

专栏——脑网络分析与类脑计算

编者按：脑网络分析是大脑与复杂网络理论相结合的产物，脑网络由脑区节点和节点间的连接边两个基本要素构成。基于图论的方法，研究者们可以从微观尺度（神经元）、中间尺度（神经集群）和大尺度（脑区）等多个不同尺度上对大脑网络进行分析。近十年，世界各国都纷纷发展以脑网络研究为主题的大型科学计划，如人脑连接组计划（Human Connectome Project, HCP）等。类脑计算是指借鉴神经元间的互相合作的工作机制，模仿人脑功能进行计算，以类脑的方式形成大规模并行处理的网络。通过对大脑网络运行机制的了解，可以促进类脑技术的创新。同时，类脑计算也为脑科学数据处理提供了新的方法。本专栏主要围绕脑网络分析和类脑计算，这两个可以相互借鉴、相互融合的前沿领域而展开论述，主要介绍了深度学习算法在脑科学领域中的前沿应用，基于网络分析的人脑认知储备的定量预测，以及小鼠大脑个体结构网络的构建。



栏目主编：林岚

林岚，北京工业大学生物医学工程系副教授。从事运用神经影像学方法，研究与年龄相关的大脑认知退化方面的研究工作。2006年于美国亚利桑那州立大学获得生物工程博士学位。2006~2011年在美国亚利桑那州立大学和亚利桑那大学从事神经影像学方面的研究工作。运用现代神经图像处理与分析方法对健康人脑的认知储备及阿尔兹海默症的发生、发展过程进行研究。2012年加入北京工业大学生物医学工程系。共发表科技论文70余篇，其中16篇文章被SCI/EI检索收录。现系【美】神经科学协会会员、【美】McKnight大脑研究协会会员。多年来还作为课题负责人和主要参与者完成了一项北京市自然科学基金，两项亚利桑那州基金，多项NIA(美国国家老年研究院基金)项目和NIH(美国国家卫生研究院基金)项目。

深度学习在神经影像中的应用研究

Study on the Application of Deep Learning in Neuroimaging

田苗，林岚，张柏雯，吴水才
北京工业大学 生命科学与生物工程学院，北京 100124

[摘要] 深度学习是机器学习方法的一个重要分支，它通过基于多层神经网络的计算模型来进行学习。一个深度学习网络可以通过组合低层特征形成更抽象的高层特征，以发现数据的复杂的内在特征。由于深度卷积网络在图像处理中的出色表现，它已成为当前研究中应用最为广泛的一种深度网络。本文首先介绍了深度学习网络的结构特征和训练方法，分析了算法的优越性，之后进一步介绍了深度卷积网络，最后讨论了深度学习在神经影像领域的最新应用现状及其发展趋势。

[关键词] 深度学习；卷积神经网络；神经影像；多层感知器；深度神经网络

Abstract: Deep learning is an important branch of a broader family of machine learning methods which attempt to learn useful features from data by using a computational model with multilayer neural networks. A deep learning network can form more abstract high-layer features through combination of the lower layer features to find the complex intrinsic characteristics of the data. Convolutional neural network is one of most researched deep networks because of its excellent performance in processing images. This paper firstly introduced the structural characteristics and training methods of deep learning networks, followed by the convolutional neural network, and finally discussed the application and development trend of deep learning network in the field of neuroimaging.

Key words: deep learning; convolutional neural network; neuroimaging; multilayer perceptron; deep neural network

TIAN Miao, LIN Lan,
ZHANG Bai-wen, WU Shui-cai
College of Life Science and Bioengineering,
Beijing University of Technology, Beijing
100124, China

[中图分类号] R445.2 [文献标识码] A
doi: 10.3969/j.issn.1674-1633.2016.12.002
[文章编号] 1674-1633(2016)12-0004-06

