

# 深度学习相关研究综述 \*

张军阳, 王慧丽, 郭 阳, 扈 啸

(国防科学技术大学 计算机学院, 长沙 410073)

**摘 要:** 随着大数据时代的到来, 深度学习技术已经成为当前人工智能领域的一个研究热点, 其已在图像识别、语音识别、自然语言处理、搜索推荐等领域展现出了巨大的优势, 并且仍在继续发展变化。为了能够及时跟踪深度学习技术的最新研究进展, 把握深度学习技术当前的研究热点和方向, 本文针对深度学习技术的相关研究内容进行综述。首先介绍了深度学习技术的应用背景、应用领域, 指出研究深度学习技术的重要性; 其次介绍了当前重要的几种神经网络模型及两种常用大规模模型训练并行方案, 其目的在于从本质上理解深度学习的模型架构和及其优化技巧; 接着对比分析了当下主流的深度学习软件工具和相关的工业界研究平台, 旨在为神经网络模型的实际使用提供借鉴; 最后详细介绍了当下几种主流的深度学习硬件加速技术和最新研究现状, 并对未来研究方向进行了展望。

**关键词:** 深度学习; 神经网络; 算法模型; 软件工具; 硬件加速

**中图分类号:** TP181

## Review of deep learning

Zhang Junyang, Wang Huili, Guo Yang, Hu Xiao

(College of Computer, National University of Defense Technology, Changsha 410073, China)

**Abstract:** With the era of big data coming, deep learning technology has become a hotspot research in the field of artificial intelligence. It has shown great advantages in image recognition, speech recognition, natural language processing, search recommendation and so on, while still continues to evolve. In order to keep track of the latest research progress of deep learning technology and grasp the current research hotspot and direction of deep learning, this paper reviews the related research contents of deep learning technology. Firstly, it introduces the application background and application field of deep learning technology and points out the importance of studying on deep learning technology; Secondly, it introduces several important neural network models and two kinds of commonly used large-scale model training parallel scheme, which aims to understand the deep learning model structure and its optimization skills; Then it analyzes the current mainstream learning tools and related industrial research platform, which aims to provide reference for the practical use of neural network model; At the end of this paper, the hardware acceleration technology and the latest research status of several kinds of deep learning hardware acceleration are introduced in detail, and the future research directions are also discussed.

**Key Words:** deep learning; neural network; algorithm model; software tools; hardware acceleration

## 0 引言

随着传感器技术、存储技术、计算机技术和网络技术的迅猛发展以及人类管理与知识水平的提高, 使得数据的膨胀趋势日益加剧, 信息技术发展的瓶颈已不仅仅存在于数据的获取、存储与传输, 而更受限与数据的加工、分析和利用。采用有效的人工智能技术从大数据中获取抽象信息并将其转换为有用的知识, 是当前大数据分析所面临的核心问题之一。大数据时代, 如何对纷繁复杂的数据进行有效分析, 让其价值得以体现和合

理的利用, 是当前迫切需要思考和解决的问题。而近期兴起的深度学习方法正是开启这扇大门的一把钥匙。深度学习是新兴的机器学习研究领域, 旨在研究如何从数据中自动地提取多层特征表示, 其核心思想是通过数据驱动的方式, 采用一系列的非线性变换, 从原始数据中提取由低层到高层、由具体到抽象、由一般到特定语义的特征。深度学习不仅改变着传统的机器学习方法, 也影响着本文对人类感知的理解, 迄今已在语音识别、图像理解、自然语言处理、视频推荐等应用领域引发了突破性的变革。

**基金项目:** 国家自然科学基金资助项目 (61572025); 国家重点研发计划资助项目 (2016YFB0200401)

**作者简介:** 张军阳 (1987-), 男, 河南平顶山人, 博士研究生, 主要研究方向为机器学习、算法优化、硬件加速、并行计算等 (zhangjunyang11@nudt.edu.cn); 王慧丽 (1986-), 博士研究生, 主要研究方向为深度学习、硬件加速器设计; 郭阳 (1971-), 男, 研究员, 博士, 博导, 主要研究方向为计算机体系结构、微处理器设计与验证; 扈啸 (1972-), 男, 副研究员, 博士, 主要研究方向为嵌入式系统。

鉴于深度学习技术的发展日新月异,国内基于深度学习的相关研究仍处于起步阶段,缺少系统的针对深度学习各方面的介绍,与当前基于深度学习的综述性文章相比,本文系统的从研究背景、应用领域、算法模型、优化方法、软件工具、硬件加速和总结展望等若干层次对当前深度学习的相关研究进行综述,为进一步深入研究深度学习理论和拓展其应用范围奠定了基础。

## 1 深度学习相关应用领域

### 1.1 图像识别

物体检测和图像分类是图像识别的两个核心问题,前者主要定位图像中特定物体出现的区域并判定其类别,后者则对图像整体的语义内容进行类别判定。Yang 等人<sup>[1]</sup>是传统图像识别算法中的代表,他们在 2009 年提出的采用稀疏编码来表征图像,通过大规模数据来训练支持向量机(support vector machine,SVM)进行图像分类,该方法在 2010 年和 2011 年的 ImageNet<sup>[2]</sup>图像分类竞赛中取得了最好成绩。图像识别是深度学习最早尝试的应用领域,早在 1989 年,LeCun 和他的同事就发表了关于卷积神经网络的相关工作<sup>[3]</sup>,在手写数字识别任务上取得了当时世界上最好的结果,并广泛应用于各大银行支票的手写数字识别任务中。百度在 2012 年将深度学习技术成功应用于自然图像 OCR 识别和人脸识别等问题上,并推出相应的移动搜索产品和桌面应用。从 2012 年的 ImageNet 竞赛开始,深度学习在图像识别领域发挥出巨大威力,在通用图像分类、图像检测、光学字符识别(optical character recognition,OCR)、人脸识别等领域,最好的系统都是基于深度学习的。图 1 为从 2010 到 2016 年 ImageNet 竞赛的识别错误率变化及人的识别错误率。2012 年是深度学习技术第一次被应用到 ImageNet 竞赛中,可以看出相对于 2011 年传统最好的识别错误率大幅降低了 41.1%,且 2015 年基于深度学习技术的图像识别率错误率已经超过了人类,2016 年最新的 ImageNet 识别错误率已经达到 2.991%。

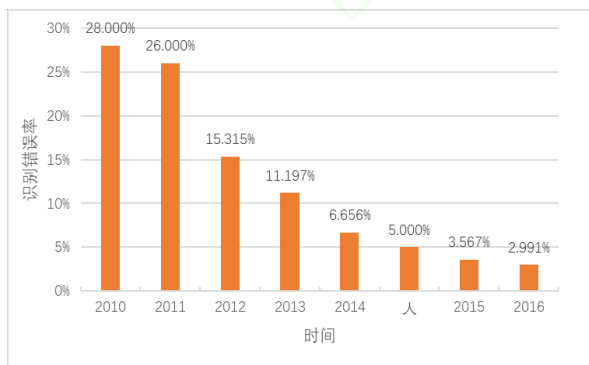


图 1 2010 至 2016 年 ImageNet 竞赛的识别错误率变化及人的识别错误率

### 1.2 语音识别

长久以来,人与机器交谈一直是人机交互领域内的一个梦想,而语音识别是其基本技术。语音识别(automatic speech

recognition,ASR)是指能够让计算机自动地识别语音中所携带信息的技术。语音是人类实现信息交互最直接、最便捷、最自然的方式之一。自人工智能(artificial intelligence,AI)的概念出现以来,让计算机甚至机器人像自然人一样实现利用语音进行交互就一直是 AI 领域研究者的梦想。

最近几年,深度学习(deep learning,DL)理论在语音识别和图像识别领域取得了令人振奋的性能提升,迅速成为了当下学术界和产业界的研究热点,为处在瓶颈期的语音等模式识别领域提供了一个强有力的工具。在语音识别领域,深度神经网络(deep neural network,DNN)模型给处在瓶颈阶段的传统的 GMM-HMM 模型带来了巨大的革新,使得语音识别的准确率又上了一个新的台阶。目前国内外知名互联网企业(谷歌、科大讯飞及百度等)的语音识别算法都采用的是 DNN 方法。2012 年 11 月,微软在中国天津的一次活动上公开演示了一个全自动的同声传译系统,讲演者用英文演讲,后台的计算机一气呵成自动完成语音识别、英中机器翻译和中文语音合成,效果非常流畅,其后台支撑的关键技术就是深度学习。近期,百度将 Deep CNN 应用于语音识别研究,使用了 VGGNet,以及包含 Residual 连接的深层卷积神经网络(convolutional neural network,CNN)等结构,并将长短期记忆网络(long short-term memory,LSTM)和 CTC 的端到端语音识别技术相结合,使得识别错误率相对下降了 10%以上。2016 年 9 月,微软的研究者在产业标准 Switchboard 语音识别任务上,取得了产业中最低的 6.3%的词错率。以及国内科大讯飞提出的前馈型序列记忆网络(feed-forward sequential memory network,FSMN)的语音识别系统,该系统使用大量的卷积层直接对整句语音信号进行建模,更好的表达了语音的长时相关性,其效果比学术界和工业界最好的双向 RNN(recurrent neural network,RNN)语音识别系统识别率提升了 15%以上。由此可见,深度学习技术对语言识别率的提高有着不可忽略的贡献。

