Department, AVIC Xi'an Aeronautics Computing Technique Research Institute, Xi'an 710068, China)

Abstract: The huge weight parameters and complex network layer structure of convolutional neural network make its computational complexity too high, and the required computing resources and storage resources also increase rapidly with the increase of network layers, so it is difficult to deploy in airborne embedded computing systems with strict requirements on resources and power consumption, which restricts the development of airborne embedded computing systems towards high intelligence. Aiming at the demand of ultra-lightweight intelligent computing in the resource-constrained airborne embedded computing system, a set of optimization and acceleration strategy of convolutional neural network model is proposed. After ultra-lightweight processing of the algorithm model, a convolutional neural network accelerator is built by combining acceleration operators, and the function verification of network model reasoning process is carried out based on FPGA. The results show that the established accelerator can significantly reduce the occupancy rate of hardware resources and obtain a good algorithm speedup ratio, which is of important significance for the design of airborne embedded intelligent computing system.

Key words: embedded computing system; convolutional neural network; ultra-lightweight; hardware accelerator; FPGA implementation

收稿日期: 2023-07-07; 修回日期: 2023-09-19

通信作者: 石添介(1997-), 男, 硕士, 工程师。E-mail: 15202955644@163.com

引用格式: 石添介, 刘飞阳, 张晓. 机载超轻量化卷积神经网络加速器设计[J]. 航空工程进展, 2024, 15(2): 188-194.

SHI Tianjie, LIU Feiyang, ZHANG Xiao. Design of airborne ultra-lightweight convolutional neural network accelerator[J]. Advances in Aeronautical Science and Engineering, 2024, 15(2): 188-194. (in Chinese)

0 引言

近年来,随着深度学习算法模型的快速发展,一方面,以卷积神经网络(CNN)为核心的算法模型可以有效解决目标识别、边缘检测、图像分割等领域的问题,对提升新一代飞机/无人机机载嵌入式计算系统的智能化任务处理能力有巨大技术价值^[1];另一方面,卷积神经网络算法虽然在性能方面效果优异,但模型规模过大,计算复杂度过高,难以将其完整部署在嵌入式计算设备当中^[2]。另外,典型嵌入式环境中的任务通常对响应时间、准确度有较高的需求,需要进一步对算法模型的推理速度进行优化。因此,为了有效提升神经网络模型在嵌入式计算领域的应用效果,对算法模型进行轻量化、硬件优化加速等方面的研究,具有重要意义。

在模型轻量化方面,二值神经网络(BNN)是深度神经网络超轻量化近两年发展的一个重要方向,其核心目标是通过算法模型优化实现低数据位宽的神经网络运算,从而大幅度降低人工智能应用的计算复杂度和存储开销,实现低成本、低功耗的智能计算。二值化通常指将深度神经网络权重值和激活函数值量化到+1和-1实现^[3]。通过二值化,深度神经网络算法模型的权重矩阵只占用一个比特,相比于单精度浮点型权重矩阵,算法模型的内存消耗理论上能减少32倍,因此二值化神经网络在模型压缩上具有很大的优势;而当权重值和激活函数值同时进行二值化之后,原来32个浮点型数据的乘加运算,可以通过一次异或运算(Xnor)和一次Popcount运算解决^[4],因此二值神经网络在模型超轻量化方面上具有巨大潜力。进一步地,2016年蒙特利尔大学首次提出二值神经网络概念,将核心的神经网络架构进行二值化处理,并提出使用随机梯度下降方法训练网络^[5],以获得超轻量化的模型,但带来的代价是精度损失严重。

在硬件部署优化及加速方面,国内外均研制出了专用的神经网络加速器/芯片^[6],用于在各种嵌入式场景中部署神经网络算法以完成各种目标任务。例如,特斯拉于2020年推出车载嵌入式

