

改进神经网络的图像识别系统设计与硬件实现

魏 东^{*①②} 董博晨^① 刘亦青^③

^①(北京建筑大学电气与信息工程学院 北京 100044)

^②(北京市科学技术委员会建筑大数据智能处理方法研究北京市重点实验室 北京 100044)

^③(北京亚鼎智能技术有限公司 北京 100071)

摘 要: 针对现有图像识别系统大多采用软件实现,无法利用神经网络并行计算能力的问题。该文提出一套基于FPGA的改进RBF神经网络硬件化图像识别系统,将乘法运算改为加法运算解决了神经网络计算复杂不利于硬件化的问题,并且提出一种基于位比较的排序电路解决了大量数据的快速排序问题,以此为基础开发了多目标图像识别应用系统。系统特征提取部分采用FPGA实现,图像识别部分采用ASIC电路实现。实验结果表明,该文所提出的改进RBF神经网络算法平均识别时间较LeNet-5, AlexNet和VGG16缩短50%;所开发的硬件系统完成对10000张样本图片识别的时间为165 μ s,对比于DSP芯片系统所需426.6 μ s,减少了60%左右。

关键词: FPGA; ASIC电路; RBF神经网络; 图像识别系统

中图分类号: TN911.73; TP183

文献标识码: A

文章编号: 1009-5896(2021)07-1828-06

DOI: 10.11999/JEIT200202

Design and Hardware Implementation of Image Recognition System Based on Improved Neural Network

WEI Dong^{*①②} DONG Bochen^① LIU Yiqing^③

^①(School of Electrical and Information Engineering, Beijing University of Civil Engineering and Architecture, Beijing 100044, China)

^②(Beijing Key Laboratory of Intelligent Processing for Building Big Data, Beijing Municipal Science and Technology Commission, Beijing 100044, China)

^③(Beijing Yading Intelligent Technology Limited Company, Beijing 100071, China)

Abstract: To solve the problem that most existing image recognition systems are implemented in software which can not utilize the parallel computing power of neural networks, this paper proposes a FPGA image recognition system based on improved RBF neural network hardware. The multiplication operation in the neural networks is complex and inconvenient for hardware implementation. Furthermore, a sort circuit based on bit comparison is designed to solve the problem of fast sorting of a large number of data. Then, a multi-target image recognition application system is developed. The feature extraction part in the developed system is implemented by FPGA, and the image recognition part is implemented by ASIC circuit. The experimental results show that the average recognition time of the improved RBF neural network algorithm proposed is 50% shorter than that of LeNet-5, AlexNet and VGG16, and the time for the developed hardware system to recognize 10000 sample pictures is 165 μ s, which is reduced by about 60% compared with 426.6 μ s required by a DSP chip system.

Key words: FPGA; ASIC circuit; RBF neural networks; Image recognition system

收稿日期: 2020-03-24; 改回日期: 2020-09-23; 网络出版: 2020-12-09

*通信作者: 魏东 weidong@bucea.edu.cn

基金项目: 北京市属高校高水平创新团队建设计划项目(IDHT20190506), 国家自然科学基金(61871020), 北京市教委科技计划重点项目(KZ201810016019)

Foundation Items: The High Level Innovation Team Construction Project of Beijing Municipal Universities (IDHT20190506), The National Natural Science Foundation of China (61871020), The Key Science and Technology Plan Project of Beijing Municipal Education Commission of China (KZ201810016019)

1 引言

神经网络已应用于多个领域,例如图像识别、语音识别、数据挖掘、智能机器人等。神经网络利用自身强大的学习能力,通过调整神经元之间的连接权重来拟合复杂的高维映射关系^[1]。目前,在图像识别中主要采用以下几种深度学习模型进行训练,包括深层置信网络(Deep Belief Network, DBN)^[2]、卷积神经网络(Convolutional Neural Network, CNN)、循环神经网络(Recurrent Neural Network, RNN)^[3]、生成式对抗网络(Generative Adversarial Network, GAN)^[4]以及胶囊网络(Capsule Network, CapsNet)^[5]。