### 1.3 自然语言处理

自然语言处理(natural language processing,NLP)也是深度学习的一个重要应用领域,经过几十年多的发展,基于统计的模型已经成为 NLP 的主流,同时人工神经网络在 NLP 领域也受到了理论界的足够重视。加拿大蒙特利尔大学教授 Bengio 等在 2003 年提出用 embedding 的方法将词映射到一个矢量表示空间,然后用非线性神经网络来表示 N-Gram 模型<sup>[4]</sup>。世界上最早的深度学习用于 NLP 的研究工作诞生于 NEC Labs American,其研究员 Collobert 和 Weston<sup>[5]</sup>从 2008 年开始采用 embedding 和多层一维卷积的结构,用于词性标注、分块、命名实体识别、语义角色标注等 4 个典型 NLP 问题。值得注意的是,他们将同一个模型用于不同的任务,都取得了与现有技术水平相当的准确率。

Mikolov 等通过对 Bengio 等提出的神经网络语言模型的进一步研究发现,通过添加隐藏层的多次递归,可以提高语言模型的性能<sup>[6]</sup>,语音识别任务中,在提高后续词预测准确率及总

体识别错误率方面都超越了当时最好的基准系统, Schwenk 等将类似的模型用在统计机器翻译任务中<sup>[7]</sup>, 采用 BLEU (bilingual evaluation understudy, BLEU) 评分机制评判, 提高了近 2 个百分点。此外, 基于深度学习模型的特征学习还在语义消歧<sup>[8]</sup>、情感分析<sup>[9,10]</sup>等自然语言处理任务中均超越了当时最优系统, 取得优异表现。

## 2 深度学习常用模型

### 2.1 自动编码器

自动编解码网络可看做是传统的多层感知器的变种, 最早由文献[11]提出。其基本想法是将输入信号经过多层神经网络后重构原始的输入, 通过非监督学习的方式挖掘输入信号的潜在结构, 将中间层的响应作为潜在的特征表示。其基本结构如图 2 所示。

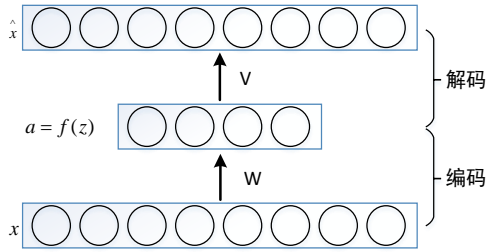


图 2 自动编码器模型结构示意图

自动编解码机由将输入信号映射到低维空间的编码器和用隐含特征重构初始输入的解码机构成。假设输入信号为  $x$ , 编码层首先将其线性映射为  $z$ , 然后再施加某种非线性变换, 这一过程可以形式化为:

$$a = f(z) = f(Wx + b) \quad (1)$$

其中:  $f(\bullet)$  为某种非线性函数, 常用的有 sigmoid 函数 ( $f(z) = 1/(1 + \exp(-z))$ ) 和修正线性单元 (rectified linear unit, ReLU) 函数 ( $f(z) = \max(0, z)$ ) 也称为激活函数。然后, 解码层通过相似的操作, 将隐含特征  $a$  映射回输入空间, 得到重构的信号  $\hat{x}$ 。自动编码器的参数即为每一层的连接权重和偏置。网络训练时的优化目标为最小化重构信号与输入信号之间的均方差:

$$\min \sum_i (\hat{x}_i - x_i)^2 \quad (2)$$

自动编码器可以通过级联和逐层训练的方式组成深层的结构, 其中只需要将前一层中隐含层的输出作为当前层的输入。深度模型通过逐层优化的方式训练后, 还可以通过让整个网络重构输入信号的原则进行精调<sup>[12]</sup>。在实际的系统中, 还经常将编码器和解码机的权重进行耦合即令  $V = W^T$ , 使得编解码的过程完全相似。在自动编码器的框架下, 很多研究者通过引入正则约束的方式, 开发了很多变种模型。一些研究者将稀疏表示的思想引入, 提出了稀疏自动编解码机<sup>[13,14]</sup>, 其中通过  $L_1$  惩罚或者鼓励输出信号的平均值与一个平均值很小的高斯分布近似来实现。为了增强自动编码器的泛化性, Vincent 等人<sup>[12]</sup>提出

了降噪自动编码器, 他们在训练之前给训练样本加入人工制造的噪声干扰, 使得网络可以从有噪声的信号中重构原始的干净输入。与之非常相似的是 Rifai 等人<sup>[15]</sup>提出的收缩自动编码器, 通过引入一个收缩惩罚项来增强模型的泛化性能, 同时降低过拟合的影响。很多研究者已经将深度自动编码器成功的应用于图像特征表示中, 文献[16]中利用深度自动编码器得到紧致的图像高层描述并基于此进行图像检索。文献[17]则训练了一个从粗到细的自动编码器, 完成人脸关键点定位的任务。

### 2.2 受限玻尔兹曼机

玻尔兹曼机 (boltzmann machine, BM) 是一种随机的递归神经网络, 由 G.E.Hinton 等<sup>[18,19,20]</sup>提出, 是能通过学习数据固有内在表示、解决复杂学习问题最早的人工神经网络之一, 受限玻尔兹曼机 (restricted boltzmann machine, RBM) 是玻尔兹曼机的扩展, 由 Hinton 等提出, 由于去掉了玻尔兹曼机同层之间的连接, 因而大大提高了学习效率。

如下图所示, RBM 是一个双向图模型, 由可视层  $v \in \{0, 1\}^{N_v}$  和隐含层  $h \in \{0, 1\}^{N_h}$  组成, 可视层和隐含层之间的联合概率分布定义为:

$$P(v, h) = \frac{1}{Z} \exp(v^T W h + V^T b v + h^T b_h) \quad (3)$$

其中  $Z$  为归一化函数,  $W \in R^{N_v \times N_h}$  表示可视层和隐含层之间的连接权重, 而  $b_v \in R^{N_v}$ ,  $b_h \in R^{N_h}$  是偏置项。此模型的优化目标和一般的概率图模型一样, 都是基于最大似然估计, 也即最小化训练数据的似然概率的负对数:

$$E(p, h) = -\log P(v, h) \quad (4)$$

如果用传统的基于 Gibbs 采样的方法求解, 则迭代次数较多效率很低, 为了克服这一问题, Hinton<sup>[19]</sup>提出了一种称为对比分歧 (contrastive divergence, CD) 的快速算法。而在文献[21]中提出了一种基于随机梯度下降法的更高效的优化算法。和稀疏编码等模型相比, RBM 模型具有一个非常好的优点, 即它的推断很快, 只需要一个简单的前向编码操作, 即

$$h = \text{sigmoid}(W \cdot v + b_h).$$

一些研究者在 RBM 基础上提出了很多扩展模型。原始的 RBM 模型中可视层为二值变量, 文献[22]中通过引入高斯核使得 RBM 支持连续变量作为输入信号。一些拓展模型修改了 RBM 的结构和概率分布模型, 使得它能模拟更加复杂的概率分布, 如 “mean-covariance RBM”<sup>[23]</sup>、“spike-slab RBM”<sup>[24]</sup>和门限 RBM<sup>[25]</sup>。这些模型中通常都定义了一个更加复杂的能量函数, 学习和推断的效率因此会有所下降。此外, 文献[26]提出在 RBM 的生成式学习算法中融入判别式学习, 使得它能更好的应用于分类等判别式任务。

通过级联多个单层的 RBM 模型可构成深层的结构, 即将前一层的隐含层作为当前层的可视层, 网络的优化采用逐层优化的方式。文献[27]中将多层的有向 Sigmoid 置信网络与 RBM



级联, 构造了一个深度信念网络 (deep belief network, DBN)。文献[28]则将 RBM 模型直接级联成多层结构, 提出了深度玻尔兹曼机网络。Lee 等人<sup>[29]</sup>用卷积操作对 DBN 网络进行扩展, 使得模型可以直接从原始的二维图像中学习潜在的特征表示。除了基于 RBM 的深度结构外, 还有其他一些层级生成式模型。Yu 等人<sup>[30]</sup>提出深度稀疏编码模型, 用于学习图像像素块的潜在结构特征。Zeiler 等人<sup>[31]</sup>通过级联多个卷积稀疏编码和最大值池化层, 构建了深度反卷积网络, 可以直接从全局图像中学习从底层到高层的层级结构特征。