SoC芯片,内部嵌入两块自主设计的神经网络加速器NNA^[7];清华大学推出的Thinker系列低功耗终端AI芯片^[8],可用于部署及加速基于神经网络的人脸识别、语音识别等算法。然而,专用加速芯片的研制难度高,研制周期较长,重构性较差,无法满足进行快速技术验证的目的。FPGA因其良好的并行计算能力已经被证实能够适应神经网络中的核心计算特征^[9],实现网络推理加速。对于密集型计算任务,FPGA在整数乘法甚至浮点乘法计算方面已经可以与中高端GPU处理卡相当;对于流式计算任务,FPGA同时拥有流水线并行和数据并行两种方式,比GPU延迟要低很多^[10]。随着相关技术的发展,越来越多的FPGA具有可重构功能,可以实现硬件资源的重配置,大幅提高了芯片的利用率^[11]。然而,如何根据不同任务需求灵活配置FPGA从而更高效地完成神经网络模型的部署优化及加速是目前仍然需要研究和解决的问题。

本文面向以卷积神经网络为核心的大规模神经网络模型,提出一套全流程的卷积神经网络模型优化加速方法,包括模型超轻量化设计、超轻量模型功能算子设计、优化以及基于FPGA器件的完整加速器设计。针对典型以卷积神经网络为核心的算法模型,采用二值化策略对模型进行充分量化,根据量化好的神经网络模型,给出一种基于高层次综合的硬件功能单元设计、优化策略,进一步设计出轻量化神经网络加速器原型;采用FPGA器件搭建轻量化神经网络加速器原型。

1 卷积神经网络模型超轻量化设计

1.1 基本概念

卷积神经网络通常由维度不同的卷积层、下采样层、全连接层等隐藏层组成,每一层将一组特征图作为输入,执行完层内响应的计算过程后生成新的特征图输入到下一层^[12]。早期的卷积神经网络模型较为简单,如1998年Lecun等^[13]提出的用于手写字符识别的LeNet网络,其网络模型的深度只有7层,且每层的特征向量很少;21世纪初,深度学习领域的三巨头之一Hinton^[14]指出,多层的

神经网络结构能够学习到更深刻、更本质的特征。因此,近年来一系列深度卷积网络模型被提出,包括 ImageNet、SSD、Faster R-CNN、Mask R-CNN、YOLO 系列等,并在各类计算机视觉任务中取得了良好的效果。

针对深度卷积网络模型规模大、计算资源耗费大的特点^[15],对算法模型采取超轻量化策略,即二值化设计。在不过多增加计算量的基础上完成对卷积神经网络前向计算过程的二值化处理,构建超轻量化的二值神经网络(BNN)。二值神经网络本质上也是一个 CNN,因其网络模型中的权重值和特征图的像素值会被二值化处理,从而大幅度降低了网络模型的大小。因此,BNN 模型可以被视为用于硬件加速的量化,是牺牲一定精度的 CNN 模型的极端实例。以典型 CNN 前向推理流程为例,BNN 和 CNN 前向推理时的差异如图 1 所示。

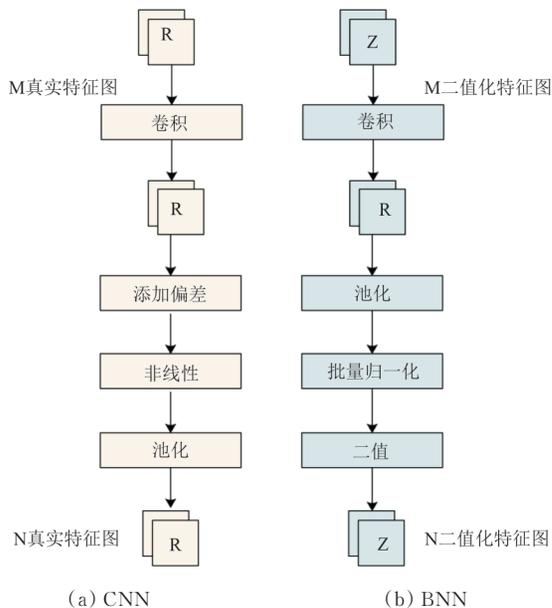


图 1 CNN 与 BNN 推理差异

Fig. 1 Difference of CNN vs BNN reasoning

BNN 中池化层在完成卷积操作后直接执行,因为添加二值偏差会进一步增加模型的误差。在 CNN 模型中,卷积层和全连接层的权重和计算结果都会使用 Sign 函数进行二值化操作,即特征图经过卷积层后从二值化变为整数,因此,二值神经网络设计时需要下一次卷积计算前增加批量归一化层对数据进行缩放,减少精度误差。

