# 深度学习最新进展综述

**摘要**：深度学习是机器学习和人工智能研究的最新趋势之一。它也是当今最流行的科学研究趋势之一。深度学习方法为计算机视觉和机器学习带来了革命性的进步。新的深度学习技术正在不断诞生，超越最先进的机器学习甚至是现有的深度学习技术。近年来，全世界在这一领域取得了许多重大突破。由于深度学习正快度发展，导致了它的进展很难被跟进，特别是对于新的研究者。在本文中，我们将简要讨论近年来关于深度学习的最新进展。

## 1. 引言

「深度学习」（DL）一词最初在1986被引入机器学习（ML），后来在2000年时被用于人工神经网络（ANN）。深度学习方法由多个层组成，以学习具有多个抽象层次的数据特征。DL方法允许计算机通过相对简单的概念来学习复杂的概念。对于人工神经网络（ANN），深度学习（DL）（也称为分层学习（Hierarchical Learning））是指在多个计算阶段中精确地分配信用，以转换网络中的聚合激活。为了学习复杂的功能，深度架构被用于多个抽象层次，即非线性操作；例如ANNs，具有许多隐藏层。用准确的话总结就是，深度学习是机器学习的一个子领域，它使用了多层次的非线性信息处理和抽象，用于有监督或无监督的特征学习、表示、分类和模式识别。

深度学习即表征学习是机器学习的一个分支或子领域，大多数人认为近代深度学习方法是从2006开始发展起来的。本文是关于最新的深度学习技术的综述，主要推荐给即将涉足该领域的研究者。本文包括DL的基本思想、主要方法、最新进展以及应用。

综述论文是非常有益的，特别是对某一特定领域的新研究人员。一个研究领域如果在不久的将来及相关应用领域中有很大的价值，那通常很难被实时跟踪到最新进展。现在，科学研究是一个很有吸引力的职业，因为知识和教育比以往任何时候都更容易分享和获得。对于一种技术研究的趋势来说，唯一正常的假设是它会在各个方面有很多的改进。几年前对某个领域的概述，现在可能已经过时了。

考虑到近年来深度学习的普及和推广，我们简要概述了深度学习和神经网络（NN），以及它的主要进展和几年来的重大突破。我们希望这篇文章将帮助许多新手研究者在这一领域全面了解最近的深度学习的研究和技术，并引导他们以正确的方式开始。同时，我们希望通过这项工作，向这个时代的顶级DL和ANN研究者们致敬：Geoffrey Hinton（Hinton）、Juergen Schmidhuber（Schmidhuber）、Yann LeCun（LeCun）、Yoshua Bengio（Bengio）和许多其他研究学者，他们的研究构建了现代人工智能（AI）。跟进他们的工作，以追踪当前最佳的DL和ML研究进展对我们来说也至关重要。

在本论文中，我们首先简述过去的研究论文，对深度学习的模型和方法进行研究。然后，我们将开始描述这一领域的最新进展。我们将讨论深度学习（DL）方法、深度架构（即深度神经网络（DNN））和深度生成模型（DGM），其次是重要的正则化和优化方法。此外，用两个简短的部分对于开源的DL框架和重要的DL应用进行总结。我们将在最后两个章节（即讨论和结论）中讨论深入学习的现状和未来。

## 2. 相关研究

在过去的几年中，有许多关于深度学习的综述论文。他们以很好的方式描述了DL方法、方法论以及它们的应用和未来研究方向。这里，我们简要介绍一些关于深度学习的优秀综述论文。

Young等人（2017）讨论了DL模型和架构，主要用于自然语言处理（NLP）。他们在不同的NLP领域中展示了DL应用，比较了DL模型，并讨论了可能的未来趋势。

Zhang等人（2017）讨论了用于前端和后端语音识别系统的当前最佳深度学习技术。

Zhu等人（2017）综述了DL遥感技术的最新进展。他们还讨论了开源的DL框架和其他深度学习的技术细节。

Wang等人（2017）以时间顺序的方式描述了深度学习模型的演变。该短文简要介绍了模型，以及在DL研究中的突破。该文以进化的方式来了解深度学习的起源，并对神经网络的优化和未来的研究做了解读。