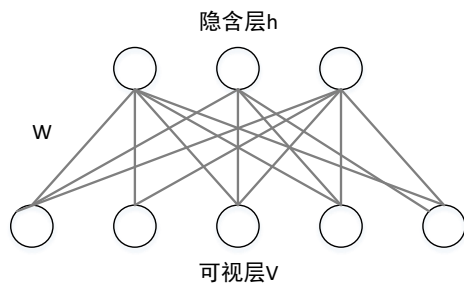


图 3 受限玻尔兹曼机

### 2.3 深度神经网络

神经网络技术起源于上世纪五、六十年代, 当时叫感知机, 是最早被设计并实现的人工神经网络, 是一种二分类的线性分类模型, 主要用于线性分类且分类能力十分有限。输入的特征向量通过隐含层变换达到输出层, 在输出层得到分类结果。早期感知机的推动者是 Rosenblatt。但是单层感知机遇到一个严重的问题, 即它对稍复杂一些的函数都无能为力 (比如最为典型的“异或”操作), 随着数学理论的发展, 这个缺点直到 20 世纪 80 年代才被 Rumelhart、Williams、Hinton、LeCun 等人发明的多层感知机 (multilayer perceptron, MLP) 克服。多层感知机可以摆脱早期离散传输函数的束缚, 使用 sigmoid 或 tanh 等连续函数模拟神经元对激励的响应, 在训练算法上则使用 Werbos 发明的反向传播 BP 算法。

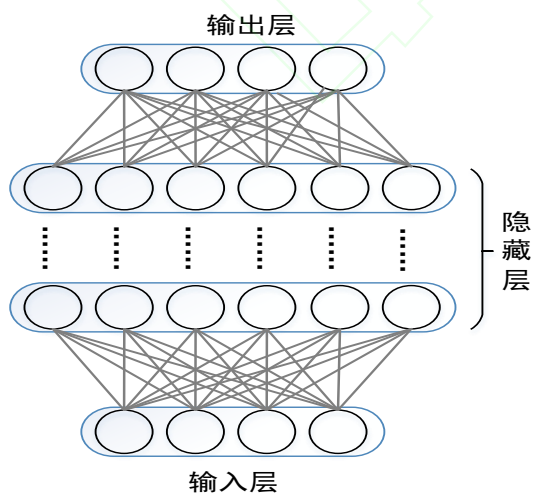


图 4 深度神经网络结构

图 4 为全连接深度神经网络结构示意图。通过增加隐含层的数量及相应的节点数, 可以形成深度神经网络, 深度神经网络一般指全连接的神经网络, 该类神经网络模型常用于图像及

语言识别等领域, 在图像识别领域由于其将图像数据变成一维数据进行处理, 忽略了图像的空间几何关系, 因此其在图像识别领域的识别率不及卷积神经网络, 且由于相邻层之间全连接, 其要训练的参数规模巨大, 因此巨大的参数量也进一步限制了全连接神经网络模型结构的深度和广度。

### 2.4 卷积神经网络

近几年, 卷积神经网络在大规模图像特征表示和分类中取得了很大的成功。标志性事件是在 2012 年的 ImageNet 大规模视觉识别挑战竞赛中<sup>[32]</sup>, Krizhevsky<sup>[33]</sup>实现的深度卷积神经网络模型将图像分类的错误率降低了近 50%。2016 年 4 月份著名的围棋人机大战中以 4:1 大比分优势战胜李世石的 AlphaGo 人工智能围棋程序就采用了 CNN+蒙特卡洛搜索树算法。卷积神经网络最早是由 leCun 等人在 1998 年提出, 用于手写字符图像的识别, 其网络结构如图 5 所示。

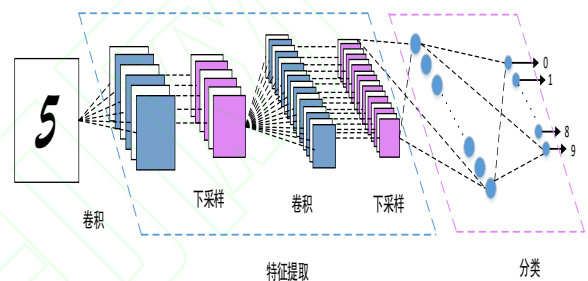


图 5 卷积神经网络

该网络的输入为原始二维图像, 经过若干卷积层和全连接层后, 输出图像在各类别下的预测概率。每个卷积层包含三种运算: 卷积、非线性激活函数和最大值池化。在卷积神经网络中, 需要学习一组二维滤波模板  $F = f_1, \dots, f_{N_k}$ , 和输入特征图  $x$  进行卷积操作, 得到  $N_k$  个二维特征图:  $z_k = f_k * x$ 。

采用卷积运算的好处有如下几点: a) 二维卷积模板可以更好的挖掘相邻像素之间的局部关系和图像的二维结构; b) 和一般神经网络中的全连接结构相比, 卷积网络通过权重共享极大的减少了网络的参数量, 使得训练大规模网络变得可行; c) 卷积操作对图像上的平移、旋转和尺度等变换具有一定的鲁棒性。得到卷积响应特征图后, 通常需要经过一个非线性激活函数来得到激活响应图, 如 sigmoid、tanh 和 ReLU 等函数。紧接着, 在激活函数响应图上施加一个最大值池化 (max pooling) 或者平均值池化 (average pooling) 运算。在这一操作中, 首先用均匀的网格将特征图划分为若干空间区域, 这些区域可以有重叠部分, 然后取每个图像区域的平均值或最大值作为输出。此外在最大值池化中, 通常还需要记录所输出最大值的位置。已有研究工作<sup>[34]</sup>证明了最大值池化操作在图像特征提取中的性能优于平均值池化, 因而近些年研究者基本都采用了最大值池化。池化操作主要有如下两个优点: a) 增强了网络对伸缩、平移、旋转等图像变换的鲁棒性; b) 使得高层网络可以在更大尺度下学习图像的更高层结构, 同时降低了网络参数使得大规模的网络训练变得可行。

由于卷积神经网络的参数量较大, 很容易发生过拟合, 影响最终的测试性能。研究者为了克服这一问题提出了很多改进的方法。Hinton 等人<sup>[35]</sup>提出了称为“dropout”的优化技术, 通过在每次训练迭代中随机忽略一半的特征点来防止过拟合, 取得了一定的效果。Wan 等人<sup>[36]</sup>进一步扩展了这一想法, 在全连接层的训练中, 每一次迭代时从网络的连接权重中随机挑选的一个子集置为 0, 使得每次网络更新针对不一样的网络结构, 进一步提升了模型的泛化性。此外还有一些简单有效的工程技巧, 如动量法、权重衰减和数据增强等。

## 2.5 循环神经网络

在全连接的 DNN 和 CNN 网络中, 每层神经元的信号只能向上一层传播, 样本的处理在各个时刻相互独立, 因此, 该类神经网络无法对时间序列上的变化进行建模, 如样本出现的时间顺序对于自然语言处理、语音识别、手写体识别等应用非常重要, 为了适应这种需求, 就出现了另一种神经网络结构—循环神经网络。RNN 中神经元的输出可以在下一个时间戳直接作用到自身, 即第  $i$  层神经元在  $t$  时刻的输入, 除了  $(i-1)$  层神经元在  $t-1$  时刻的输出外, 还包括其自身在  $t$  时刻的输入。如图 6 所示,  $(t+1)$  时刻网络的最终结果  $O(t+1)$  是该时刻输入和所有历史共同作用的结果, 这就达到了对时间序列建模的目的。

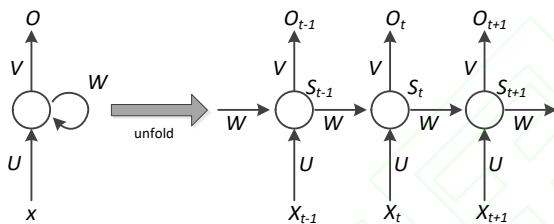


图 6 RNN 在时间上进行展开

为了适应不同的应用需求, RNN 模型出现了不同的变种, 主要包括以下几种:

a) 长短期记忆模型<sup>[37]</sup>。该模型通常比 vanilla RNNs 能够更好地对长时依赖进行表达, 主要为了解决通过时间的反向传播 (backpropagation through time, BPTT) 算法无法解决长时依赖问题, 因为 BPTT 会带来梯度消失或梯度爆炸问题。传统的 RNN 虽然被设计成可以处理整个时间序列信息, 但其记忆最深的还是最后输入的一些信号, 而受之前的信号影响的强度越来越低, 最后可能只起到一点辅助作用, 即 RNN 输出的还是最后的一些信号, 这样的缺陷使得 RNN 难以处理长时依赖的问题, 而 LSTM 就是专门为解决长时依赖而设计的, 不需要特别复杂地调试超参数, 默认就可以记住长期的信息, 其不足之处是模型结构较 RNN 复杂。一般包括输入门、遗忘门、输出门, “门”的结构就是一个使用 sigmoid 神经网络和一个按位做乘法操作, sigmoid 激活函数可以使得神经网络输出一个 0 到 1 之间的数值, 该值描述了当前输入有多少信息量可以通过这个结构, 类似一个门的功能, 当门打开时, sigmoid 神经网络的输出为 1, 全部信息都可以通过; 当门关上时, sigmoid 神经网络层输出为 0, 任何信息都无法通过。其中“遗忘门”的作