1.2 网络结构设计

1) 卷积单元设计

在二值神经网络中,卷积和激活值均被二值化,具体地,二值化将网络模型中原始 float 类型数据转换为二进制数据,该操作可以基于符号函数快速实现。

数据被二值化后,前向推理中复杂的矩阵运算可以由简单的按位 Xnor 运算和 Popcount 运算代替,以原始全精度神经网络来说,上述二值化计算过程如图 2 所示。

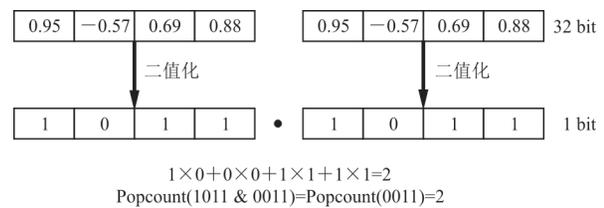


图 2 矩阵乘法二值化运算示意图

Fig. 2 Schematic diagram of binary operation of matrix multiplication

二值神经网络中除部分全精度卷积外均可以由二值卷积代替。具体来说,由于全精度卷积神经网络模型在采取二值化的量化策略后会不可避免地造成精度的下降。为了确保网络模型的精度,通常不对网络模型的原始输入及其参与运算的卷积层进行二值化,以尽可能多地让网络学习到输入的信息。因此,在资源较为紧缺的计算设备中,原始的卷积网络部分可采用一层全精度卷积层和多层二值卷积层的组合代替,二值卷积层的部分可以大幅降低计算时延和硬件资源占有率。

2) 其他单元设计

在二值神经网络的整体设计中,除卷积结构外,其他神经网络通用层如池化层、归一化层等均与实值神经网络相同。根据设计的卷积模块,与池化层等通用网络模块配合,即可搭建出完整的神经网络模型。

2 超轻量化神经网络加速器设计

本文提出的超轻量化神经网络加速器的设计流程如图 3 所示。

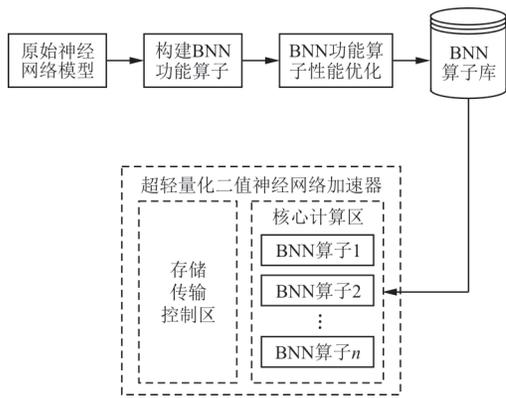


图3 超轻量化神经网络加速器设计流程图
Fig. 3 Design flow chart of ultra-lightweight neural network accelerator

针对超轻量化后的卷积神经网络模型,抽取核心网络层的计算特征,通过使用高层次综合技术及工具,采用标准软件语言(C/C++)完成对BNN前向推理过程中核心结构功能的实现及相关优化等工作,最终形成面向FPGA的硬件逻辑的功能算子库,进一步根据具体网络模型组合功能算子构建出二值神经网络加速器的核心计算区,结合存储、传输组件等搭建出完整的二值神经网络加速器,以支持不同结构的卷积神经网络模型超轻量化后在FPGA上的快速部署和加速工作。

2.1 功能算子设计

根据神经网络的计算过程和原理,BNN可以被分为三种不同的核心结构,分别是:全精度卷积层、二值卷积层以及二值全连接层,各层之间根据模型内部的数据流向还需要增加池化、批归一化、二值化等其他算子操作。