尽管神经网络在功能上可以满足图像识别的需要,但是目前图像识别大多采用软件实现,无法体现神经网络并行处理的优点,并且CPU或GPU在计算能力上有限,限制了神经网络结构的复杂性^[6]。为解决上述问题,本文提出一种硬件资源占用率低的改进RBF神经网络算法,并采用硬件化的改进RBF神经网络实现了图像特征提取和图像识别,其中特征提取采用FPGA实现,图像识别采用ASIC电路实现。RBF神经网络具有基函数运算复杂、中间层到输出层的分类运算量大的缺点,本文将RCE(Reilly, Cooper和Elbaum)算法和K最近邻(K-Nearest Neighbor, KNN)算法思想引入到RBF网络中,将乘法运算改为加法运算,以减低网络的运算量,便于硬件实现,同时采用最小均方算法(Least Mean Square, LMS)对权值进行调整,保证了网络的收敛性。实验结果表明改进RBF神经网络算法在不同数据集中识别准确率均高于LeNet-5和AlexNet,并与VGG16准确率持平,而平均识别时间较3种网络模型均缩短50%。本文提出的系统与基于DSP芯片的图像识别系统相比,在同为1024张相同数据集的前提下,识别时间可以缩短60%。

2 多目标图像识别系统设计方案

2.1 图像识别系统架构

传统图像识别技术中的图像预处理是对图像的特征进行增强,根据特征差异对图像进行分割,将要识别的目标从整个图像中分离,再针对分离出的目标进行特征提取^[7]。与传统技术不同的是,本文采用基于FPGA的图像扫描引擎代替传统结构中图像预处理等步骤,对图像进行全局特征提取,特征识别采用ASIC电路实现,系统结构如图1所示。

系统每次处理的数据量为

$$D = 256N, N \in [0, +\infty) \quad (1)$$

其中, D 表示单次处理的数据量,单位为Byte; N 表示系统中的样本数量。

2.2 图像识别系统业务流程

识别系统包含学习和识别两个部分,其中系统学习部分主要包括图像预处理以及学习环境设置。当一张图片输入进学习系统中首先被转换成Bitmap单色灰度图,之后需要手动截取取景范围。学习系统支持对图片进行截取、放大、缩小、平移等操作,详细描述如图2所示。

系统识别部分主要包括图像输入、图像预处理和系统识别参数设定。用户在图中确定识别目标后,可对识别结果进行逻辑组合,使得识别结果更加精准。

3 改进RBF神经网络算法研究

3.1 改进RBF神经网络模型结构

RBF, RCE和KNN网络都具有描述样本原型的特征向量,输入层将新数据与样本原型特征向量进行差异值计算,然后输出层根据差异值按照一定规则输出类别。3种网络的区别为外部数据和样本原型向量差异值计算方式不同,RBF神经网络是叠加外部数据和原型样本间的差异值^[8],RCE算法在中间层计算曼哈顿距离,KNN算法依靠数据点周围临近样本进行类别划分。本文将RCE算法和KNN算法引入RBF神经网络,将乘法运算改为加法运算,改进后网络结构如图3所示,其中外部输入矢量用 V 表示,样本矢量用 S 表示。

输入层到中间层的连接权值为 W_{ij} ,令 $W_{ij} = S_j$ 。外部矢量 V 进入网络后,进行如下计算:

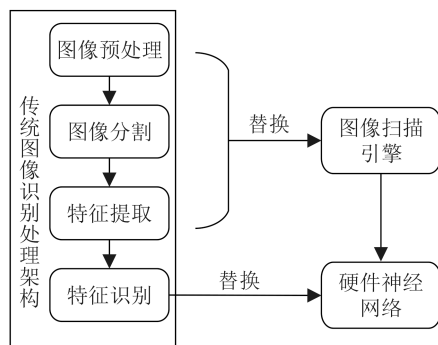


图1 基于改进RBF神经网络电路的图像识别系统架构

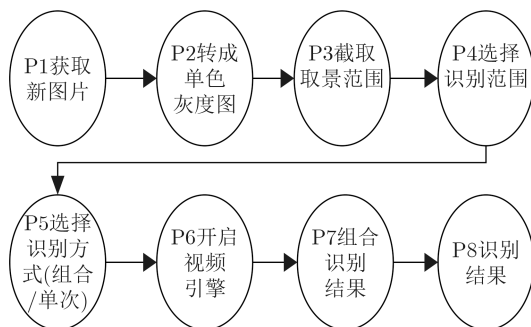


图2 多目标识别系统学习流程

$$\lambda_j \mathbf{W}_{ij} \mathbf{V} = \lambda_j \mathbf{W}_{ij} \mathbf{V} \quad (2)$$

式(2)计算输入向量与样本原型向量之间的相似程度, 如果 \mathbf{V} 和某个样本的相似度足够高, 则令

$$\lambda_j \mathbf{S}_j \mathbf{V} = \lambda_j, \lambda_j > 1 \quad (3)$$

当 $\lambda_j \leq 1$ 时, 中间层第 j 个神经元就会处于激活状态。

网络整体输入输出关系为

$$\mathbf{y}(k) = f(g(\mathbf{x}(k), \mathbf{c}, \sigma), \mathbf{W}) \quad (4)$$

其中, $\mathbf{x}(k)$, \mathbf{c} , $g(\cdot)$, \mathbf{W} 和 $f(\cdot)$ 分别为外部输入矢量、样本向量矢量、中间层激活函数、连接权值矩阵和输出层激活函数。

改进RBF神经网络和RBF神经网络结构相似, 输入层和中间层为全连接, 中间层的输入输出关系为

$$\text{Dist} = (|x_1 - c_{j1}| + |x_2 - c_{j2}|) \cdot g(|x_1 - c_{j1}| + |x_2 - c_{j2}| - \sigma_j) \quad (5)$$

$$t_j = d_j \cdot g(|x_1 - c_{j1}| + |x_2 - c_{j2}| - \sigma_j) \quad (6)$$

$$o_j = [t_j - \text{Dist}_j] \quad (7)$$

其中, o_j , x_1 , x_2 , c_{j1} , c_{j2} , σ_j , d_j , Dist_j 分别代表中间层神经元的输出、第1维度值、第2维度值、样本向量第1维度值、样本向量第2维度值、响应阈值、函数值和曼哈顿距离。

神经元的输出如式(8), 式(9)和式(10)所示:

$$y = \text{rank}(o_1, o_2, \dots, o_j, \dots) \quad (8)$$

$$o_j = [t_j, \text{Dist}_j] \quad (9)$$

$$y = \text{rank}([t_1, \text{Dist}_1], [t_2, \text{Dist}_2], \dots, [t_j, \text{Dist}_j]) \quad (10)$$

其中, $\text{rank}(\cdot)$ ——升序排队函数, 根据 Dist 的大小将输入 o 从小到大依次排序。

3.2 改进RBF神经网络学习模型

网络的学习使用监督学习的方法, 必须具备一组训练样本数据对集, 训练向量 $\mathbf{X}(k)$ 和 $\mathbf{X}(k)$ 对应的目标值 $d(k)$ 组成训练样本数据。将 $\mathbf{X}(k)$ 和 $d(k)$ 输入到系统中, 通过调整参数使得目标和输出值之间的误差 e 达到某一阈值, 如图4所示。