用是让循环神经网络“忘记”之前没有用的信息, 而“输入门”的作用是在循环神经网络“忘记”部分之前的状态后, 还需要从当前的输入补充最新的“记忆”, 而“输出门”会根据最新的状态  $C_t$ 、上一时刻的输出  $h_{t-1}$  和当前的输入  $x_t$  来决定该时刻的输出  $h_t$ , 如图 7 所示。因此 LSTM 结构可以更加有效的决定哪些信息应该被遗忘, 哪些信息应该得到保留, 因此成为当前语音识别、机器翻译、文本标注等领域常用的神经网络模型。

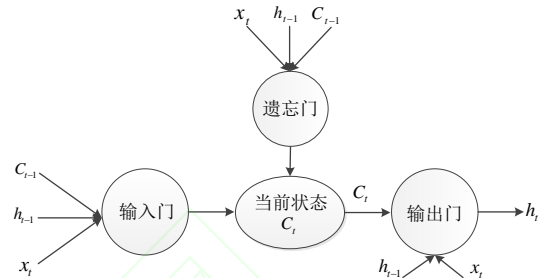


图 7 LSTM 单元结构示意图

b) Simple RNNs (SRNs)。SRNs 是 RNNs 的一种特例, 它是一个三层网络, 并且在隐藏层增加了上下文单元, 图 8 中的  $y$  便是隐藏层,  $u$  便是上下文单元。上下文节点与隐藏层节点一一对应, 且值是确定的。在每一步中, 使用标准的前向反馈进行传播, 然后使用学习算法进行学习。上下文每一个节点保存其连接的隐藏层节点的上一步的输出, 即保存上文, 并作用于当前步对应的隐藏层节点的状态, 即隐藏层的输入由输入层的输出与上一步的自己的状态所决定的。因此 SRNs 能够解决标准的多层感知机无法解决的对序列数据进行预测的任务。

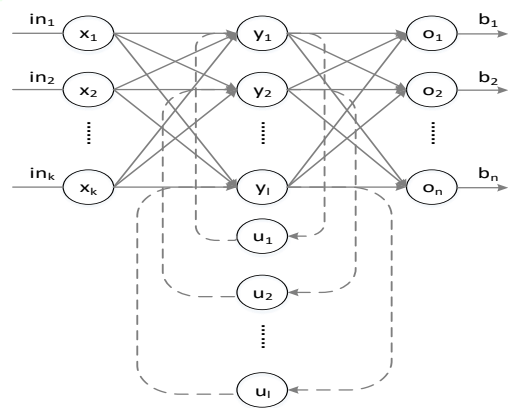


图 8 SRN 网络结构

c) Bidirectional RNNs。Bidirectional RNNs<sup>[38]</sup>是一个相对简单的 RNNs 如图 9 所示, 是由两个 RNN 上下叠加在一起组成的, 其输出由这两个 RNN 的隐藏层状态决定, 双向 RNN 网络模型可以用来根据上下文预测一个语句中缺失的词语, 即当前的输出不仅仅与前面的序列有关, 而且还与后面的序列有关。

此外针对不同的应用需求还出现了一些包括深度 RNN 模型 (DeepRNN)<sup>[39]</sup>、回声状态网络 (Echo State Networks)<sup>[40]</sup>、门控 RNN 模型 (gated recurrent unit, GRU)<sup>[41]</sup>、时钟频率驱动的 RNN (Clockwork RNN)<sup>[42]</sup> 模型等。

## 2.6 多模型融合的神经网络

除了单个的神经网络模型, 还出现了不同神经网络模型组

合的神经网络, 比如 CNN 和 RBM、CNN 和 RNN 等, 通过将各个网络模型的优势组合起来可以达到最优的效果, 文献 [43] Li Fei-Fei 等人将 CNN 与 RNN 相结合用于对图像描述的自动生成, 使得该组合模型能够根据图像的特征生成文字描述或者根据文字产生相应内容的图片。

随着深度学习技术的发展, 相信会有越来越多的性能优异的神经网络模型出现在大众的视野, 如近期火热的生成对抗网络 (generative adversarial networks, GAN) 及相应变种模型为无监督学习的研究开启了一扇门窗。

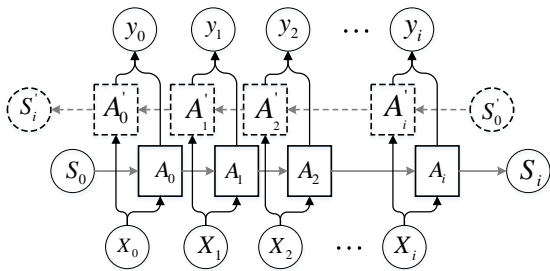


图 9 双向循环网络结构

### 3 基于深度学习的优化方法

随着神经网络模型层数越来越深、节点个数越来越多, 需要训练的数据集越来越大, 模型的复杂度也越来越高, 因此在模型的实际训练中单 CPU 或单 GPU 的加速方案存在着严重的性能不足, 一般需要十几天的时间才能使得模型的训练得到收敛, 已远远不能满足训练大规模神经网络、开展更多实验的需求。故多 CPU 或多 GPU 的加速方案成为训练大规模神经网络模型的首选。但是由于在图像识别或语言识别类应用中, 深度神经网络模型的计算量十分巨大, 且模型层与层之间存在的一定的数据相关性, 因此如何划分任务量以及计算资源是设计 CPU 或 GPU 集群加速框架的一个重要问题。本节主要介绍两种常用的基于 CPU 集群或 GPU 集群的大规模神经网络模型训练的常用并行方案。

#### 3.1 数据并行

当训练的模型规模比较大时, 可以通过数据并行<sup>[44]</sup>的方法来加速模型的训练, 数据并行可以对训练数据做切分, 同时采用多个模型实例, 对多个分块的数据同时进行训练。数据并行的大致框架如图 10 所示:

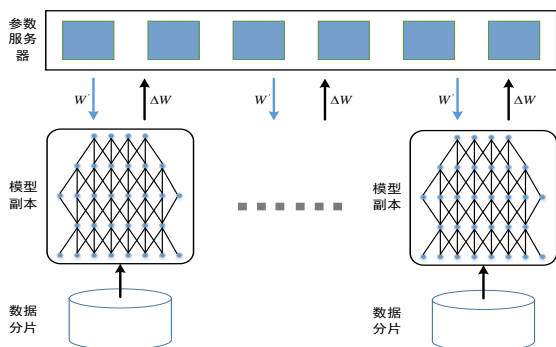


图 10 数据并行的基本架构

在训练过程中, 由于数据并行需要进行训练参数的交换, 因此通常需要一个参数服务器, 多个训练过程相互独立, 每个训练的结果, 即模型的变化量  $\Delta W$  需要提交给参数服务器, 参数服务器负责更新最新的模型参数  $W' = W - \eta \cdot \Delta W$ , 之后再最新的模型参数  $W'$  广播至每个训练过程, 以便各个训练过程可以从同一起点开始训练。在数据并行的实现中, 由于是采用同样的模型不同的数据进行训练, 影响模型性能的瓶颈在于多 CPU 或多 GPU 间的参数交换, 根据参数更新公式, 需要将所有模型计算出的梯度提交到参数服务器并更新到相应参数上, 因此数据片的划分以及与参数服务器的带宽可能会成为限制数据并行效率的瓶颈。

#### 3.2 模型并行

除了数据并行, 还可以采用模型并行的方式来加速模型的训练。模型并行是指将大的模型分拆成几个分片, 由若干个训练单元分别持有, 各个训练单元相互协作共同完成大模型的训练。图 11 为模型并行的基本框架。

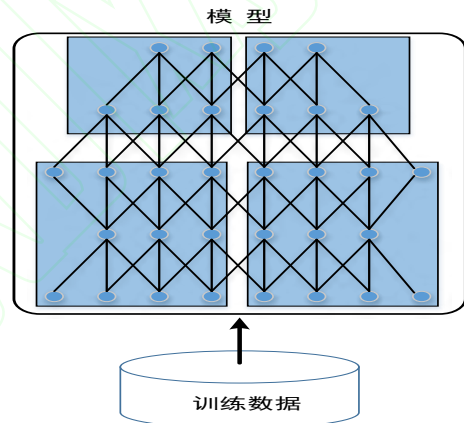


图 11 模型并行的基本架构

一般来说, 模型并行带来的通信和同步开销多于数据并行, 因此其加速比不及数据并行, 但对于单机内存无法容纳的大模型来说, 模型并行也是一个很好的方法, 2012 年 Imagenet 冠军模型 Axnet 就是采用两块 GPU 卡进行模型并行训练。