学习是人类所具有的一种重要的智能行为,是人脑获取知识或技能的一个持续性变化的过程。机器学习^[1]最初源于人工智能,是研究计算机如何通过模拟实现人脑的学习行为的科学领域。它运用算法重组已有的知识结构,获取新的知识或技能,从而对现实世界的事物进行学习、测定或预测。机器学习算法按照学习方式主要分为两大类:非监督学习和监督学习。非监督式学习一般指数据是未知,样本没有被特别预定义,机器模型通过学习探索出隐含在数据训练样本集中的特殊模式。常见的非监督学习算法包括关联规则学习^[2]、聚类^[3]等。而监督式学习指的是输入数据被特别标记,有明确的标识,通过学习模型可将预测结果与实际标识结果进行比较,通过不断的调整模型参数,使得预测结果不断接近预期目标。监督式学习一般比非监督的预训练模型具有更好的性能,这归功于监督模型对数据集特性能够更好的编码,但需要大量的时间对数据进行手工标记。监督式学习是分类和回归问题的绝佳选择,其中常见算法有逻辑回归^[4]、决策树^[5]、支持向量机^[6]和人工神经网络(Artificial Neural Network, ANN)^[7-8]等。ANN已有几十年的历史,它的灵感源于我们对大脑中神经元工作模式的理解。感知器^[9]是将神经元用数学的方法表达出来,它是ANN中最基本的结构。多层感知器(Multilayer Perceptron, MLP)的基本架构包含输入层、输出层和隐含层,是只含有一层隐藏层节点的浅层模型。输入的特征向量经隐含层变换到达输出层,在输出层得到结果。各类ANN的结构具有很大的差异性,但其一般都具有两个主要特征:可调整的权重集合和模拟输入信号到神经元的非线性变换的能力。层数的增加提高了ANN网络的复杂度,使得优化函数越来越容易偏离全局最优,陷入局部最优解。如果利用有限数据来训练深层网络,深层网络的性能可能还不如较浅层的网络。同时,另一个不可忽略的问题是随着网络层数增加,梯度消失现象会变得更加严重。ANN中通常使用sigmoid函数作为神经元的输入输出函数,使用链式规则来计算梯度(通过微分)。在利用反向传播算法(Backward Propagation, BP)^[10]反向传播梯度时,随着网络层数的不断增加,梯度呈指数衰减,低层接收训练信号的能力不断减弱,层数过多时很有可能使低层几乎接收不到有效的训练信号。在深度学习未被提出之前,ANN的研究长期停留在浅层模型阶段。浅层的ANN模型虽然在过去为一些特定问题的解决提供了一种有效的手段,但仍存在很多的局限性,浅层ANN的研究相对沉寂。

收稿日期:2016-11-15

基金项目:国家科技支撑计划课题(2015BAI02B03);北京工业大学研究生科技基金(ykj-2016-00009);北京市自然科学基金资助项目(7143171)。

通讯作者:林岚,副教授。

通讯作者邮箱:lanlin@bjut.edu.cn

深度神经网络(Deep Neural Network, DNN)^[11-15]是在ANN基础上而发展起来的,包含多个隐藏层的ANN。从广义上说,DNN是一种特征学习方法,采用海量数据进行训练,通过一些简单的非线性模型将原始数据转变为低层特征,并逐步形成更加抽象的高层表示属性类别或特征,得到对数据更具本质的刻画,可以使得分类和预测更加有效准确。若通过足够多的转换组合,即使非常复杂的函数也可以被学习。单从结构上来说,全连接的DNN和多层感知机是没有任何本质区别的,DNN也可以理解为ANN一个更贴切的名称。但是如何有效地训练深层模型是长期以来困扰DNN发展的难题。2006年,Hinton等^[16]在《科学》上发表文章提出了深度置信网络(Deep Belief Network, DBN),使用非监督贪心逐层训练算法^[17]。深度网络在训练上的难题,可以通过无监督“逐层初始化”来有效克服,并将隐含层推动到了7层,这为解决与深层结构相关的优化难题带来希望。从此ANN在真正意义上有了“深度”的概念,深度学习的热潮也由此被揭开。在过去的几十年里,神经网络使用最多的一些例如sigmoid、tanh(z)等非线性函数,为了解决梯度消失的难题,修正线性单元(Rectified Linear Unit, ReLU)等传输函数的出现有效的加速了网络的收敛,让一个多层神经网络可以更快的学习。Lecun等^[17-18]提出的卷积神经网络(Convolutional Neural Network, CNN)将图像处理中的二维离散卷积运算与多层神经网络结合起来。卷积运算模仿人脑对信号的分级处理原理,对输入信息进行自动地特征提取。具体操作就是在原来的全连接层前面加入了部分连接的卷积层与降维层,利用空间相对关系减少参数数目以提高训练性能。实践中,DNN的解空间中存在着大量的鞍点,同时鞍点周围大部分曲面都是方向朝上,所以DNN算法就算是陷入了这些局部最小值,对结果也不会产生较大的干扰。如今,随着大数据时代的到来,大批量的神经影像数据为神经科学提供了研究素材,而大数据分析需要将这些影像数据运用深度学习等复杂的算法合成到更广泛的框架中进行分析。在本文中我们首先介绍了深度学习的一些关键的基本概念,随后介绍了图像处理中最为常用的CNN,最后展示并讨论了深度学习算法在神经影像领域的一些具体应用。