网络中用 $\mathbf{X}(k)$ 作为中间层的样本向量, 通过调整基宽度 δ_j 和中连接权值 w_j , 使网络的输出 y 趋近于目标 d , 使误差达到设定值。基宽度 δ_j 的调整近似于RBF神经网络中响应域 σ_j 的调整, 基宽度 δ_j 的调整规则是通过调整基宽度 δ_j 使样本点相邻的神经元之间互不干扰。因此, 网络中响应域 σ_j 的调整方式为

$$\sigma_j = \min(|c_{i1} - c_{j1}| + |c_{i1} - c_{j2}|), i \neq j, i = 0, 1, 2, \dots \quad (11)$$

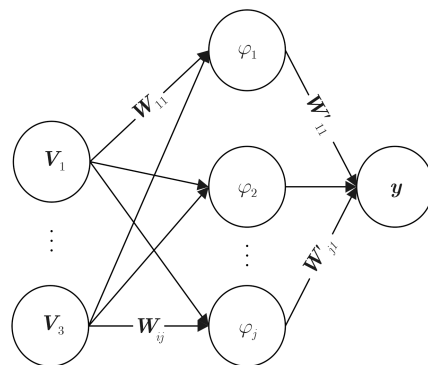


图3 改进RBF神经网络模型

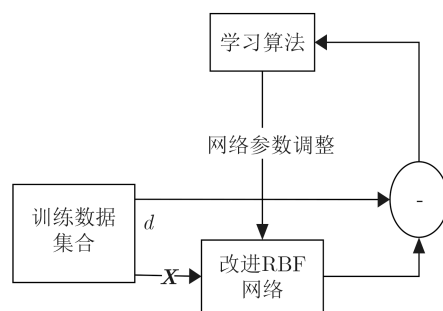


图4 改进RBF神经网络学习模型

权值 w_j 则采用最小均方算法(LMS)进行调整, 即

$$w_j(k+1) = w_j(k) + \alpha \cdot [d(k) - y(k)] \cdot g\left(\sum_n |X - C_j| - \sigma_j\right) \quad (12)$$

网络关系为

$$y = \sum_j w_j \cdot g\left(\sum_n |X - C_j| - \sigma_j\right) \quad (13)$$

将式(12)代入式(13)得

$$\begin{aligned} w_j(k+1) &= w_j(k) + \alpha \cdot \left[d(k) - \sum_j w_j(k) \cdot g\left(\sum_n |X - C_j| - \sigma_j\right) \right] \cdot g\left(\sum_n |X - C_j| - \sigma_j\right) \end{aligned} \quad (14)$$

3.3 改进RBF神经网络识别算法

当外部向量进入网络后, 首先将其与每个神经元保存的样本向量计算曼哈顿距离, 将得到的距离值与神经元 σ_j 做比较, 如果距离值小于 σ_j , 则神经元输出其类别值。

定义一个训练样本集合 $\mathbf{D} = \{1\mathbf{P}, 2\mathbf{P}, \dots, k\mathbf{P}\}$, 其中:

$$k\mathbf{P} = [k\mathbf{X}, k\mathbf{TCat}] \quad (15)$$

$k\mathbf{X}$ 表示第 k 个外部输入训练向量; $k\mathbf{TCat}$ 表示第 k 个训练向量对应的类别值

网络整体函数关系为

$$y = f(\mathbf{O}) = f(\varphi(k\mathbf{X})) \quad (16)$$

$$f(\mathbf{O}) = \text{rank}(\mathbf{O}) \quad (17)$$

神经元函数关系为

$$(\mathbf{O}) = \varphi(\mathbf{X}) = \begin{bmatrix} \mathbf{O}_1 \\ \vdots \\ \mathbf{O}_J \\ \vdots \\ \mathbf{O}_N \end{bmatrix} \quad (18)$$

神经网络识别结果为

$$y = \text{rank}(\mathbf{o}_1, \mathbf{o}_2, \dots, \mathbf{o}_j, \dots, \mathbf{o}_n) \quad (19)$$