### 4 深度学习常用软件工具及平台

#### 4.1 常用软件工具

当前基于深度学习的软件工具有很多, 由于每种软件工具针对的侧重点不同, 因此, 根据需求的不同, 如图像处理、自然语言处理或是金融领域等, 因人而异、因项目而异采用合适的深度学习架构。本节主要介绍当下常用的深度学习软件工具。

第一类是 Tensorflow。由 Google 基于 DistBelief 进行研发的第二代人工智能系统, 该平台吸取了已有平台的长处, 既能让用户触碰底层数据, 又具有现成的神经网络模块, 可以使用户非常快速的实现建模, 是一个非常优秀的跨界平台, 该软件库采用数据流图模式实现数值计算, 流图中的节点表示数学运算, 边表示数据阵列, 基于该软件库开发的平台, 架构灵活, 代码一次开发可无须修改即可在单机、可移动设备或服务器等设备上运行, 同时可支持多 GPU/CPU 并行训练。



第二类是以 Keras 为主的深度学习抽象化平台。其本身不具有底层运算协调能力, 而是依托于 TensorFlow 或 Theano 进行底层运算, Keras 提供神经网络模块抽象化和训练中的流程优化, 可以让用户在快速建模的同时, 具有很方便的二次开发能力, 加入自己喜欢的模块。

第三类是以 Caffe、Torch、MXNet、CNTK 为主的深度学习功能性平台。该类平台提供了完备的基本模块, 支持快速神经网络模型的创建和训练, 不足之处是用户很难接触到这些底层运算模块。

第四类是 Theano, Theano 是深度学习领域最早的平台, 专注于底层基本运算。该平台有以下几个特点:

- a) 集成 NumPy 的基于 Python 实现的科学计算包, 可以和稀疏矩阵运算包 Scipy 配合使用, 全面兼容 Numpy 库函数。
- b) 易于使用 GPU 进行加速, 具有比 CPU 实现相对较大的加速比。
- c) 具有优异可靠性和速度优势。
- d) 可支持动态 C 程序生成。
- e) 拥有测试和自检单元, 可方便检测和诊断多类型错误。

表 1 为当前常用的几种软件工具, 可见基于深度学习的软件工具很多, 相应的编程语言也有很多, 没有哪一种编程平台或语言可以一统江湖。相信未来, 更新的、效率更好的编程语言或平台也可能会出现。

表 1 常用软件工具的相关比较

平台	底层语言	操作语言
TensorFlow	C++, Python	C++, Python
Keras	Python	Python
Caffe	C++	C++, Matlab, Python
Torch	C, Lua	Lua, C++
MXNet	C++, Python 等	C++, Python, Julia, Scala
CNTK	C++	C++, Python
Theano	Python, C	Python

4.2 工业界平台

随着深度学习技术的兴起, 不仅在学术界, 工业界如 Google、Facebook、百度、腾讯等科技类公司都实现了自己的软件平台, 主要有以下几种:

DistBelief 是由 Google 用 CPU 集群实现的数据并行和模型并行框架, 该集群可使用上万 CPU core 训练多达 10 亿参数的深度网络模型, 可用于语音识别和 2.1 万类目的图像分类<sup>[44]</sup>。此外 Google 还采用了由图像处理器 (graphics processing unit, GPU) 实现的 COTS HPC 系统, 也是一个模型并行和数据并行的框架, 由于采用了众核 GPU, 该 COTS 可以用 3 台 GPU 服务器在数天内完成对 10 亿参数的深度神经网络训练。

Facebook 实现了多 GPU 训练深度卷积神经网络的并行框架, 结合数据并行和模型并行的方式来训练卷积神经网络模型, 使用 4 张 NVIDIA Titan GPU 可在数天内训练 ImageNet 1000 分类的网络<sup>[45]</sup>。

Paddle ( parallel asynchronous distributed deep learning,Paddle) 是由国内的百度公司搭建的多机 GPU 训练平台<sup>[46]</sup>, 其将数据放置于不同的机器, 通过参数服务器协调各机

器的训练, Paddle 平台也可以支持数据并行和模型并行。

腾讯为加速深度学习模型训练也开发了并行化平台—Mariana, 其包含深度神经网络训练的多 GPU 数据并行框架、深度卷积神经网络的多 GPU 模型并行和数据并行框架, 以及深度神经网络的 CPU 集群框架。该平台基于特定应用的训练场景, 设计定制化的并行训练平台, 用于语音识别、图像识别、及在广告推荐中的应用<sup>[47]</sup>。

通过对以上几种工业界平台的介绍可以发现, 不管是基于 CPU 集群的 DistBelief 平台还是基于多 GPU 的 Paddle 或 Mariana 平台, 针对大规模神经网络模型的训练基本上都是采用基于模型的并行方案或基于数据的并行方案, 或是同时采用两种并行方案<sup>[48]</sup>。由于神经网络模型在前向传播及反向传播计算过程存在一定的数据相关性, 因此当前其在大规模 CPU 集群或者 GPU 集群上训练的方法并不多。

5 深度学习相关加速技术

近年来, 随着深度神经网络模型层数的增加, 与之相对应的权重参数成倍的增长, 从而对硬件的计算能力有着越来越高的需求, 尤其在数据训练的阶段。因此, 针对深度学习处理器的研究再次在工业界和学术界中崛起。目前针对数据训练阶段, 被业内广泛接受的是“CPU+GPU”的异构模式和 MIC (many integrated core,MIC) 众核同构来实现高性能计算。而针对数据推断阶段, 则较多地依赖于“CPU+FPGA”或“ASIC”。

5.1 CPU 加速技术

CPU 作为通用处理器, 本身不用做任何改变就可以完成神经网络算法的计算, 然而由于通常 CPU 的并行度低, 本身的计算能力也有限。现在常用的方式是进行分布式计算, 通过集合多个 CPU 从而提升计算的并行度。

CPU 作为传统的计算单元, 一开始就作为深度学习的计算平台, 但是由于深度学习的超大规模计算量以及高度的并行性, CPU 越来越难以适应深度学习的计算需求, 只能通过多核 CPU 或者 CPU 集群进行深度学习算法的加速。2012 年 6 月, 《纽约时报》披露了 Google Brain 项目, 该项目由著名的斯坦福大学机器学习教授 Andrew Ng 和在大规模计算机系统方面的世界顶尖专家 JeffDean 共同主导, 用 16000 个 CPU Core 的并行计算平台训练一种称为“深度神经网络”的机器学习模型 (内部约有 10 亿个节点), 该训练过程进行了 7 天才能完成猫脸识别任务, 因此并行能力的缺乏是限制 CPU 加速深度学习应用的主要因素, 当前基于 CPU 的多是异构平台, 如 CPU+GPU 或 CPU+FPGA 的异构加速平台, 复杂控制及串行部分由 CPU 执行, 并行部分由 GPU 或 FPGA 执行。

5.2 GPU 加速技术

对于深度学习来说, 目前硬件加速主要靠使用图形处理单元 (GPU)。相比传统的通用处理器 (CPU), GPU 的核心计算能力要多出几个数量级, 也更容易进行并行计算。尤其是 NVIDIA 通用并行计算框架 (compute unified device

architecture,CUDA), 作为最主流的 GPU 编写平台, 各主要的深度学习工具均用其来进行加速。

GPU 的众核体系结构包含几千个流处理器, 可将运算并行化执行, 大幅缩短模型的运算时间。随着 NVIDIA、AMD 等公司不断推进其 GPU 的大规模并行架构支持, 面向通用计算的 GPU (general-purpose GPU, GPGPU) 已成为加速并行应用程序的重要手段。得益于 GPU 众核体系结构, 程序在 GPU 系统上的运行速度相较于单核 CPU 往往提升几十倍乃至上千倍。目前 GPU 已经发展到了较为成熟的阶段。利用 GPU 来训练深度神经网络, 可以充分发挥其数以千计计算核心的高效并行计算能力, 在使用海量训练数据的场景下, 所耗费的时间大幅缩短, 占用的服务器也更少。如果针对适当的深度神经网络进行合理优化, 一块 GPU 卡可相当于数十甚至上百台 CPU 服务器的计算能力, 因此 GPU 已经成为业界在深度学习模型训练方面的首选解决方案。