Goodfellow等人（2016）详细讨论了深度网络和生成模型，从机器学习（ML）基础知识、深度架构的优缺点出发，对近年来的DL研究和应用进行了总结。

LeCun等人（2015）从卷积神经网络（CNN）和递归神经网络（RNN）概述了深度学习（DL）模型。他们从表征学习的角度描述了DL，展示了DL技术如何工作、如何在各种应用中成功使用、以及如何对预测未来进行基于无监督学习（UL）的学习。同时他们还指出了DL在文献目录中的主要进展。

Schmidhuber（2015）从CNN、RNN和深度强化学习(RL)对深度学习做了一个概述。他强调了序列处理的RNN，同时指出基本DL和NN的局限性，以及改进它们的技巧。

Nielsen(2015)用代码和例子描述了神经网络的细节。他还在一定程度上讨论了深度神经网络和深度学习。

Schmidhuber(2014)讨论了基于时间序列的神经网络、采用机器学习方法进行分类，以及在神经网络中使用深度学习的历史和进展。

Deng和Yu(2014)描述了深度学习类别和技术，以及DL在几个领域的应用。

Bengio(2013)从表征学习的角度简要概述了DL算法，即监督和无监督网络、优化和训练模型。他聚焦于深度学习的许多挑战，例如：为更大的模型和数据扩展算法，减少优化困难，设计有效的缩放方法等。

Bengio等人(2013)讨论了表征和特征学习即深度学习。他们从应用、技术和挑战的角度探讨了各种方法和模型。

Deng(2011)从信息处理及相关领域的角度对深度结构化学习及其架构进行了概述。

Arel等人(2010)简要概述了近年来的DL技术。

Bengio(2009)讨论了深度架构，即人工智能的神经网络和生成模型。

最近所有关于深度学习（DL）的论文都从多个角度讨论了深度学习重点。这对DL的研究人员来说是非常有必要的。然而，DL目前是一个蓬勃发展的领域。在最近的DL概述论文发表之后，仍有许多新的技术和架构被提出。此外，以往的论文从不同的角度进行研究。我们的论文主要是针对刚进入这一领域的学习者和新手。为此，我们将努力为新研究人员和任何对这一领域感兴趣的人提供一个深度学习的基础和清晰的概念。

## 3. 最新进展

在本节中，我们将讨论最近从机器学习和人工神经网络(ANN)的中衍生出来的主要深度学习(DL)方法，人工神经网络是深度学习最常用的形式。

### 3.1 深度架构的演变

人工神经网络(ANN)已经取得了长足的进步，同时也带来了其他的深度模型。第一代人工神经网络由简单的感知器神经层组成，只能进行有限的简单计算。第二代使用反向传播，根据错误率更新神经元的权重。然后支持向量机(SVM)浮出水面，在一段时间内超越ANN。为了克服反向传播的局限性，人们提出了受限玻尔兹曼机（RBM），使学习更容易。此时其他技术和神经网络也出现了，如前馈神经网络(FNN)、卷积神经网络(CNN)、循环神经网络(RNN)等，以及深层信念网络、自编码器等。从那时起，为实现各种用途，ANN在不同方面得到了改进和设计。

Schmidhuber(2014)、Bengio(2009)、Deng和Yu(2014)、Goodfellow等人(2016)、Wang等人(2017)对深度神经网络(DNN)的进化和历史以及深度学习(DL)进行了详细的概述。在大多数情况下，深层架构是简单架构的多层非线性重复，这样可从输入中获得高度复杂的函数。

## 4. 深度学习方法

深度神经网络在监督学习中取得了巨大的成功。此外，深度学习模型在无监督、混合和强化学习方面也非常成功。

### 4.1 深度监督学习

监督学习应用在当数据标记、分类器分类或数值预测的情况。LeCun等人(2015)对监督学习方法以及深层结构的形成给出了一个精简的解释。Deng和Yu(2014)提到了许多用于监督和混合学习的深度网络，并做出解释，例如深度堆栈网络(DSN)及其变体。Schmidthuber (2014)的研究涵盖了所有神经网络，从早期神经网络到最近成功的卷积神经网络(CNN)、循环神经网络(RNN)、长短期记忆(LSTM)及其改进。