3.4 网络收敛性证明

当输入 $\mathbf{X} \neq \mathbf{C}_j$ 时,

$$\varphi(\mathbf{X}, \mathbf{C}_j, \sigma_j) = g\left(\sum_n |C_i - C_j| - \sigma_j\right) = 0 \quad (20)$$

将式(20)代入式(14)中, 得

$$w_j(k+1) = w_j(k) \quad (21)$$

当输入 $\mathbf{X} = \mathbf{C}_j$ 时,

$$\varphi(\mathbf{X}, \mathbf{C}_j, \sigma_j) = g\left(\sum_n |C_i - C_j| - \sigma_j\right) = 1 \quad (22)$$

将式(22)代入式(14)中, 得

$$w_j(k+1) = w_j(k) + \alpha \cdot [d(k) - w_j(k)] \quad (23)$$

当输入等于 \mathbf{C}_j 时, w_j 进行调整, 此时的 $d(k)$ 就是将 \mathbf{C}_j 作为输入, 所以可以表示为

$$w_j(k) = F(\mathbf{C}_j) = d_{C_j} \quad (24)$$

4 改进RBF神经网络硬件设计

4.1 硬件化需要解决的问题

对改进RBF神经网络硬件化时, 首先要决定样本相似性计算的方法和如何对输出层进行快速排序, 本文在相似性计算中采用曼哈顿距离, 而对于结果的快速排序问题本文提出一种基于位比较的快速排序电路。

4.1.1 相似性计算

外部向量进入网络后与网络内部的原始样本向量进行相似性计算, RBF神经网络计算欧几里得距离, 针对欧几里得距离运算量大的问题, 本文用曼哈顿距离替代欧几里得距离进行相似性计算, 可表示为

$$d_{jk} = \sum_{i=1}^n |(x_i^j - x_i^k)|, \quad n \in N \quad (25)$$

其中, d_{jk} , i , x_i^j , x_i^k 分别表示为向量 \mathbf{X}^j , \mathbf{X}^k 之间的

曼哈顿距离、向量 \mathbf{X} 的维度、向量 \mathbf{X}^j 的第 i 个维度值和向量 \mathbf{X}^k 的第 i 个维度值。

4.1.2 输出排序计算

网络计算外部输入和原始样本向量的曼哈顿距离后, 输出层要对大量距离值进行排序输出, 并使用经典的冒泡法等进行排序。当数据过多时, 典型的排序算法在实现时存在资源占用大、运算时间久的问题。

改进RBF网络采用的排序输出方式为

$$y = \text{rank}(w_{11} \cdot \text{Dist}_1, w_{21} \cdot \text{Dist}_2, \dots, w_{j1} \cdot \text{Dist}_j, \dots) \quad (26)$$

其中, $\text{rank}(\cdot)$ 表示数组从小到大排序; w_{j1} 表示权值; Dist_j 表示中间层第 j 个神经元和输入 x 的距离。

本文采用图5所示电路来实现对无限值的迅速排序, 首先在寄存器RX中存储距离值(X 表示1, 2, 3, ...), 将 R 中的数据经过处理 P 传递给寄存器 RX^1 , 之后从 RX^1 中低位开始读取数据到寄存器 R , 当 R 中任意一个位有1出现时停止取数, 出现1的位数所对应的距离即为最小距离。

4.2 改进RBF神经网络FPGA实现

4.2.1 算法状态机

FPGA实现改进RBF神经网络算法最重要的准备工作是将算法转换为算法状态机(ASM)。本文设计了一种新的系统算法状态机, 如图6所示。

图6可以看出, 系统包括6种状态, 系统运行之后会永远处于6个状态之一, 符合硬件电路工作特性。

4.2.2 模块分割

针对在FPGA实现过程中容易丢失设计功能的问题, 本系统采用模块分割的方式对系统模块自上而下进行描述。系统模块包括控制单元和数据路径, 数据通过控制单元向网络发出指令, 网络通过控制总线

