

大数据时代下深度学习理论综述

河南工业贸易职业学院信息工程系 邱俊玲

深度学习是机器学习的一个分支，是人工智能的一部分。深度学习通过模拟人的学习活动，达到获取新知识或技能的目的。本文首先论述了深度学习的理论及实现框架，重点介绍了卷积神经网络，然后分析了深度学习现阶段的应用，最后提出深度学习尚待解决的问题，以期该理论有更长远的发展。

一、引言

据 Domo 公司在 2016 年统计数据显示，全球每分钟在亚马逊上产生 222283 美元的交易，YouTube 上有 400 小时

的视频共享资源，谷歌翻译出 69500000 个字符，气象频道收到 13888889 次预报请求，Dropbox 上传 833333 个新文件等。面对如此海量的数据，如何处理这些日益膨胀的数据，

综上，情况如表 1 所示，Y 轴方向的刚度最大，X 轴和 Y 轴方向的刚度稍低。接下来再根据各个关键件的最大变形量以及所占比重来查找机床的薄弱环节。

表 1 卧式加工中心整机变形和刚度

力 (F_x, F_y, F_z) /N	刀具端最大变形 量 (mm)	工件端最大变形 量 (mm)	机床刚度 (106N/m)
(10000,0,0)	0.20854	0.0946	33
(0,10000,0)	0.173	0.011	54
(0,0,10000)	0.099	0.162	38

四、薄弱环节查找

数控机床工艺系统是一个机床、刀具、工件组成的串联系统，工艺系统刚性低于整机中刚性最低的部件，因此分析各部件对工艺系统刚性的影响，找出薄弱环节，提高机床刚度。

根据有限元分析的数据，总结 HMC50e 卧式加工中心各部件在 X、Y、Z 方向分别作用下所引起变形的影响，分析各部件引起机床的位移值，如图 8 所示；分析各部变形在整个系统中变形的比重，如图 9 所示。

通过图 8 和图 9 得知，床身对 X、Y 方向的变形影响较大。由此可通过改进床身件，可提高机床整体刚度。

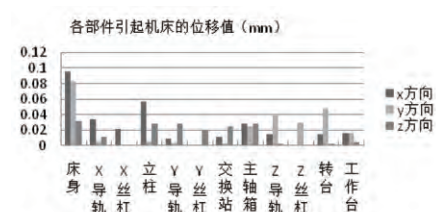


图 8 机床各部件变形引起工艺系统变形的真实值

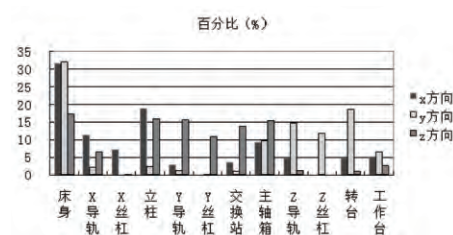


图 9 机床各部件变形在工艺系统变形的比重

五、结语

从上述分析可以看出，应用 ANSYS 软件对 HMC50e 卧式加工中心进行有限元分析，计算机床在三个方向切削力作用下的静刚度，找到机床床身件为整机的薄弱环节。这种有限元软件分析机床整机的刚度比较准确和简便，有利于快速修改设计缺陷，减少设计成本。这对于提高机床产品的质量和技术水平，具有重要的理论意义和实用价值。

如何挑选出适合自己的数据，如何利用这些数据更好地服务客观世界，找出规律进行分类，实现对大数据的合理利用是研究者应考虑和解决的当务之急，在此条件下，深度学习应运而生。

深度学习的基本思路是通过算法的设置，训练机器能从大量的历史数据中找出规律，从而对新的样本进行智能识别或预测未来。这是传统的机器学习方法，利用样本中的特征向量来训练数据进而建立模型。但随着大数据时代的到来，为深度学习创造了条件，无监督学习方法成为研究人员更为热衷的方向。