1 监督学习和反向传播算法对多层神经网络的训练

1.1 监督学习

深度学习的训练通常使用自下而上的非监督学习和自顶向下的监督学习算法。对于非监督学习方法,并不需要使用带标签的数据来检测特征。通常先利用非监督学习方法对每一隐藏层进行逐步训练去学习特征,每次单独训练一层,并将低层的训练结果作为更高一层的输入,最后一个输出层被添加到该网络的顶部,这时改用监督学习的BP

算法从上到下进行微调去不断优化模型。

在 DNN 中,最常用的形式还是监督学习。它的工作方式是当向模型输入训练数据时,模型对训练实例进行预测,通过计算预测结果与期望结果间的预测误差,根据预测误差更新模型的权重,以减少下一个预测的误差。在典型的深度学习系统中,可能有数以百万计的带有标签的样本和权值被用来训练网络。为了保证调整权重向量的正确性,学习算法需要计算每个权值的梯度向量,表示随着权值变化的一个微量,误差会随之增加或减少的量,然后在梯度矢量的相反方向上对权值向量进行调整。在实际应用中,随机梯度下降(Stochastic Gradient Descent Algorithm, SGD)的算法是一种最常用的方法。它通过计算样本的平均梯度来调整相应权值的大小,可以基于小的样本集合来重复这个过程用以训练网络。模型中每输入一个新样本,该算法就迭代更新一次,不需要训练集中所有的数据参与每次迭代,单次迭代的结果不一定是全局最优,但是多次迭代的最终目标是寻求全局最优解,所以最终的结果往往是在全局最优解附近。这种迭代过程不断重复,直到目标函数停止增长。同其它优化技术相比,SGD 具有极快的训练速度,大大缩短了训练时间。通常在训练完成后,系统的性能就因此通过测试集(用于训练的数据样本)被确定了,这时可以通过未经过训练的新样本来评定算法的泛化能力。

1.2 BP算法

BP 算法是 Rumelhart 等^[19]于 1986 年针对具有非线性连续转移函数的多层前馈网络提出的一种训练算法,在多层网络的训练中举足轻重,是复合函数的链式法则的具体应用。深度学习的训练目标就是通过调整网络中的每一个权值来使得目标函数达到最小。目标函数也可以看成是由所有待求权值为自变量的复合函数。BP 算法正是用来求解这种多层复合函数的所有变量的偏导数的利器,目标函数对于某层输入的导数(或者梯度)可以通过向后传播的方式对该层输出(或者下一层输入)的导数求得。BP 算法可以通过传播梯度的方式被重复的应用于 DNN 的每一层。它的学习过程主要分为两个阶段(图 1):信号的正向传播阶段和误差的反向传播阶段。当信号正向传播时,样本数据从输入层进入网络,经各个隐层逐层处理后,从输出层传出。若传出后的实际值与期望值不符,则转入误差的反向传播阶段,即将输出误差以某种形式通过隐层向输入层逐层反向传播,并分摊到各层的所有神经元,从而获得各层神经元的误差信号,根据此误差信号来修正各神经元的权值。正向传播与误差反向传播阶段周而复始的交替进行,因此权值在此过程中得到不断地调整,直到网络的输出误差减少到可接受的水平,或达到预设定的学习次数。