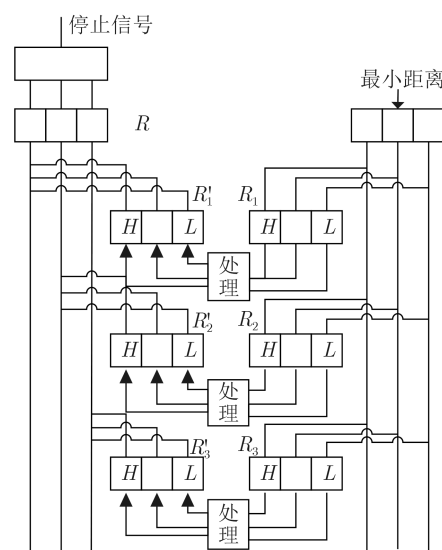


图5 快速排序电路

将命令传送给神经元，神经元接收到信号后从数据总线接收数据进行处理，将处理结果送回数据总线。

5 实验研究

5.1 改进RBF神经网络图像识别算法对比实验

为验证改进RBF神经网络图像识别算法的性能，本文分别在MNIST^[9]、CIFAR-10^[10]和VOC2012^[11]数据集上与LeNet-5、AlexNet和VGG16网络模型进行准确率及平均识别时间的对比实验。实验中采用Tensorflow作为深度学习框架，硬件环境包括一台CPU为I7-8750的笔记本电脑，其主频为2.20 GHz内存为8GB，显卡为NVIDIA GTX 1050TI；软件采用Visual Studio 2017，并使用Python语言进行实现，实验结果如表1—表3所示。

从上述表格中可以看出，在3种不同数据集中VGG16网络识别准确率最高，但是由于网络自身参数量大网络平均识别时间最长；而LeNet-5和AlexNet网络平均识别时间较VGG16网络缩短2~3 s，但是准确率较VGG16网络降低3%~6%。而本文提出的改进RBF神经网络算法在不同数据集中准确率优于LeNet-5和AlexNet，并且与VGG16网络持平，但是在平均识别时间上缩短了50%。因此本文提出的改进RBF神经网络算法在识别精度以及识别时间上均有一定的提升。

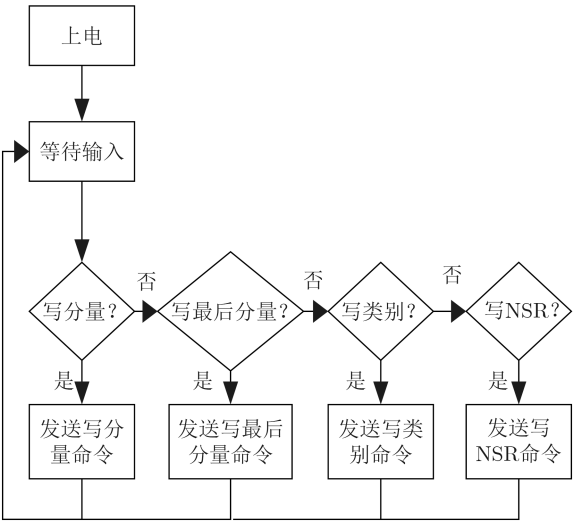


图 6 系统算法状态机

表 1 测试数据集为MNIST时不同网络模型对比实验

网络模型	准确率	平均识别时间(s)
LeNet-5	0.989	1.2
AlexNet	0.991	1.5
VGG16	0.997	2.3
改进RBF神经网络	0.996	0.9

5.2 改进RBF神经网络图像识别系统与传统识别技术比较

本文针对视频检索任务，采用ASIC电路实现系统。并与基于DSP芯片的图像识别系统进行对比试验，设定对比实验的环境为：(1)样本库中包含10000个256 Byte大小的样本；(2)外部输入一个256 Byte大小的数据；(3)目标是计算该外部数据和样本库中每个样本的差值；(4)输出结果包含样本名和差值；(5)要求输出结果找出差异值最小的样本以及相应的差值。