二、深度学习的发展历程

深度学习的研究起源于人工神经网络，是由输入层、隐藏层、输出层组成的三层模型。一个神经网络由节点和有向边表示，每个神经元节点的输出作为连入的下一个节点的输入。每次输出前，每个神经元的自运算是运用了一个非线性激活函数，正因此神经网络的模型具有非线性的特征。深度学习的发展经历了漫长的历史时期。早在 1943 年，McCulloch 和 Pitts 建立了第一个神经网络计算模型，称为 MP 模型。1949 年，Donald 利用心理学知识，提出了自组织学习的规则 Hebbian Learning，此后 Wesley 等人用计算机仿真了 Hebbian 网络。随后的 1958 年，Frank 将 MP 模型进行推广，提出了第一个人工神经网络模型——感知机模型，这是一种单层神经网络模型，并从理论上证明了这一模型处理线性可分类问题的收敛性。1962 年 Rosenblatt 总结了这一研究成果，并出版了《神经动力学原理：感知机和大脑机制的理论》，自此机器学习的研究进入了第一个高潮时期。然而在 1969 年，Minsky 和 Paper 从感知器的局限性入手，以一个逻辑异或运算的例子推翻了感知机模型，证明了不包含隐层单元的感知器无法处理异或这种非线性分类问题，自此神经网络的研究坠入低谷。

1982 年美国生物物理学家 Hopfield 采用了新型的全互连神经网络，利用所定义的计算能量函数成功给出了 TSP 问题的近似解，这一理论给了神经网络新的春天。1986 年 David E. Rumelhart 发明了反向传播 BP 算法，但由于其梯度弥散的问题导致该算法的应用有其局限性，在神经网络初始参数完全随机的情况下，BP 算法被论证出在浅层神经网络中是有效的，然而在深层网络中它的处理效果并

不理想。

2006 年，Hinton 等人提出通过非监督学习方法对原始数据进行逐级训练，再采用监督方法精调整个网络。非监督学习进行的训练为神经网络提供了较好的初始参数，降低了网络的优化难度。2008 年，Vincent 等人提出了降噪自编码器，2011 年 Rafir 等人提出了收缩自编码器，2012 年微软公司建立了深度神经网络——隐马尔科夫混合模型，并将其应用于语音识别；同年，Ng 在“GoogleBrain”项目中使用无监督深度学习方法。2015 年 Xu 提出了集合注意力的场景识别方法，同年 He 等人提出了拥有 152 层的深度残差网络模型。深度学习理论的探索在曲折中前进。

三、深度学习的理论及框架结构

1. 深度学习的理论

传统的机器学习需要对数据进行预处理，然后对处理的数据进行特征提取和选择，以达到学习的目的，实现推理、预测的功能。但该方法的准确性依赖于特征提取和特征选择的好坏，而人类的学习往往是自动的，为了使机器也能实现自动学习，模拟人脑进行分层分析学习的神经网络延伸出了深度学习的概念。

深度学习沿用了神经网络分层结构，但与神经网络又有不同。神经网络采用 BP 迭代算法来训练整个网络，随机设定初值计算当前网络的输出，利用该输出与真实样本的差值来改变前面各层的参数，直到收敛。该训练方法适合层数不多的情况，且随机的初值对系统也有一定的影响。深度学习含多层隐层，单纯采用传统的 BP 算法会导致模型时间复杂度很高，层层训练也会出现过拟合的情况，因此深度学习采用的是由上而下的监督训练方法，由下而上的非监督学习，常用的模型包括：自动编码器（Auto Encoder）、稀疏编码（Sparse Coding）、限制波尔兹曼机（Restricted Boltzmann Machine）、深度信任网络（Deep Belief Networks）和卷积神经网络（Convolutional Neural Networks），本文重点阐述卷积神经网络的原理及优势。

卷积神经网络是由 Hubel 和 Wiesel 根据视觉神经科学对动物大脑视觉皮层简单细胞核复杂细胞的反应设计而来的，是第一个真正成功训练多层网络结构的算法，主要应用于二维图像的识别或人脸识别。