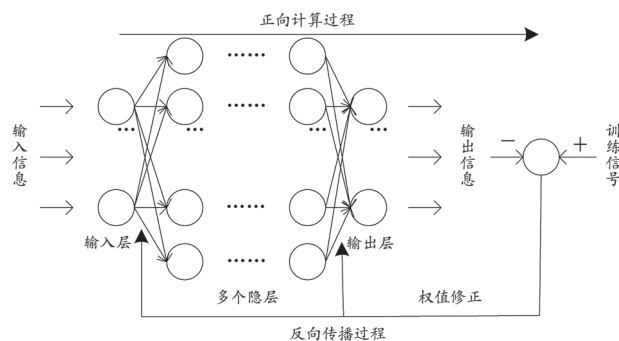


图1 BP算法原理图

2 卷积神经网络

在非周期性连接的 ANN 结构中,前馈网络是最简单的形式。前馈网络的整个网络中无反馈,信息从输入节点开始单向传输,计算前一层神经元输入数据的权值的和,然后把权值和传给一个非线性激活函数,如 ReLU 激活函数。ReLU 激活函数因为其分段线性性质,导致其前传、后传、求导都是分段线性,可以让一个深度网络跳过无监督预训练的过程,直接进行有监督的训练。在全连接 DNN 的结构里,其下层神经元和所有的上层神经元都能够形成连接,由此也会带来参数数量膨胀的潜在问题,海量权重在训练过程中不仅容易过拟合,而且极易陷入局部最优。

与其它大多数深度前馈式 ANN 相比,CNN 是深度学习算法中应用最成功的算法之一,且更易于训练,比全连接的 DNN 有更好的泛化性能,在计算机视觉和图像识别领域都得到了很好的应用^[20-21]。CNN 作为一种特殊的图像识别方式,其中的卷积及池化层的灵感直接来源于视觉皮层中的细胞排列。视觉皮层中存在两类细胞:一类细胞是简单细胞,主要用于识别物体边缘的模式,还有一类复杂细胞,具有位置不变性,可以表达模式的具体位置,这些细胞是以层级结构形成了视觉回路。CNN 的设计中包含了四个关键的概念-局部连接、权值共享、池化^[22]以及多网络层,网络由三个重要的部分构成:卷积层、池化层、全连接层。卷积层和池化层交替构成了卷积神经网络的基本架构。卷积层运用若干个卷积核完成对输入数据的特征提取,其作用是探测上一层特征的局部连接。卷积核并不一次性解析所有的训练数据,每一个节点仅与其临近区域的神经元进行局部连接,一旦该局部特征被提取,它与其它特征间的位置关系也随之确定。CNN 的这种结构归功于两点:首先,在图像数据中相邻的值经常是高度相关的,可以形成易于被探测到的具有区分性的局部特征;其次,不同位置局部统计特征不太相关,所以不同位置的单元可以共享权值,并可以探测相同的样本。正因为神经元共享权值,因而大大减少了网络自由参数的个数,降低了网络参数选择的复杂度,缩减了输入数据的规模。应用卷积核之后得到的结果被称为特征图,特征图的数目和卷积核的数目相等,在一个特征图中的全部单元享用相同的卷积核,不同

层的特征图使用不同的卷积核。池化是一种滤除细节的方法,可以把相似的特征合并起来,应用中一般采用最大池化的方法。当输入数据在前一层中的位置有变化的时候,池化操作让这些特征表示对这些变化具有鲁棒性,对输入样本有较高的畸变容忍能力。卷积层和池化层不断交替进行,后面再加上一个更多卷积和全连接层,从而实现相应的功能。BP算法在CNN上的应用和在一般的深度网络上的用法相同,可以让所有的在过滤器中的权值得到训练。