### 5.3 FPGA 加速技术

作为 GPU 在算法加速上强有力的竞争者, 现场可编程逻辑门阵列 (field programmable gate array, FPGA) 近年来受到了越来越多的关注, FPGA 作为深度学习加速器具有以下几点优势:

a) 可重构。FPGA 芯片可以被重复编程, 用户可以针对不同应用的计算特征定制阵列结构、计算单元、数据并行策略和存储结构。因此, FPGA 能够灵活的适应高性能计算领域的不同计算应用、算法以及模型, 实现快速的更新、升级以及调试。此外, 新一代的 FPGA 芯片还具有动态可重构的能力, 可以在系统不掉电和不干扰当前任务的前提下实现快速的切换。

b) 低功耗。目前主流的通用处理器在满负荷状态下的功耗大约为 60-80W, 而 FPGA 的平均功耗不超过 20W, 远低于 GPU 和 CPU 的功耗, 低功耗是 FPGA 当前受到极大关注的重要一点。

c) 可定制。FPGA 可以根据应用需求灵活的对数据位宽进行配置, 满足不同精度的计算需求, 由于 FPGA 具有丰富的逻辑资源、存储资源和 DSP 资源, 因此可以在一个 FPGA 芯片内部定制多种运算单元。

d) 高性能。FPGA 芯片上具有大量的片上存储资源, 可以提供强大的带宽和并行访存能力。针对特定的应用定制计算通路和存储结构, 同时开发粗粒度线程级并行和细粒度的指令级并行, 可以最大限度为开发 FPGA 芯片提供计算和访存能力。鉴于 FPGA 的以上优势, 在 15、16 年的 ISCA、Micro、NIPS 等顶会上出现了不少针对深度学习的 FPGA 加速器。而在刚刚结束的 FPGA2017 中获得最佳论文的深鉴科技 ESE 语音识别引擎, 结合深度压缩 (deep compression)、专用编译器以及 ESE 专用处理器架构, 在中端的 FPGA 上即可取得比 Pascal Titan X GPU 高 3 倍的性能, 并将功耗降低 3.5 倍。据悉, 该 ESE 语音识别引擎, 也是深鉴科技 RNN 处理器产品的原型。

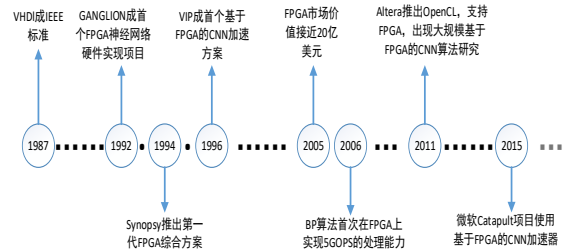


图 12 FPGA 深度学习研究路线图

当然 FPGA 也并非完美无瑕, 同样面临一系列挑战, 比如硬件编程困难, FPGA 的开发需要对底层硬件有一定的知识且使用硬件描述语言 (hardware description language, HDL) 进行开发, 需要开发人员具有长期的经验积累, 虽然已经有高级编程语言可以使用 C 或 C++ 进行开发, 但由于其还不完善, 性能还没达到硬件描述语言的程度, 因此还有一定的局限性。此外 FPGA 存在许多编程模式<sup>[49]</sup>, 还未形成统一的编程模型, 且模块的重用也是一大难题, 因此, FPGA 在深度学习的大规模应用甚至替代 GPU 还有很长的路要走。

鉴于成本上的考虑, 基于 FPGA 的低功耗优势, 因此使用 FPGA 做深度学习加速的多是企业用户, 如百度、微软、IBM 等公司都有专门做 FPGA 的团队为服务器加速, 图 12 为基于 FPGA 的深度学习研究大致过程。

### 5.4 ASIC 加速技术

与 FPGA 的可编程性不同, 专用集成电路 (application-specific integrated circuit, ASIC) 一旦设计制造完成后电路结构就固定了, 无法再改变。其主要代表公司是 Movidius。ASIC 具有以下几个特点: a) 需要大量设计时间以及验证和物理设计周期, 因此需要相对多的上市时间; b) 同一时间点上用最好的工艺实现的 ASIC 加速器的速度会比用同样工艺 FPGA 实现的加速器速度快 5-10 倍, 且量产后 ASIC 的成本会远远低于 FPGA 方案 (便宜 10 到 100 倍); 因此 FPGA 主要用于服务器市场, 而 ASIC 主要用于移动终端的消费电子领域。表 2 为 FPGA 和 ASIC 的相关比较。

表 2 FPGA 和 ASIC 的相关比较

	上市速度	性能	成本	量产成本	可配置	目标市场
FPGA	快	差	低	高	完全	企业军工
ASIC	慢	好	高	低	有限	消费电子

当前在专用神经网络加速器方面做得最好的当属中科院计算所的陈云霁团队, 其设计的寒武纪系列神经网络加速器连续在 2013 年 ASPLOS、2014 年 MICRO、2015 年 ASPLOS、ISCA、2016 年 ISCA、MICRO 等国际顶级会议发表, 并在国际上产生了重要影响, 已成为国际上专用神经网络加速器的代表。2013 年提出的 DianNao<sup>[50]</sup> 成为国际上首个深度学习处理器, 并获得体系结构 A 类会议最佳论文, 2014 年提出的 DaDianNao<sup>[51]</sup> 是国际上首个多核深度学习处理器, 并获得 MICRO14 最佳论文, 2015 年提出的 PuDianNao<sup>[52]</sup> 可以支持多种神经网络模型, 成为国际上首个通用机器学习处理器, 2015 年提出的 ShiDianNao<sup>[53]</sup> 是一个可以嵌入在手机等终端面向视频、图像智能助理具有极低功耗的专用神经网络处理器, 其相比主流 GPU 有 28 倍的



性能, 4700 倍的性能功耗比。2016 年提出的 Cambricon<sup>[54]</sup>, 一种神经网络指令集是国际上首个神经网络通用指令集, 且获得 ISCA 评审最高分, 足见其研究的受重视程度, 该通用指令集可以高效的实现当前所有的神经网络模型, 通过该指令集可以编写出不同的神经网络模型, 该工作使得专用神经网络处理器具有了可编程的能力。

此外, 名震一时的“AlphaGo”除了配备 1920 颗 CPU 和 280 颗 GPU 外, 谷歌披露它还安装一定数量的张量处理单元 (tensor processing unit, TPU)。谷歌称 TPU 是专为谷歌开源项目 TensorFlow 而优化的硬件加速器, 属于一款 ASIC 加速器。业内普遍认为“AlphaGo”对围棋局势的预判所使用的价值网络就是依赖 TPU 的发挥。谷歌指出, 在深度学习方面, TPU 兼具了 CPU 与 ASIC 的特点, 可编程, 高效率, 低能耗, 因此 TPU 可以兼具桌面机与嵌入式设备的功能。

另外, 中星微“数字多媒体芯片技术”国家重点实验室宣布, 中国首款嵌入式神经网络处理器 (neural processing unit, NPU) 芯片诞生并实现量产。这款 NPU 芯片采用了“数据驱动并行计算”架构, 这种数据流类型的处理器, 极大地提升了计算能力与功耗的比例, 特别擅长处理视频、图像类的海量多媒体数据, 使得人工智能在嵌入式机器视觉应用中可以大显身手。

通过最近国际顶会的相关论文以及商业产品可知, 基于专用的神经网络加速器也是当前的一个研究热点, 尤其是针对嵌入式平台, 如手机、无人机、无人车等。相信随着研究的进一步深入, 拥有不同体系结构的专用神经网络加速器会越来越多。

## 5.5 其他技术研究

除了传统的硬件加速器, 随着半导体技术的发展, 新型的加速方案不断涌现。

IBM 的 TrueNorth<sup>[55]</sup> 计算平台, 号称只有邮票大小, 重量只有几克, 但却集成了 54 亿个硅晶体管, 内置了 4096 个内核, 100 万个“神经元”、256 亿个“突触”, 能力相当于一台超级计算机, 功耗却只有 65 毫瓦。与传统冯诺依曼结构不同, 芯片把数字处理器当作神经元, 把内存作为突触, 它的内存、CPU 和通信部件是完全集成在一起。因此信息的处理完全在本地进行, 而且由于本地处理的数据量并不大, 传统计算机内存与 CPU 之间的瓶颈不复存在, 因此, 有人把 IBM 的芯片称为是计算机史上最伟大的发明之一, 将会引发技术革命, 颠覆从云计算到超级计算机乃至智能手机等一切。IBM 不久前发表于 PNAS 的论文, 描述了 IBM 研究员训练卷积神经网络在神经形态硬件上分类图像和语音, 在 8 个标准数据集上达到了接近目前最先进的精度, 每秒 1200~2600 帧的速度处理, 能耗 25~275 毫瓦。这是首次将深度学习算法的力量和神经形态处理器的高能效相结合, 向着实现嵌入式类脑智能计算又迈进了一步。但是短期看来, 情况并非那么乐观。首先芯片的编程困难, 这种芯片要颠覆传统的编程思想, 因此需要一套全新的配套开发工具, 由于其相关资料尚未完全公开, 因此, 该芯片的能力有待进一步证实。

2017 年年初高通披露了其最新的 Snapdragon 835 的相关信息, 新增加了机器学习方面的功能, 包括支持客户生成神经网络层、同时还支持谷歌的机器学习架构 TensorFlow。据称 Hexagon 682 是首个支持 TensorFlow 和 Halide 架构的移动数字信号处理器 (digital signal processing, DSP)。而早在 2013 年, 高通就展示了一款内置 Zeroth 芯片的机器人, 它能够在接受外界信息之后学会选择正确的路线行进。另外, DSP 供应商 CEVA 也于近两年在机器学习领域进行了研究, 并推出了多款适应于深度学习的 DSP 芯片。