### 4.2 深度无监督学习

当输入数据没有标记时，可应用无监督学习方法从数据中提取特征并对其进行分类或标记。LeCun等人(2015)预测了无监督学习在深度学习中的未来。Schmidthuber(2014)也描述了无监督学习的神经网络。Deng和Yu(2014)简要介绍了无监督学习的深度架构，并详细解释了深度自编码器。

### 4.3 深度强化学习

强化学习使用奖惩系统预测学习模型的下一步。这主要用于游戏和机器人，解决平常的决策问题。Schmidthuber(2014)描述了强化学习(RL)中深度学习的进展，以及深度前馈神经网络(FNN)和循环神经网络(RNN)在RL中的应用。Li(2017)讨论了深度强化学习(DeepReinforcement Learning, DRL)、它的架构(例如Deep Q-Network, DQN)以及在各个领域的应用。

Mnih等人(2016)提出了一种利用异步梯度下降进行DNN优化的DRL框架。

van Hasselt等人(2015)提出了一种使用深度神经网络(deep neural network, DNN)的DRL架构。

## 5. 深度神经网络

在本节中，我们将简要地讨论深度神经网络(DNN)，以及它们最近的改进和突破。神经网络的功能与人脑相似。它们主要由神经元和连接组成。当我们说深度神经网络时，我们可以假设有相当多的隐藏层，可以用来从输入中提取特征和计算复杂的函数。Bengio(2009)解释了深度结构的神经网络，如卷积神经网络(CNN)、自编码器(AE)等及其变体。Deng和Yu(2014)详细介绍了一些神经网络架构，如AE及其变体。Goodfellow等(2016)对深度前馈网络、卷积网络、递归网络及其改进进行了介绍和技巧性讲解。Schmidhuber(2014)提到了神经网络从早期神经网络到最近成功技术的完整历史。

### 5.1 深度自编码器

自编码器(AE)是神经网络(NN)，其中输出即输入。AE采用原始输入，编码为压缩表示，然后解码以重建输入。在深度AE中，低隐藏层用于编码，高隐藏层用于解码，误差反向传播用于训练.。

#### 5.1.1 变分自编码器

变分自动编码器(VAE)可以算作解码器。VAE建立在标准神经网络上，可以通过随机梯度下降训练(Doersch,2016)。

#### 5.1.2 多层降噪自编码器

在早期的自编码器(AE)中，编码层的维度比输入层小（窄）。在多层降噪自编码器(SDAE)中，编码层比输入层宽(Deng and Yu, 2014)。

#### 5.1.3 变换自编码器

深度自动编码器(DAE)可以是转换可变的，也就是从多层非线性处理中提取的特征可以根据学习者的需要而改变。变换自编码器(TAE)既可以使用输入向量，也可以使用目标输出向量来应用转换不变性属性，将代码引导到期望的方向(Deng and Yu,2014)。

### 5.2 深度卷积神经网络

四种基本思想构成了卷积神经网络(CNN)，即：**局部连接、共享权重、池化和多层使用。CNN的第一部分由卷积层和池化层组成，后一部分主要是全连接层**。卷积层检测特征的局部连接，池层将相似的特征合并为一个。CNN在卷积层中使用卷积而不是矩阵乘法。

Krizhevsky等人(2012)提出了一种深度卷积神经网络(CNN)架构，也称为AlexNet，这是深度学习(Deep Learning, DL)的一个重大突破。网络由5个卷积层和3个全连接层组成。该架构采用图形处理单元(GPU)进行卷积运算，采用线性整流函数(ReLU)作为激活函数，用Dropout来减少过拟合。