除了CNN以外,DNN还有很多其它结构。传统DNN无法对时间序列上的变化进行建模。然而,样本出现的时间顺序在自然语言处理^[23]、语音识别^[24]、手写体识别^[25]等应用中都起着非常重要的作用。为了适应这种需求,就出现了循环神经网络(Recurrent Neural Networks,RNN)^[26]。它的基本原理就是神经元的输出可以在下一个时间直接作用到自身。事实上,无论是哪种网络,在实际应用中它们常常都是混合使用的,比如CNN和RNN在上层输出之前往往会接上全连接层,因此很难说某个DNN到底属于哪个类别。不难想象随着深度学习热度的延续,更灵活的组合方式、更多的网络结构将会被发展出来。

3 深度学习在神经影像中的应用

近年来,深度学习方法在机器学习的很多领域都取得了显著的改进。与其它方法相比,深度学习的方法有两个重要的优点:首先,深度学习是一种数据驱动的自动学习特征的方法,这一重要的能力大幅度减少了当存在大量特征时,相关特征选择中存在的主观性;其次,深度学习模型深度与传统的浅层模型相比,它应用了具有层次性结构的非线性层,从而能够更好的建立非常复杂的数据模式的模型。

神经影像学的一个重要目标是基于非侵入性的神经影像测量数据来更好地理解大脑神经系统的工作原理。在大脑成像领域,数据驱动的特征学习方法已经被运用了多年。与传统的数据驱动方法相比较,深度学习的方法在进行特征选择时并不一定需要预处理步骤。近年来,DBN被用来分析结构磁共振(Magnetic Resonance Imaging,MRI)和功能磁共振数据(Functional MRI,fMRI)。Plis等^[27]发现深度学习的方法可以揭示高维神经影像数据间的关系,并生成具有实际生理意义的特征。Hjelm等^[28]应用限制玻尔兹曼机(Restricted Boltzmann Machine,RBM)的概率模型来确定fMRI数据的内在功能网络,RBM通过拟合数据的概率分布模型,将线性因素从fMRI中分离出来。更重要的是,更深层的级联模型可以用于多模态影像的扩展,从而解决了仅利用传统矩阵分解模型一直无法解决的问题。

脑结构的自动分割是结构和功能图像分析中的一项重要任务。传统的脑组织分割一般是采用基于图谱和高精度配准的方法^[29]。这些方法往往由于高精度的配准,耗时较

长,同时对于一些个体差异大的小组织及肿瘤造成的病变组织,效果并不是很显著^[30]。Choi等^[31]提出了一种基于级联CNN(全局CNN和局部CNN)的深度网络用于大脑中纹状体的分割。全局CNN被用于确定纹状体的近似位置,裁剪后的MRI和全局CNN的输出图像被用于作为局部CNN的输入,局部CNN被用于细分割。结果显示该方法较传统基于Freesurfer的分割方法精度更高,速度更快。高危组织的分割是脑肿瘤外科治疗计划中的关键一步,Dolz等^[32]提出了一种基于堆叠降噪自动编码器(Stacking Denoising Auto-Encoders,SDAE)的半监督深度学习方法。在这个混合体系结构中,首先用无监督学习的方法预训练网络参数,随后用有监督学习的方法对深度网络的参数进行更精细的调节,达到预期目标。最终结果显示该方法与专家手工划分的结果相近似,但它却可以大幅度的削减分割的时间。Kleesiek等^[33]采用3-D CNN对多模态的脑神经影像进行了非脑组织去除,结果显示该方法显著优于六种广为使用的非脑组织去除工具。婴儿MRI图像中的脑组织分割,由于脑灰质、白质的灰度分布近似,是神经影像处理中的一大难题。Zhang等^[34]采用CNN对婴儿MRI图像进行了组织分割,结果显示该方法显著优于其它分割方法。