1) 全精度卷积层

通常来讲,BNN中第一层为全精度卷积层,确保输入的数据是原图像以保证网络模型的精度,非单比特卷积层可以利用较为成熟的行缓存结构完成定点2D卷积功能的实现。简单地说,尽管该层输入为原图像的大小,但该层的权重仍为二值化形式,因此可以使用符号反转操作直接替换传统卷积操作中的乘法,从而减少网络层推理时的时延。

2) 二值卷积层

二值卷积层是BNN在FPGA上实现时最关

键的结构,该结构占用了BNN网络模型中的较大一部分,因此该单元需要保持高吞吐量和资源效率。另外,由于每层输入的特征图大小不同,该核心算子需要动态支持不同尺寸的输入。由于特征图及相关权重已经变成了单比特数据,因此本文提出一种重新组织数据的方法,将来自不同特征图的数据重组为该二值卷积层模块的输入数据,即模块输入数据的每一位来自不同的特征图。为了在进行卷积计算时保证计算逻辑与特征图输入无关,本文在二值卷积层模块设计一个数据排序模块,用于重新排序模块输入数据的位。

二值卷积层计算的基本流程如图4所示,流程分为两个阶段。第一阶段,来自片上缓存的特征图输入到数据排序模块内,进行重新排序后被放到缓存区中;第二阶段,进行卷积计算并将每位的结果进行累加后缓存到Buffer中。

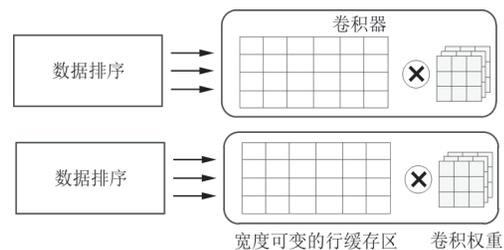


图4 二值卷积层计算过程示意图
Fig. 4 Schematic diagram of calculation process of two-value convolution layer

3) 二值全连接层及其他操作算子

二值全连接层单元设计比较简单,即在每个周期读取和输入相等数量的权重数据位,由于在全连接层中没有边缘填充,因此在二值全连接层内可以完成真正的单比特计算。换句话说,通过将数据和权重进行按位异或运算,然后将结果位与Popcount相加,以实现传统神经网络全连接层中的乘加运算过程。另外,池化、批标准化等操作在FPGA算子设计的过程中没有被过多的特殊设计,按照各算子的计算原理开发硬件逻辑代码即可。

2.2 功能算子优化

完成二值神经网络各个核心算子的功能设计后,可以基于高层次综合优化工具对各个核心算

子进行优化,以保证生成的 Verilog 代码具有更好的性能,可采用高层次综合优化策略。

1) 循环展开优化

高层次综合开发过程在默认情况下不展开循环,此时循环内部的运算采用的是相同资源,每一次计算都要等上一次循环计算结束,采用循环展开类似“复制”资源的操作,即采用几个相同的资源做运算。通过添加循环展开优化,能够缩短运行时间,充分利用空闲资源。循环展开优化指令如下: #pragma HLS unroll。

假设一个循环次数为 6 的 for 循环,采用 unroll 的优化策略,可以选择将它展开成 3 个 for 循环,每个循环只计算两步,即第 0 次和第 3 次、第 1 次和第 4 次、第 2 次和第 5 次分别为一次循环。在实际应用中选择展开为几个循环可以根据需求来设定。

2) 数据流优化

数据流优化要求数据按照顺序流动,没有数据反馈存在。大部分卷积神经网络模型前向推理过程中正好满足这个要求,可以引入数据流优化策略。采用数据流优化策略之后,函数之间会加入数据缓存区,在前一个计算还没完成时就开始下一次计算,从而改善吞吐量。数据流优化指令如下: #pragma HLS DATAFLOW & #pragma HLS pipeline。

以二值全连接层为例,一个完整的 BNN 中通常包括多个二值全连接层,而二值全连接层的计算过程依次包括读取数据、按位异或运算、Pop-count 相加、写出数据四个步骤。在该层的一次前向推理过程中,网络结构中所有二进制全连接层内的计算都是按照这四个步骤进行,同时所有的计算过程之间也是串联执行的,即做完一次完整的推理运算再进行下一次推理运算,这种运算方式时间较长,也没有达到并行的效果。针对上述问题,可以通过添加流水线技术,缩短循环延迟周期,用以增加系统的吞吐量。总体来说,上述优化策略可以根据在开发和调试过程中应用到 BNN 的不同功能算子以达到更优的硬件资源占用和更好的加速效果。