Iandola等人(2016)提出了一个小型的CNN架构，叫做「SqueezeNet」。

Szegedy等人(2014)提出了一种深度CNN架构，名为Inception。Dai等人(2017)提出了对Inception-ResNet的改进。

Redmon等人(2015)提出了一个名为YOLO(You Only Look Once)的CNN架构，用于均匀和实时的目标检测。

Zeiler和Fergus(2013)提出了一种将CNN内部激活可视化的方法。

Gehring等人(2017)提出了一种用于序列到序列学习的CNN架构。

Bansal等人(2017)提出了PixelNet，使用像素来表示。

Goodfellow等人(2016)解释了CNN的基本架构和思想。Gu等人(2015)对CNN的最新进展、CNN的多种变体、CNN的架构、正则化方法和功能以及在各个领域的应用进行了很好的概述。

#### 5.2.1 深度最大池化卷积神经网络

最大池化卷积神经网络(MPCNN)主要对卷积和最大池化进行操作，特别是在数字图像处理中。MPCNN通常由输入层以外的三种层组成。卷积层获取输入图像并生成特征图，然后应用非线性激活函数。最大池层向下采样图像，并保持子区域的最大值。全连接层进行线性乘法。在深度MPCNN中，在输入层之后周期性地使用卷积和混合池化，然后是全连接层。

#### 5.2.2 极深的卷积神经网络

Simonyan和Zisserman(2014)提出了非常深层的卷积神经网络(VDCNN)架构，也称为VGG Net。VGG Net使用非常小的卷积滤波器，深度达到16-19层。Conneau等人(2016)提出了另一种文本分类的VDCNN架构，使用小卷积和池化。他们声称这个VDCNN架构是第一个在文本处理中使用的，它在字符级别上起作用。该架构由29个卷积层组成。

### 5.3 网络中的网络

Lin等人(2013)提出了网络中的网络(Network In Network,NIN)。NIN以具有复杂结构的微神经网络代替传统卷积神经网络(CNN)的卷积层。它使用多层感知器(MLPConv)处理微神经网络和全局平均池化层，而不是全连接层。深度NIN架构可以由NIN结构的多重叠加组成。

### 5.4 基于区域的卷积神经网络

Girshick等人(2014)提出了基于区域的卷积神经网络(R-CNN)，使用区域进行识别。R-CNN使用区域来定位和分割目标。该架构由三个模块组成：定义了候选区域的集合的类别独立区域建议，从区域中提取特征的大型卷积神经网络(CNN)，以及一组类特定的线性支持向量机(SVM)。

#### 5.4.1 Fast R-CNN

Girshick(2015)提出了快速的基于区域的卷积网络(Fast R-CNN)。这种方法利用R-CNN架构能快速地生成结果。Fast R-CNN由卷积层和池化层、区域建议层和一系列全连接层组成。

#### 5.4.2 Faster R-CNN

Ren等人(2015)提出了更快的基于区域的卷积神经网络(Faster R-CNN)，它使用区域建议网络(Region Proposal Network, RPN)进行实时目标检测。RPN是一个全卷积网络，能够准确、高效地生成区域建议(Ren et al.，2015)。

#### 5.4.3 Mask R-CNN

何恺明等人(2017)提出了基于区域的掩模卷积网络(Mask R-CNN)实例目标分割。Mask R-CNN扩展了R-CNN的架构，并使用一个额外的分支用于预测目标掩模。

#### 5.4.4 Multi-Expert R-CNN

Lee等人(2017)提出了基于区域的多专家卷积神经网络(ME R-CNN)，利用了Fast R-CNN架构。ME R-CNN从选择性和详尽的搜索中生成兴趣区域(RoI)。它也使用per-RoI多专家网络而不是单一的per-RoI网络。每个专家都是来自Fast R-CNN的全连接层的相同架构。

### 5.5 深度残差网络

He等人(2015)提出的残差网络(ResNet)由152层组成。ResNet具有较低的误差，并且容易通过残差学习进行训练。更深层次的ResNet可以获得更好的性能。在深度学习领域，人们认为ResNet是一个重要的进步。