近年来,深度学习的方法在处理神经影像数据方面发展迅速,运用这些数据对患者进行分类及预测,使得深度学习已经成为神经影像领域中炙手可热的方法。Suk等^[35]提出了基于堆栈自动编码器(Stacked Auto Encoder,SAE)的深度学习模型,并将其用于AD和MCI的分类。该方法首先分别基于MRI、PET和CSF的低级特征,采用SAE的非监督预训练和有监督参数微调,得到潜在特征,随后基于多任务学习选出可以代表分类和临床评分的特征,最后通过支持向量机进行预测与分类,结果表明该算法可以较好的区分AD和MCI患者。Kawahara等^[36]提出了一种改进的CNN模型BrainNetCNN,它被用于预测早产儿脑网络的神经发育。与传统的基于图像的CNN(空间局部卷积)不同,BrainNetCNN是为脑网络数据专门设计的深度学习框架。它利用结构性大脑网络的拓扑位置,具有边到边,边到节点,节点到图的卷积核的独特设计。预测认知及运动发育的结果表明,BrainNetCNN框架优于其它同类方法。由于CNN在层内共享权重,从而减少了学习的参数的数量,BrainNetCNN在相同数量的模型参数下优于完全连接的神经网络。此外,深度学习还被用在一些其它疾病的诊断中,如Kim等^[37]基于静息态fMRI数据,采用SAE预训练的DNN对健康对照组和精神分裂症患者进行分类。结果显示该DNN远优于支持向量机的分类效果。

对于从高维神经影像数据中提取隐藏的模式这类问题,深度学习也是一个非常有前途的工具^[38]。Ithapu等^[39]将3-D CNN模型用于完成和整合多模态神经影像学数据,

预测从 MRI 数据丢失的 PET 特征模式。模型包括两个隐藏层,采用一个体数据 MRI 模态作为输入, PET 模态作为输出,不同模态间的非线性关系可通过大量训练捕获。结果表明,预测的 PET 数据实现了与真实 PET 图像相似的分类性能。在实验中随机抽取一半的受试者样本数据来训练模型,然后使用该模型来预测剩余受试者的 PET 图像。3-D CNN 模型的结果与真实 PET 图像的结果相当,即该模型可以成功地提取 MRI 和 PET 图像之间的高度非线性关系。

4 发展前景

多年来,大脑神经科学的发展揭示了神经计算的基本原理,这些神经计算的基本原理也被广泛应用于神经网络的构建,其中深度学习算法的提升与改进大大促进了神经科学的发展。现在,深度学习已成为传统机器学习策略的有效补充,并在神经影像领域得到了越来越广泛的应用。即便如此,它也还存在一些不足之处:首先,深度学习算法在实际应用过程中往往是若干个模型间的组合,从而达到结构优化。但是不同的神经影像的分析一般代表着不同的实际问题,因此针对不同模态的影像特征往往需要设计完全不同的模型来适应,这对算法的构建就提出了更高的要求。其次,深度学习模型的搭建需要调整大量的参数,只有非常庞大的训练数据,才能防止过度拟合,达到良好的性能。如果将 DNN 应用于神经影像,目前面对的问题是满足深度网络需求的大规模的神经影像训练数据集还相当的少见。ADNI^[40-41]、HCP^[42] 这些神经影像大科学计划的发展和新计划的产生,在一定程度上可以缓解这个问题。另外一种可行的代替方法是迁移学习,可以在一定程度上减少对神经影像数据集规模的需求。基于普通的大规模图像数据集如 ImageNet^[43] (包含了大约 100 万张自然图像和 1000 个标签/分类)对网络进行预训练,然后使用这一训练的网络的权重作为初始设置对参数进行调优。迁移学习策略的效率取决于神经影像数据集与 ImageNet 图像数据集的相似度。前面几层网络特征需要更加具有通用性,并不与神经影像的分类任务直接相关。当然,如果训练出的特征具有特异性,转移学习就未必有好的效果。预计随着算法的持续改进以及相关软硬件的发展,深度学习的方法将会适用于越来越多的神经影像研究问题。

[参考文献]