设定上述环境后，基于DSP图像识别系统和本文提出系统的各项工作参数计算结果如表4所示。

从表4可以看出，本文所开发的系统算术逻辑部件数量(ALU)达到1024个，相比较于基于DSP芯片的系统ALU数量有明显提升。由于ALU数量的提升，系统每个周期进行加法运算的次数达到2048次，计算的数据量达到2048 Byte。系统完成所有样本的时间为165 μ s，较DSP芯片减少了60%左右。

表 2 测试数据集为CIFAR-10时不同网络模型对比实验

网络模型	准确率	平均识别时间(s)
LeNet-5	0.787	2.4
AlexNet	0.810	3.3
VGG16	0.832	5.5
改进RBF神经网络	0.828	1.3

表 3 测试数据集为VOC2012时不同网络模型对比实验

网络模型	准确率	平均识别时间(s)
LeNet-5	0.757	2.7
AlexNet	0.783	3.5
VGG16	0.813	6.7
改进RBF神经网络	0.808	2.6

表 4 基于DSP图像识别系统与本文提出图像识别系统性能比较

计算量	DSP芯片	本文系统
时钟频率(MHz)	500	15
ALU数量	6	1024
运算位宽(bit)	32	8
单个样本大小(Byte)	256	256
每个周期能进行加法次数	6	$1024 \times 2 = 2048$
每次加法处理数据(Byte)	$(32/8) \times 6 = 24$	$(8/8) \times 2048 = 2048$
完成两个样本的比较需要周期数	$(2 \times 256 / 24) = 21.33$	$(2 \times 256 / 2048) = 0.25$
完成两个样本比较的时间(ns)	$21.33 \times 2 = 42.66$	$0.25 \times 66 = 16.5$
和所有样本比较所需时间(μ s)	426.6	165

6 结论

本文提出一套基于FPGA的改进RBF神经网络图像识别系统,其中算法引入了RCE和KNN算法思想将乘法运算改为加法运算。算法部分与LeNet-5、AlexNet和VGG16网络模型进行对比实验,实验结果证明本文所提出算法在识别准确率优于LeNet-5和AlexNet网络,并与VGG16网络准确率持平,而平均识别时间比3种网络模型缩短了50%;硬件部分采用FPGA实现特征提取过程,图像识别过程采用ASIC电路实现,最终实现图像识别功能。通过与基于DSP芯片的图像识别系统进行比较,本文提出的系统在同为1024张图片的前提下,识别时间可缩短60%左右。

今后研究的方向有以下几个:

(1) 改进RBF神经网络虽然简化了网络结构,但是网络内部神经元数量众多,需要进一步降低神经元资源的消耗,即如何在保证识别准确的条件下让神经元用量达到最小。

(2) 提高神经元计算能力,随着集成电路工艺的不断提升,神经元可以拥有更高量级的运算位宽,这对于神经网络硬件化性能的提升有着不可小觑的作用。

(3) 目前图像识别算法发展快、种类多,目前没有明确的硬性指标来评估算法的优劣。未来是否可以在设计方案相同的情况下,通过指标对算法硬件实现的复杂度进行评估,还需要进一步研究。