以二值卷积层前向推理的过程为例,结合高层次综合优化语句后的伪代码实现为:

二值卷积层前向计算伪代码

```

VaribaleBuffer linebuf;
ConvWeights wts;
IntergerBuffer outbuf;
for(i=0,i<n_input_words: i++){
# pragma HLS pipeline
//读取输入字,更新行缓存区
WordType word=input_data[i];
Bit-Sort(linebuf,word input width);
//每次更新权重
//开始处理下一个新的特征图
if (i%words per fimap=0)
    wts=weights[i/words per fimap]
//跨行缓存区执行转换
for (c=0;c<LINE BUF_COLS:c++){
# pragma HLS unroll
Outbuf [i% words per_fmap][c]+=conv(c, linebuf, wts);
}
}

```

该功能算子的基本操作是通过内部循环调用本文设计的数据排序模块以及卷积模块完成二值卷积层的计算过程。在最内层循环的卷积操作,由于数据矩阵各个位没有数据依赖,即可以使用循环展开的方式让原本串行循环完成的程序并行完成。对于整个功能函数,采用流水线优化,进一步降低整个计算过程的时延。

2.3 完整加速器设计

与目前主流的神经网络加速器设计不同的是,本文设计的加速器不是在固定的加速架构之上根据目标网络模型的结构生成硬件调度指令实现网络推理的功能,而是为不同算法直接在 FPGA 中构建不同的逻辑代码,从而更加高效地完成加速各类卷积神经网络模型的任务。简单来说,基于目标网络模型的结构,合理使用并组合功能算子,从而为目标卷积神经网络的超轻量化实现定制化设计加速器。加速的每一层由功能算子组成,每一层是独立的计算架构,层与层之间通过片内总线数据流通信,一旦上一层计算结束得到输出,下一层的计算就开始启动。

BNN 加速器的基本架构中(如图 5 所示),除了由功能算子构成的网络计算模块之外,还由片内数据缓存区、片内权重存储区以及片外一片内通信模块组成。整个加速器设计遵循数据驱动的

模式,即各个模块接收到数据后就开始进行计算,没有额外的控制模块,为简化加速器的设计,同时减少对 FPGA 片上资源的占用和对二值神经网络推理速度的响应。

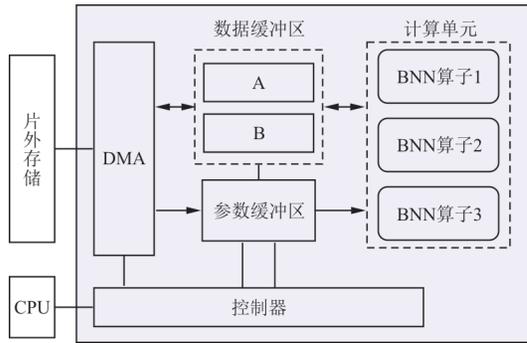


图 5 基于 FPGA 的超轻量化神经网络加速器架构设计

Fig. 5 Architecture design of ultra-lightweight neural network accelerator based on FPGA

在传统全精度 CNN 加速器设计中,神经网络隐藏层的输入特征图及层输出(下一层输入)大小通常超过了 FPGA 片上存储的大小,即传统 CNN 加速器的工作过程中网络中间输入/输出需要与片外 RAM 之间进行连续的数据通信,大幅增加模型前向推理的时延。在超轻量化二值神经网络中,由于数据被几十倍地压缩,因此网络模型的中间输出完全可以放在片内 RAM 中。因此,本文设计两个大小相等的数据缓存区,两个缓存区可以被交替使用,即网络模型的其中一层从缓存区 A 读取数据后计算得到的结果写入缓存区 B 中,下一层即可从缓存区 B 读取输入,将该层推理的结果写入缓存区 A 中。该设计避免了网络推理过程中层与层之间数据读取时的冲突,进一步提高整个网络推理逻辑的流水化程度,从而提升加速效果。