- [1] 刘更代,潘志庚,程熙,等.人体运动合成中的机器学习技术综述[J].计算机辅助设计与图形学学报,2010,22(9):1619-1627.
- [2] 王曙燕,周明全,耿国华,等.医学图像的关联规则挖掘方法研究[J].计算机应用,2005,25(6):1408-1409.
- [3] 贺玲,吴玲达,蔡益朝,等.数据挖掘中的聚类算法综述[J].计算机应用研究,2007,24(1):10-13.
- [4] Brimacombe M. Large sample convergence diagnostics for likelihood based inference: Logistic regression[J]. *Stat Methodol*, 2016,33:114-130.
- [5] Kim K. A hybrid classification algorithm by subspace partitioning through semi-supervised decision tree[J]. *Pattern Recogn*, 2016, 60:157-163.
- [6] Chen D, Tian Y, Liu X. Structural Nonparallel Support Vector Machine for Pattern Recognition[J]. *Pattern Recogn*, 2016, 60: 296-305.
- [7] Yu X, Ye C, Xiang L. Application of artificial neural network in the diagnostic system of osteoporosis[J]. *Neurocomputing*, 2016, 214:376-381.
- [8] Lin L, Jin C, Fu Z, et al. Predicting healthy older adult's brain age based on structural connectivity networks using artificial neural networks[J]. *Comput Meth Prog Bio*, 2016, 125:8-17.
- [9] Pastor M, Song J, Hoang D T, et al. Minimal perceptrons for memorizing complex patterns[J]. *Physica A*, 2015, 462:31-37.
- [10] Lillicrap TP, Cownden D, Tweed DB, et al. Random synaptic feedback weights support error backpropagation for deep learning[J]. *Nat Commun*, 2016, 7:13276.
- [11] Du J, Xu Y. Hierarchical deep neural network for multivariate regression[J]. *Pattern Recogn*, 2016, 63:149-157.
- [12] Ekins S. The Next Era: Deep Learning in Pharmaceutical Research[J]. *Pharm Res*, 2016, 33(11):2594-2603.
- [13] Gawehn E, Hiss JA, Schneider G. Deep Learning in Drug Discovery[J]. *Mol Inform*, 2016, 35(1):3-14.
- [14] Angermueller C, Pärnamäa T, Parts L. Deep learning for computational biology[J]. *Mol Syst Biol*, 2016, 12(7):878.
- [15] Mamoshina P, Vieira A, Putin E, et al. Applications of Deep Learning in Biomedicine[J]. *Mol Pharm*, 2016, 13(5):1445-1454.
- [16] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786):504-507.
- [17] LeCun Y, Boser B, Denker JS, et al. Handwritten digit recognition with a back-propagation network[C]. Proceedings of Advances in Neural Information Processing Systems, 1990:396-404.
- [18] LeCun Y, Boser B, Denker JS, et al. Backpropagation applied to handwritten zip code recognition[J]. *Neural Comput*, 1989, 1(4): 541-551.
- [19] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors[J]. *Cognitive Modeling*, 1988, 5(3):1.
- [20] Yang W, Chen Y, Liu Y, et al. Cascade of multi-scale convolutional neural networks for bone suppression of chest radiographs in gradient domain[J]. *Med Image Anal*, 2017, 35:421-433.
- [21] Ypsilantis PP, Siddique M, Sohn HM, et al. Predicting Response to Neoadjuvant Chemotherapy with PET Imaging Using Convolutional Neural Networks[J]. *PLoS One*, 2015, 10(9):e0137036.