另外, 新型材料如忆阻器 (memristor) 也被用于神经网络的构建, 2016 年 Rajeev Balasubramanian 教授课题组与 HP 实验室合作, 提出了一种基于忆阻器交叉开关的卷积神经网络加速器<sup>[56]</sup>, 基于流水线的组织方式来加速神经网络的不同计算层, 并采用 eDRAM 来实现流水线段间数据寄存。同样基于新型材料的 ReRAM 被认为是今后替代当前 DRAM 作为密度更大、功耗更小的下一代存储的技术之一。其独特的交叉网络结构和多比特存储性质, 能以很高的能量效率加速神经网络计算中的主要计算模块。加州大学课题组结合 ReRAM 的这种特性, 设计了一种可以在“存储”状态和“神经网络加速器”状态之间灵活切换的内存计算架构。新型材料可以融合数据存储与计算, 在较低的功耗下达到很高的计算性能。然而这类芯片及硬件设计, 由于受到制造工艺的影响, 也存在许多限制。

此外, 三维堆叠技术也被引入到深度学习加速器的设计中, 设计以存储为中心的总体结构, 在 CPU 周围设置大量的加速器单元。Saibal Mukhopadhyay 教授课题组提出了一种基于三维堆叠存储的可编程神经网络加速器计算结构 Neurocube<sup>[57]</sup>, 采用以三维堆叠存储为基础的内存计算架构, 在三维堆叠内存的最下层 (逻辑层) 中添加计算单元, 可以通过存储内部的巨大带宽, 消除不必要的搬运; 并且使用定制逻辑模块加速神经网络的计算 (包括训练部分)。

除了硬件结构上的加速, 16 年的顶会上还提出一些算法层次上的加速, 如 2016MICRO 会议上, 纽约州立大学石溪分校的 Manoj Alwani 等人<sup>[58]</sup>提出一种 Fused-layer 的卷积神经网络加速器, 通过融合两个或两个以上的卷积层, 使得 DRAM 只加载输入特征图, 而不需要将中间结果写回, 只保存计算结果, 该方法可以大幅减少层与层之间的片外数据移动, 进而大幅降低可移动的数据量。此外, 在 MICRO2016 上, 中科院计算所陈云霄等人提出了一种稀疏的神经网络加速器, 通过对神经网络的分析, 找出神经网络模型相邻层之间的稀疏连接, 在不降低模型识别率的基础上, 将全连接网络变成稀疏连接, 进而压缩神经网络模型, 只计算和存储连接的神经元, 因此可以大幅降低模型的计算量和存储需求。

通过近两年国际顶级会议的相关论文可以发现, 神经网络加速器的研究是当前的一个研究热点, 不仅有基于硬件的神经网络加速也有基于软件算法层次上的神经网络加速研究, 由于深度神经网络模型的超大规模计算量, 因此未来需要从硬件和

软件算法层次等方面一起来加速神经网络算法模型的计算。

## 6 存在问题与研究展望

前途光明, 道路曲折, 尽管深度学习技术在图像处理、语音识别、自然语言处理等领域取得了突破性的进展, 但是仍旧有许多问题亟待解决。

a) 无监督数据的特征学习。当前, 标签数据的特征学习仍然占据主导地位, 而真实世界存在着海量的无标签数据, 将这些无标签数据逐一添加人工标签, 显然是不现实的, 因此, 随着深度学习技术的发展, 必将越来越重视对无标签数据的特征学习, 以及将无标签数据进行自动添加标签技术的研究。

b) 基于模型融合的深度学习方法。相关研究表明, 单一的深度学习模型往往不能带来最好的效果, 而通过增加深度来提高模型效果的方法往往会有一定的局限性, 如, 梯度消失问题、计算过于复杂、模型的并行性有限等问题, 因此通过融合其他模型或者多种简单模型进行平均打分, 可能会带来更好的效果。

c) 迁移学习。迁移学习<sup>[59,60,61]</sup>可以说是一种“站在巨人肩上”的学习方法, 可以在不同领域中进行知识迁移, 使得在一个领域中已有的知识得到充分的利用, 不需要每次都求解问题视为全新的问题。一个好的迁移学习方法可以大大加快模型的训练速度。

d) 嵌入式设备。目前深度学习技术正往嵌入式设备靠近, 即原来的训练往往在服务器或者云端, 而嵌入式设备通过网络将待识别的任务上传至云端, 再由云端将处理结果发送到嵌入式端, 随着嵌入式设备计算能力的提升、新型存储技术以及储电技术的进步, 在嵌入式端完成实时训练是完全可能的, 到时就可能实现真正的人工智能。因此, 嵌入式设备成为将来的研究重点, 包括军/民用无人机、无人车/战车、无人潜水器等智能化装备。

e) 低功耗设计。鉴于嵌入式设备对功耗非常敏感, 因此具有功耗优势的 FPGA 芯片可能成为研究的一个热点, 设计基于 FPGA 类似 Caffe 的可编程深度学习软件平台会是一个研究方向。

f) 算法层优化。由于深度学习技术巨大的计算量和存储需求, 不仅要在硬件上进行加速, 算法模型优化上也可以锦上添花, 如稀疏编码、层级融合、深度压缩等相关技术也会继续研究。

g) 脉冲神经网络。脉冲神经网络目前虽然在精度上并不具有和机器学习算法一样的水准, 一方面因为学习算法, 另一方面因为信息编码, 然而脉冲神经网络是更接近生物学现象和观察的模型, 因此, 未来在脉冲神经网络研究上的突破也是人工智能研究上的一个重点。

h) 非精确计算。鉴于神经网络模型对计算精度不是特别敏感, 因此, 非精确计算越来越引人瞩目, 被认为是降低能耗最有效的手段之一, 通过牺牲可接受的实验精度来换取明显的资源节约(能耗、关键路径延迟、面积), 可以将非精确计算和硬

件神经网络相结合来扩大应用范围、提高错误恢复能力和提高能源节约程度, 使得该神经网络成为未来异构多核平台的热门备选加速器。

i) 模型压缩。深度学习仍在不断进步, 目前网络的规模开始朝着更深但是参数更少的方向发展, 如微软提出的深度残差网络和 Stanford 提出的稀疏神经网络, 该研究体现了深度神经网络中存在参数的冗余性, 可以预见未来的算法研究会进一步压缩冗余参数的存在空间, 从而网络可能具有更好的精度但是却拥有更少的参数。

## 7 结束语

深度学习作为机器学习领域的一个重要研究方向, 近年来受到了越来越多的关注, 鉴于深度学习研究领域的发展变化日新月异, 本文系统的介绍了深度学习的相关研究现状, 从深度学习的应用领域入手, 重点介绍了深度学习的常用神经网络模型, 分析了两种常用的深度学习模型并行训练方法, 比较了两种模型训练方法的优缺点, 对比分析了 7 种常用的深度学习开源软件工具的应用特点及几种工业界的研究平台, 并重点介绍了当前神经网络硬件加速器的研究现状, 对比分析了 CPU、GPU、FPGA、ASIC 等常用的硬件加速器, 并对深度学习的未来研究方向进行了展望, 可以预见, 随着新型存储、光互连技术、半导体工艺等新技术、新工艺的使用, 相信真正的人工智能一定能够实现。总之深度学习技术未来的发展仍然是充满不同的机遇和挑战, 也是大有可为的。

## 参考文献:

- [1] Yang J, Yu K, Gong Y, et al. Linear spatial pyramid matching using sparse coding for image classification [J]. 2009: 1794-1801.
- [2] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks. [J]. Science, 2006, 313 (5786): 504.
- [3] Lecun Y, Boser B, Denker J S, et al. Backpropagation applied to handwritten zip code recognition [J]. Neural Computation, 1990, 2 (4): 541-551.
- [4] Bengio Y, Vincent P, Janvin C. A neural probabilistic language model [J]. Journal of Machine Learning Research, 2003, 3 (6): 1137-1155.
- [5] Collobert R, Weston J, Karlen M, et al. Natural Language Processing (Almost) from Scratch [J]. Journal of Machine Learning Research, 2011, 12 (1): 2493-2537.
- [6] Mikolov T, Deoras A, Kombrink S, et al. Empirical evaluation and combination of advanced language modeling techniques [C]// Proc of Conference of the International Speech Communication Association. 2011: 605-608.
- [7] Schwenk H, Rousseau A, Attik M. Large, pruned or continuous space language models on a GPU for statistical machine translation [C]// Will We Ever Really Replace the N-Gram Model? on the Future of Language Modeling for Hlt. Association for Computational Linguistics, 2012: 11-19.