#### 5.5.1 Resnet in Resnet

Targ等人(2016)在Resnet in Resnet(RiR)中提出将ResNets和标准卷积神经网络(CNN)结合到深层双流架构中。

#### 5.5.2 ResNeXt

Xie等人(2016)提出了ResNeXt架构。ResNext利用ResNets来重复使用分割-转换-合并策略。

### 5.6 胶囊网络

Sabour等人(2017)提出了胶囊网络(CapsNet)，即一个包含两个卷积层和一个全连接层的架构。CapsNet通常包含多个卷积层，胶囊层位于末端。CapsNet被认为是深度学习的最新突破之一，因为据说这是基于卷积神经网络的局限性而提出的。它使用的是一层又一层的胶囊，而不是神经元。激活的较低级胶囊做出预测，在同意多个预测后，更高级的胶囊变得活跃。在这些胶囊层中使用了一种协议路由机制。Hinton之后提出EM路由，利用期望最大化(EM)算法对CapsNet进行了改进。

### 5.7 循环神经网络

循环神经网络(RNN)**更适合于序列输入**，**如语音、文本和生成序列**。一个重复的隐藏单元在时间展开时可以被认为是具有相同权重的非常深的前馈网络。由于梯度消失和维度爆炸问题，RNN曾经很难训练。为了解决这个问题，后来许多人提出了改进意见。

Goodfellow等人(2016)详细分析了循环和递归神经网络和架构的细节，以及相关的门控和记忆网络。

Karpathy等人(2015)使用字符级语言模型来分析和可视化预测、表征训练动态、RNN及其变体(如LSTM)的错误类型等。

J´ozefowicz等人(2016)探讨了RNN模型和语言模型的局限性。

#### 5.7.1 RNN-EM

Peng和Yao(2015)提出了利用外部记忆(RNN-EM)来改善RNN的记忆能力。他们声称在语言理解方面达到了最先进的水平，比其他RNN更好。

#### 5.7.2 GF-RNN

Chung等(2015)提出了门控反馈递归神经网络(GF-RNN)，它通过将多个递归层与全局门控单元叠加来扩展标准的RNN。

#### 5.7.3 CRF-RNN

Zheng等人(2015)提出条件随机场作为循环神经网络(CRF-RNN)，其将卷积神经网络(CNN)和条件随机场(CRF)结合起来进行概率图形建模。

#### 5.7.4 Quasi-RNN

Bradbury等人(2016)提出了用于神经序列建模和沿时间步的并行应用的准循环神经网络(QRNN)。

### 5.8 记忆网络

Weston等人(2014)提出了问答记忆网络(QA)。记忆网络由记忆、输入特征映射、泛化、输出特征映射和响应组成。

#### 5.8.1 动态记忆网络

Kumar等人(2015)提出了用于QA任务的动态记忆网络(DMN)。DMN有四个模块:输入、问题、情景记忆、输出。

### 5.9 增强神经网络

Olah和Carter(2016)很好地展示了注意力和增强循环神经网络，即神经图灵机(NTM)、注意力接口、神经编码器和自适应计算时间。增强神经网络通常是使用额外的属性，如逻辑函数以及标准的神经网络架构。

#### 5.9.1 神经图灵机

Graves等人(2014)提出了神经图灵机(NTM)架构，由神经网络控制器和记忆库组成。NTM通常将RNN与外部记忆库结合。

#### 5.9.2 神经GPU

Kaiser和Sutskever(2015)提出了神经GPU，解决了NTM的并行问题。

#### 5.9.3 神经随机存取机

Kurach等人(2015)提出了神经随机存取机，它使用外部的可变大小的随机存取存储器。

#### 5.9.4 神经编程器

Neelakantan等人(2015)提出了神经编程器，一种具有算术和逻辑功能的增强神经网络。

#### 5.9.5 神经编程器-解释器

Reed和de Freitas(2015)提出了可以学习的神经编程器-解释器(NPI)。NPI包括周期性内核、程序内存和特定于领域的编码器。

### 5.10 长短期记忆网络

Hochreiter和Schmidhuber(1997)提出了长短期记忆(Long short-Short-Term Memory, LSTM)，克服了循环神经网络(RNN)的误差回流问题。LSTM是基于循环网络和基于梯度的学习算法，LSTM引入自循环产生路径，使得梯度能够流动。