- [22] Shi Z, Ye Y, Wu Y. Rank-based pooling for deep convolutional neural networks[J]. *Neural Networks*, 2016, 83: 21-31.
- [23] Gudivada V, Rao D, Raghavan V. Big Data-Driven Natural Language-Processing Research and Applications[J]. *Big Data Analytics*, 2015, 33: 203.
- [24] Maas AL, Qi P, Xie Z, et al. Building DNN acoustic models for large vocabulary speech recognition[J]. *Comput Speech Lang*, 2015, 41: 195-213.
- [25] Elleuch M, Maalej R, Kherallah M. A New Design Based-SVM of the CNN Classifier Architecture with Dropout for Offline Arabic Handwritten Recognition[J]. *Procedia Computer Science*, 2016, 80: 1712-1723.
- [26] Mikolov T, Karafiát M, Burget L, et al. Recurrent neural network based language model[C]. Conference of the International Speech Communication Association, 2010: 1045-1048.
- [27] Plis SM, Hjelm DR, Salakhutdinov R, et al. Deep learning for neuroimaging: a validation study[J]. *Front Neurosci*, 2014, 8(8): 229-229.
- [28] Hjelm RD, Calhoun VD, Salakhutdinov R, et al. Restricted Boltzmann machines for neuroimaging: an application in identifying intrinsic networks[J]. *NeuroImage*, 2014, 96: 245-260.
- [29] Fu Z, Lin L, Jin C. Symmetric image normalization for mouse brain magnetic resonance microscopy[C]. International Conference on Advances in Mechanical Engineering and Industrial Informatics, 2016: 941945.
- [30] Lin L, Wu S. An automated template-based adaptive threshold approach for measuring ventricular volume enlargement in mouse brain MR microscopy[J]. *J Microsc*, 2012, 248(3): 260-265.
- [31] Choi H, Jin KH. Fast and robust segmentation of the striatum using deep convolutional neural networks[J]. *J Neurosci Meth*, 2016, 274: 146-153.
- [32] Dolz J, Betrouni N, Quidet M, et al. Stacking denoising auto-encoders in a deep network to segment the brainstem on MRI in brain cancer patients: A clinical study[J]. *Comput Med Imag Grap*, 2016, 52: 8-18.
- [33] Kleesiek J, Urban G, Hubert A, et al. Deep MRI brain extraction: A 3D convolutional neural network for skull stripping[J]. *Neuroimage*, 2016, 129: 460-469.
- [34] Zhang W, Li R, Deng H, et al. Deep convolutional neural networks for multi-modality iso-intense infant brain image segmentation[J]. *Neuroimage*, 2015, 108: 214-224.
- [35] Suk HI, Shen D. Deep Learning-Based Feature Representation for AD/MCI Classification[M]. Medical Image Computing and Computer-Assisted Intervention-MICCAI 2013. Springer: Berlin Heidelberg, 2013: 583-590.
- [36] Kawahara J, Brown CJ, Miller SP, et al. BrainNetCNN: Convolutional Neural Networks for Brain Networks; Towards Predicting Neurodevelopment[J]. *Neuroimage*, 2016.
- [37] Kim J, Calhoun VD, Shim E, et al. Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: Evidence from whole-brain resting-state functional connectivity patterns of schizophrenia[J]. *Neuroimage*, 2016, 124(PtA): 127-146.
- [38] Kriegeskorte N. Deep neural networks: A new framework for modeling biological vision and brain information processing[J]. *Annual Review of Vision Science*, 2015, 1: 417-446.
- [39] Ithapu VK, Singh V, Okonkwo OC, et al. Imaging-based enrichment criteria using deep learning algorithms for efficient clinical trials in mild cognitive impairment[J]. *Alzheimers Dement*, 2015, 11(12): 1489-1499.
- [40] 齐志刚, 李坤成, 王军. 为更早识别阿尔茨海默病: 阿尔茨海默病神经影像学计划简介[J]. 中国现代神经疾病杂志, 2014, 14(4): 277-280.
- [41] 靳聪, 林岚, 付振荣, 等. MRI在阿尔茨海默氏病中的应用研究进展[J]. 智慧健康, 2015, 1(1): 52-56.
- [42] Glasser MF, Smith SM, Marcus DS, et al. The Human Connectome Project's neuroimaging approach[J]. *Nat Neurosci*, 2016, 19(9): 1175-1187.
- [43] Deng J, Dong W, Socher R, et al. Imagenet: A large-scale hierarchical image database[C]. IEEE Int Conf Computer Vision and Pattern Recognition, 2009: 248-255.