- [8] Socher R, Huval B, Manning C D, et al. Semantic compositionality through recursively-matrix-vector spaces [C]// Proc of Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning. 2012: 1201-1211.
- [9] Le Q V, Mikolov T. Distributed Representations of Sentences and Documents [J]. 2014, 4: II-1188.
- [10] Chen Y. Convolutional Neural Network for Sentence Classification [J]. 2015.
- [11] Bourlard H, Kamp Y. Auto-association by multilayer perceptrons and singular value decomposition [J]. Biological Cybernetics, 1988, 59 (4): 291-294.
- [12] Vincent P, Larochelle H, Bengio Y, et al. Extracting and composing robust features with denoising autoencoders [C]// Proc of International Conference. 2008: 1096-1103.
- [13] Le Q V, Ngiam J, Coates A, et al. On optimization methods for deep learning [C]// Proc of International Conference on Machine Learning. 2011: 265-272.
- [14] Zou W Y, Ng A Y, Yu K. Unsupervised learning of visual invariance with temporal coherence [C]// Proc of Workshop on Deep Learning and Unsupervised Feature Learning. 2011, 3.
- [15] Rifai S, Vincent P, Muller X, et al. Contrastive auto-encoders: Explicit invariance during feature extraction [C]// Proc of the 28th International Conference on Machine Learning. 2011: 833-840.
- [16] Krizhevsky A, Hinton G E. Using very deep autoencoders for content-based image retrieval [C]// Proc of ESANN. 2011.
- [17] Zhang J, Shan S, Kan M, et al. Coarse-to-fine auto-encoder networks (cfan) for real-time face alignment [C]// Proc of European Conference on Computer Vision. [S. l.]: Springer International Publishing, 2014: 1-16.
- [18] Hinton G E. A Practical Guide to Training restricted Boltzmann machines [J]. Momentum, 2012, 9 (1): 599-619.
- [19] Hinton G E. Training products of experts by minimizing contrastive divergence [J]. Neural computation, 2002, 14 (8): 1771-1800.
- [20] Ackley D H, Hinton G E, Sejnowski T J. A learning algorithm for Boltzmann machines [J]. Cognitive science, 1985, 9 (1): 147-169.
- [21] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks [C]// Aistats. 2010, 9: 249-256.
- [22] 赵元庆, 吴华. 多尺度特征和神经网络相融合的手写体数字识别 [J]. 计算机科学, 2013, 40 (8): 316-318.
- [23] Hinton G E. Modeling pixel means and covariances using factorized third-order Boltzmann machines [C]// Proc of Computer Vision and Pattern Recognition Conference. 2010: 2551-2558.
- [24] Courville A, Bergstra J, Bengio Y. A spike and slab restricted Boltzmann machine [C]// Proc of the 14th International Conference on Artificial Intelligence and Statistics. 2011: 233-241.
- [25] Memisevic R, Hinton G. Unsupervised learning of image transformations [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2007: 1-8.
- [26] Larochelle H, Bengio Y. Classification using discriminative restricted Boltzmann machines [C]// Proc of the 25th international conference on Machine learning. 2008: 536-543.
- [27] 孙志军, 薛磊, 许阳明. 基于深度学习的边际 Fisher 分析特征提取算法 [J]. 电子与信息学报, 2013, 35 (4): 805-811.
- [28] Salakhutdinov R, Hinton G. Deep Boltzmann machines [C]// Artificial Intelligence and Statistics. 2009: 448-455.
- [29] Lee H, Grosse R, Ranganath R, et al. Co-convolutional deep belief networks for scalable unsupervised learning of hierarchical representations [C]// Proc of the 26th Annual International Conference on Machine Learning. 2009: 609-616.
- [30] Yu K, Lin Y, Lafferty J. Learning image representations from the pixel level via hierarchical sparse coding [C]// Proc of IEEE Conference on Computer Vision and Pattern Recognition. 2011: 1713-1720.
- [31] Zeiler M D, Taylor G W, Fergus R. Adaptive deconvolutional networks for mid and high level feature learning [C]// Proc of IEEE International Conference on Computer Vision. 2011: 2018-2025.
- [32] Russakovsky O, Deng J, Su H, et al. ImageNet large scale visual recognition challenge [J]. International Journal of Computer Vision, 2015, 115 (3): 211-252.
- [33] 10 Breakthrough Technologies 2013 [N]. MIT Technology Review, 2013-04-23
- [34] Boureau Y L, Ponce J, Lecun Y. A theoretical analysis of feature pooling in visual recognition [C]// Proc of International Conference on Machine Learning. 2010: 111-118.
- [35] Hinton G E, Srivastava N, Krizhevsky A, et al. Improving neural networks by preventing co-adaptation of feature detectors [J]. Computer Science, 2012, 3 (4): 212-223.
- [36] Wan L, Zeiler M, Zhang S, et al. Regularization of neural networks using drop connect [C]// Proc of International Conference on Machine Learning. 2013: 1058-1066.
- [37] Graves A. Supervised Sequence Labelling with Recurrent Neural Networks [M]. Springer Berlin Heidelberg, 2012.
- [38] Kiperavasser E, Goldberg Y. Simple and accurate dependency parsing using bidirectional LSTM feature representations [J]. 2016.
- [39] Graves A, Mohamed A R, Hinton G. Speech recognition with deep recurrent neural networks [J]. 2013, 38 (2003): 6645-6649.
- [40] York S N. Echo State Networks [J]. Springer New York, 2015.
- [41] Cho K, Merriënboer B V, Gulcehre C, et al. Learning phrase representations using rnn encoder-decoder for statistical machine translation [J]. Computer Science, 2014.
- [42] Koutník J, Greff K, Gomez F, et al. A Clockwork RNN [J]. Computer Science, 2014: 1863-1871.
- [43] Karpathy A, Li F F. Deep visual-semantic alignments for generating image descriptions [J]. IEEE Trans on Pattern Analysis & Machine Intelligence,



- 2017, 39 (4): 664.
- [44] Dean J, Corrado G S, Monga R, et al. Large scale distributed deep networks [C]// Proc of International Conference on Neural Information Processing Systems. Curran Associates Inc. , 2012: 1223-1231.
- [45] Yadan O, Adams K, Taigman Y, et al. Multi-GPU Training of ConvNets [J]. Computer Science, 2014.
- [46] Yu K. Large-scale deep learning at Baidu [C]// Proc of ACM International Conference on Information & Knowledge Management. 2013: 2211-2212.
- [47] 腾讯公司. 深度学习在腾讯的平台化和应用实践 [EB/OL]. <http://www.36dsj.com/archives/20176>.
- [48] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [C]// Proc of International Conference on Neural Information Processing Systems. Curran Associates Inc. 2012: 1097-1105.
- [49] Dehon A, Adams J, Delormier M, et al. Design patterns for reconfigurable computing [C]// Proc of IEEE Symposium on Field-Programmable Custom Computing Machines. IEEE Computer Society, 2004: 13-23.
- [50] Chen T, Du Z, Sun N, et al. DianNao: a small-footprint high-throughput accelerator for ubiquitous machine-learning [J]. ACM SIGPLAN Notices, 2014, 49 (4): 269-284.
- [51] Chen Y, Luo T, Liu S, et al. DaDianNao: A Machine-Learning Supercomputer [C]// Proc of IEEE/ACM International Symposium on Micro-architecture. IEEE Computer Society, 2014: 609-622.
- [52] Liu D, Chen T, Liu S, et al. PuDianNao: a polyvalent machine learning accelerator [J]. ACM SIGPLAN Notices, 2015, 43 (1): 369-381.
- [53] Du Z, Shi DianNao. shifting vision processing closer to the sensor [C]// Proc of International Symposium on Computer Architecture. 2015: 92-104.
- [54] Liu S, Du Z, Tao J, et al. Cambricon: An Instruction Set Architecture for Neural Networks [J]. ACM SIGARCH Computer Architecture News, 2016, 44 (3): 393-405.
- [55] Esser S K, Merolla P A, Arthur J V, et al. From the cover: convolutional networks for fast, energy-efficient neuromorphic computing [J]. Proceedings of the National Academy of Sciences of the United States of America, 2016, 113 (41): 11441.
- [56] Chi P, Li S, Xu C, et al. PRIME: a novel processing-in-memory architecture for neural network computation in ReRAM-based main memory [C]// Proc of International Symposium on Computer Architecture. IEEE Press, 2016: 27-39.
- [57] Kim D, Kung J, Chai S, et al. Neurocube: a programmable digital neuromorphic architecture with high-density 3D memory [C]// Proc of the 43rd ACM/IEEE International Symposium on Computer Architecture. 2016.
- [58] Alwani M, Chen H, Ferdman M, et al. Fused-layer CNN accelerators [C]// Proc of IEEE/ACM International Symposium on Microarchitecture. IEEE Computer Society, 2016: 1-12.
- [59] Weiss K, Khoshgoftar T M, Wang D D. A survey of transfer learning [J]. Journal of Big Data, 2016, 3 (1): 9.
- [60] 庄福振, 罗平, 何清, 等. 迁移学习研究进展 [J]. 软件学报, 2015, 26 (1): 26-39.
- [61] 周志华, 杨强. 机器学习及其应用 [M]. 北京: 清华大学出版社, 2011.