除了网络模型的原始输入和中间输入特征图之外,由于 BNN 紧凑的模型尺寸,本文将所有神经网络参数都保存在片上存储器中,以避免对磁盘存储器的大量访问。简单来说,根据二值神经网络的模型结构,为每一层构建大小可变的存储区,以避免片上存储资源的浪费。片外一片内通信模块仅用于存储输入图像、输出网络的预测值以及更新网络时输入新的权重值,通过使用 DMA 模块实现。

3 功能验证

为了对超轻量化卷积神经网络加速器进行性

能评估,本文针对典型的 LeNet-5 卷积神经网络模型,对其进行二值化处理,进而在 Xilinx 7z100 FPGA 器件上分别完成原始网络模型的部署和轻量化后的二值神经网络加速器搭建。LeNet-5 包含三层卷积层,两层下采样层和两次全连接层,每层有多个特征映射,每个特征映射有多个神经元。卷积层中通过卷积操作,可以使原信号特征增强,并且降低噪音。下采样层中,利用图像局部相关性的原理,对图像进行子抽样,可以减少数据处理量并且保留有用信息。

将不同图像数据输入原始 LeNet-5 模型和超轻量化后的 LeNet-5 模型进行训练,训练完成后的模型即可有效提取待测图像中的特征,用以有效完成图像识别任务。模型优化前后的推理时延及资源占用率如表 1 所示。

表 1 推理时延及资源占用率对比
Table 1 Comparison of inference delay and resource occupancy

模型	时延/ ms	块 RAM 占 用率/%	DSP 占 用率/%	触发器占 用率/%	查找表占 用率/%
原始模型 LeNet-5	18.861	42	33	48	51
超轻量化 后模型	3.645	7	2	3	6

从表 1 可以看出:本文设计的超轻量化卷积神经网络加速器,对网络推理任务的加速比达到了 5.17,对 FPGA 各类硬件的资源占有率大幅度降低。另外,本文选用 Cifar-100、Fashion-Minst 和 Labelme 三种典型图像识别数据集对原始 LeNet-5 模型和超轻量化后 LeNet-5 模型的图像识别准确率进行对比,对比结果如图 6 所示。

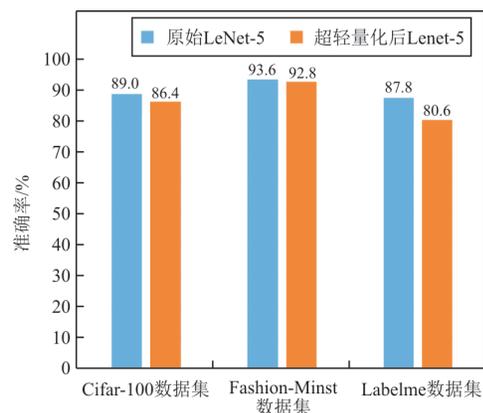


图 6 优化前后模型在不同数据集上图像识别的准确率
Fig. 6 Image recognition accuracy of the model before and after optimization on different datasets

相比于原始神经网络模型的图片检测准确度,加速器网络推理的结果虽然在两个数据集上的目标检测精度均有一定损失,但处于可接受的范围之内。

4 结 论

1) 针对模型规模大、计算复杂度高的原始卷积神经网络模型,对模型进行超轻量化处理,进一步设计和优化功能算子,功能算子的功能及性能均有良好表现。

2) 通过结合使用功能算子,在 FPGA 器件上搭建出超轻量化卷积神经网络加速器原型,可以在满足精度要求的前提下有效加速神经网络推理过程,并减少对计算资源的占用。