总之,利用FPGA进行神经网络算法硬件化还有巨大的上升空间,对于网络结构的优化、编译器设计具有很高的研究价值。

参 考 文 献

- [1] 李国良,周煊赫,孙估,等. 基于机器学习的数据库技术综述[J]. 计算机学报, 2020, 43(11): 2019-2049.
LI Guoliang, ZHOU Xuanhe, SUN Ji, et al. A survey of machine learning based database techniques[J]. *Chinese Journal of Computers*, 2020, 43(11): 2019-2049.
- [2] 刘方园,王水花,张煜东. 深度置信网络模型及应用研究综述[J]. 计算机工程与应用, 2018, 54(1): 11-18, 47. doi: 10.3778/j.issn.1002-8331.1711-0028.
LIU Fangyuan, WANG Shuihua, and ZHANG Yudong. Review of deep confidence network model and application research[J]. *Computer Engineering and Applications*, 2018, 54(1): 11-18, 47. doi: 10.3778/j.issn.1002-8331.1711-0028.
- [3] LIANG Tian and AFZEL N. Software reliability prediction using recurrent neural network with Bayesian regularization[J]. *International Journal of Neural Systems*, 2004, 14(3): 165-174. doi: 10.1142/S0129065704001966.
- [4] GOODFELLOW I J, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]. Proceedings of the 27th International Conference on Neural Information Processing Systems, Montreal, Canada, 2014: 2672-2680.
- [5] SABOUR S, FROSST N, and HINTON G E. Dynamic routing between capsules[C]. Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, USA, 2017: 3856-3866.
- [6] 任源,潘俊,刘京京,等. 人工智能芯片的研究进展[J]. 微纳电子与智能制造, 2019, 1(2): 20-34.
REN Yuan, PAN Jun, LIU Jingjing, et al. Overview of artificial intelligence chip development[J]. *Micro/Nano Electronics and Intelligent Manufacturing*, 2019, 1(2): 20-34.
- [7] 秦华标,曹钦平. 基于FPGA的卷积神经网络硬件加速器设计[J]. 电子与信息学报, 2019, 41(11): 2599-2605. doi: 10.11999/JEIT190058.
QIN Huabiao and CAO Qinpeng. Design of convolutional neural networks hardware acceleration based on FPGA[J]. *Journal of Electronics & Information Technology*, 2019, 41(11): 2599-2605. doi: 10.11999/JEIT190058.
- [8] 韩栋,周聖元,支天,等. 智能芯片的评述和展望[J]. 计算机研究与发展, 2019, 56(1): 7-22. doi: 10.7544/issn1000-1239.2019.20180693.
HAN Dong, ZHOU Shengyuan, ZHI Tian, et al. A survey of artificial intelligence chip[J]. *Journal of Computer Research and Development*, 2019, 56(1): 7-22. doi: 10.7544/issn1000-1239.2019.20180693.
- [9] 王巍,周凯利,王伊昌,等. 基于快速滤波算法的卷积神经网络加速器设计[J]. 电子与信息学报, 2019, 41(11): 2578-2584. doi: 10.11999/JEIT190037.
WANG Wei, ZHOU Kaili, WANG Yichang, et al. Design of convolutional neural networks accelerator based on fast filter algorithm[J]. *Journal of Electronics & Information Technology*, 2019, 41(11): 2578-2584. doi: 10.11999/JEIT190037.
- [10] 伍家松,达臻,魏黎明,等. 基于分裂基-2/(2a)FFT算法的卷积神经网络加速性能的研究[J]. 电子与信息学报, 2017, 39(2): 285-292. doi: 10.11999/JEIT160357.
WU Jiasong, DA Zhen, WEI Liming, et al. Acceleration performance study of convolutional neural network based on split-radix-2/(2a) FFT algorithms[J]. *Journal of Electronics & Information Technology*, 2017, 39(2): 285-292. doi: 10.11999/JEIT160357.
- [11] 张烨,许艇,冯定忠,等. 基于难分样本挖掘的快速区域卷积神经网络目标检测研究[J]. 电子与信息学报, 2019, 41(6): 1496-1502. doi: 10.11999/JEIT180702.
ZHANG Ye, XU Ting, FENG Dingzhong, et al. Research on faster RCNN object detection based on hard example mining[J]. *Journal of Electronics & Information Technology*, 2019, 41(6): 1496-1502. doi: 10.11999/JEIT180702.

魏 东:女,1968年生,教授,研究方向为人工神经网络优化计算。
董博晨:男,1995年生,硕士,研究方向为神经网络芯片设计。
刘亦青:男,1987年生,硕士,研究方向为控制科学与工程。

责任编辑:陈 倩