一般而言，卷积神经网络包括两层结构，第一层是卷积层，通过多卷积滤波器组对输入图像的各种特征进行卷积操作，从而生成特征图。假设一个卷积核只提取一种特征，要得到图像的多个特征，就需要多个卷积核来提取不同的特征，每一层对应多张特征图。第二层是池化层，通过对图像某一个区域的特征取平均值实现聚合操作。首先为实现主体的可靠性先粗粒化各个特征的位置，然后计算某个特征图内局部块的最大值，通过平移行或列的块来获取相邻单元的输入数据，最后根据实际需要串联多个卷积、非线性变换和池化有监督地训练所有过滤器中的权值参数。

卷积神经网络的权值分享使得其结构与生物神经网络更类似，分层结构降低了模型的复杂度，并减少了权值的数量。其优点在于多维图像的输入可以直接作为网络的输入，与网络的拓扑结构能够很好地吻合，且特征提取和模式分类同时进行，避免了传统算法中复杂的特征提取和数据重建过程，从训练数据中进行学习，它可以直接处理灰度图片，能够直接用于处理基于图像的分类。

2. 深度学习的框架对比

目前，深度学习的主流框架主要包括 Caffe、TensorFlow、MXNet、Torch 和 Theano。

(1) Caffe。Caffe 于 2013 年由 Yangqing Jia 基于卷积神经网络编写和维护，是第一个主流的工业级深度学习工具，其开发语言是 C++，速度快，在计算机视觉领域很流行，但其缺点在于灵活性一般，而且对递归网络和语言建模的支持度较差。

(2) TensorFlow。TensorFlow 是由谷歌公司开发的第二代深度学习工具，它使用了向量运算的符号图方法，支持快速开发，且因其开源性和灵活性，深受开发人员喜爱。它的局限性在于速度较慢，内存占用相对较大，而且只能在 Linux 和 MAC OS 系统上开发。

(3) MXNet。MXNet 是另一款开源的深度学习工具，重在灵活性和高效性，同时强调提高内存使用率，其主要的开发语言也是 C++，是一款非常好用的国产开源工具。

(4) Torch。Torch 是 Facebook 力推的深度学习框架，具有较好的灵活性和高效性，主要的开发语言是 C 和 Lua，其核心的单元都已经过优化，使用者不需要花多余的时间在计算机优化方面，可以专注于实现算法。

(5) Theano。Theano 是 2008 年由蒙特利尔理工学

院研发，其最大的特点就是灵活性很高，适合学术研究，对递归网络和语言建模都有很好的支持，同时基于开发语言 Python 派生了大量的深度学习软件包，而且是在 Windows 系统上运行的深度学习框架，唯一的缺点就是其速度不高。

四、深度学习现阶段的研究及应用

语音处理方面，2012 年，微软公司研究人员建立深度神经网络与隐马尔科夫结合的混合模型，并将其应用于语音识别系统，大大降低了语音识别的错误率。2013 年，Graves 将长短时递归算法应用于语音处理，相较于传统算法更高效。在图像识别方面，Krizhevsky 等人提出了应用于 ImageNet 的 AlexNet 方法，基于大数据集将深度卷积神经网络应用于图像识别，将错误率降至 37.5%。2016 年，谷歌公司的 AlphaGo 通过人机对弈击败了人类所配置的视觉组成部分。其他方面，谷歌的 Ground Truth 团队使用深度神经网络从多国收集到的街景图中读取街道名，在 FSNS 数据集中达到了 84.2% 的准确率。

五、结语

深度学习的概念自被提出以来，其理论及学习架构都有很大的发展，并取得了很显著的成果，以谷歌公司为代表的一系列国内外研究人员做了大量的工作，比如在图像、语音和视频等方面的发展，甚至在自动驾驶、智能机器人等领域取得的成绩。基于深度学习的非监督理论学习方法也被广大研究人员不断改进。然而目前深度学习领域还有很多待解决的问题，比如理论支撑薄弱、无监督学习能力还有待提高等问题，要想机器学习真正做到像人类一样学习还有很长的路要走。