3) 本文所提出的全流程的卷积神经网络模型优化加速方法对于提升卷积神经网络在机载嵌入式计算系统等硬件资源和功耗严格受限的领域应用具有重要意义。

参 考 文 献

- [1] 程嘉晖. 基于深度卷积神经网络的飞行器图像识别算法研究[D]. 杭州: 浙江大学, 2017.
CHENG Jiahui. Research on aircraft image recognition algorithm based on depth convolutional neural network [D]. Hangzhou: Zhejiang University, 2017. (in Chinese)
- [2] LIN K T, CHIU C T, CHANG J Y, et al. High utilization energy-aware real-time inference deep convolutional neural network accelerator[C]// 2021 IEEE International Symposium on Circuits and Systems. US: IEEE, 2021: 1-7.
- [3] RASTEGARI M, ORDONEZ V, REDMON J, et al. XNOR-Net: ImageNet classification using binary convolutional neural networks[EB/OL]. [2023-07-07]. <https://arxiv.org/abs/1603.05279>.
- [4] 孙孝辉, 宋庆增, 金光浩, 等. 基于 ARM+FPGA 平台的二值神经网络加速方法研究[J]. 计算机应用研究, 2020, 37(3): 779-783.
SUN Xiaohui, SONG Qingzeng, JIN Guanghao, et al. Research on acceleration method of binary neural network based on ARM+FPGA platform [J]. Computer Application Research, 2020, 37(3): 779-783. (in Chinese)
- [5] YIN S Y, OUYANG P, YANG J X, et al. An energy-efficient reconfigurable processor for binary-and ternary-weight neural networks with flexible data bit width[J]. IEEE Journal of Solid-State Circuits, 2019, 54(4): 1120-1136.
- [6] 邬贵明. FPGA 矩阵计算并行算法与结构[D]. 长沙: 国防科学技术大学, 2011.
WU Guiming. Parallel algorithm and structure of FPGA matrix calculation [D]. Changsha: National University of Defense Technology, 2011. (in Chinese)
- [7] TALPES E, DAS SARMA D, VENKATARAMANAN G, et al. Compute solution for Tesla's full self-driving computer[J]. IEEE Transactions on Micro, 2020, 40(2): 25-35.
- [8] HU W W. Reconfigurable technology in the application of virtual instrument[J]. Advanced Materials Research, 2014, 1049/1050: 1137-1140.
- [9] COURBARIAUX M, HUBARA I, SOUDRY D, et al. Binarized neural networks: training deep neural networks with weights and activations constrained to +1 or -1 [EB/OL]. [2023-07-07]. <https://arxiv.org/abs/1602.02830>.
- [10] NIE G, XIAO L, ZHU M, et al. Binary neural networks as a general-propose compute paradigm for on-device computer vision [EB/OL]. [2023-07-07]. <https://paperswithcode.com/paper/binary-neural-networks-as-a-general-propose-review/#:~:text=Binary%20Neural%20Networks%20as%20a%20general-propose%20compute%20paradigm,similar%20degree%20of%20general%20applicability%20in%20vision%20tasks>.
- [11] 谷奎, 徐贵力, 王友仁. FPGA 动态可重构理论及其研究进展[J]. 计算机测量与控制, 2007, 15(11): 1415-1418.
GU Luan, XU Guili, WANG Youren. FPGA dynamic reconfigurable theory and its research progress [J]. Computer Measurement and Control, 2007, 15(11): 1415-1418. (in Chinese)
- [12] 李旭冬, 叶茂, 李涛. 基于卷积神经网络的目标检测研究综述[J]. 计算机应用研究, 2017, 34(10): 2881-2886.
LI Xudong, YE Mao, LI Tao. Overview of target detection based on convolutional neural network [J]. Computer Application Research, 2017, 34(10): 2881-2886. (in Chinese)
- [13] LECUN Y, BOTTOU L. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [14] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313: 504-507.
- [15] 高远, 闫妍, 周磊, 等. 基于卷积神经网络和多项式混沌方法的翼型鲁棒性优化[J]. 航空工程进展, 2021, 12(2): 80-87.
GAO Yuan, YAN Yan, ZHOU Lei, et al. Airfoil robust optimization based on convolution neural network and polynomial chaos method [J]. Advanced in Aeronautical Science and Engineering, 2021, 12(2): 80-87. (in Chinese)

(编辑: 丛艳娟)