Greff等人(2017)对标准LSTM和8个LSTM变体进行了大规模分析，分别用于语音识别、手写识别和复调音乐建模。他们声称LSTM的8个变种没有显著改善，而只有标准LSTM表现良好。

Shi等人(2016b)提出了深度长短期记忆网络(DLSTM)，它是一个LSTM单元的堆栈，用于特征映射学习表示。

#### 5.10.1 批-归一化 LSTM

Cooijmans等人(2016)提出了批-归一化LSTM(BN-LSTM)，它对递归神经网络的隐藏状态使用批-归一化。

#### 5.10.2 Pixel RNN

van den Oord等人(2016b)提出像素递归神经网络(Pixel-RNN)，由12个二维LSTM层组成。

#### 5.10.3 双向 LSTM

W¨ollmer等人(2010)提出了双向LSTM(BLSTM)的循环网络与动态贝叶斯网络(DBN)一起用于上下文敏感关键字检测。

#### 5.10.4 Variational Bi-LSTM

Shabanian等人(2017)提出了变分双向LSTM（Variational Bi-LSTM），它是双向LSTM体系结构的变体。Variational Bi-LSTM使用变分自编码器(VAE)在LSTM之间创建一个信息交换通道，以学习更好的表征。

### 5.11 谷歌神经机器翻译

Wu等人(2016)提出了名为谷歌神经机器翻译(GNMT)的自动翻译系统，该系统结合了编码器网络、解码器网络和注意力网络，遵循共同的序列对序列(sequence-to-sequence)的学习框架。

### 5.12 Fader Network

Lample等人(2017)提出了Fader网络，这是一种新型的编码器-解码器架构，通过改变属性值来生成真实的输入图像变化。

### 5.13 超网络

Ha等人(2016)提出的超网络（Hyper Networks）为其他神经网络生成权值，如静态超网络卷积网络、用于循环网络的动态超网络。

Deutsch(2018)使用超网络生成神经网络。

### 5.14 Highway Networks

Srivastava等人(2015)提出了高速路网络（Highway Networks），通过使用门控单元来学习管理信息。跨多个层次的信息流称为信息高速路。

#### 5.14.1 Recurrent Highway Networks

Zilly等人(2017)提出了循环高速路网络(Recurrent Highway Networks，RHN)，它扩展了长短期记忆(LSTM)架构。RHN在周期性过渡中使用了Highway层。

### 5.15 Highway LSTM RNN

Zhang等人(2016)提出了高速路长短期记忆(high-Long short Memory, HLSTM)RNN，它在相邻层的内存单元之间扩展了具有封闭方向连接(即Highway)的深度LSTM网络。

### 5.16 长期循环 CNN

Donahue等人(2014)提出了长期循环卷积网络(LRCN)，它使用CNN进行输入，然后使用LSTM进行递归序列建模并生成预测。

### 5.17 深度神经 SVM

Zhang等人(2015)提出了深度神经SVM(DNSVM)，它以支持向量机(Support Vector Machine, SVM)作为深度神经网络(Deep Neural Network, DNN)分类的顶层。

### 5.18 卷积残差记忆网络

Moniz和Pal(2016)提出了卷积残差记忆网络，将记忆机制并入卷积神经网络(CNN)。它用一个长短期记忆机制来增强卷积残差网络。

### 5.19 分形网络

Larsson等人(2016)提出分形网络即FractalNet作为残差网络的替代方案。他们声称可以训练超深度的神经网络而不需要残差学习。分形是简单扩展规则生成的重复架构。

### 5.20 WaveNet

van den Oord等人(2016)提出了用于产生原始音频的深度神经网络WaveNet。WaveNet由一堆卷积层和softmax分布层组成，用于输出。

Rethage等人(2017)提出了一个WaveNet模型用于语音去噪。

### 5.21 指针网络

Vinyals等人(2017)提出了指针网络(Ptr-Nets)，通过使用一种称为「指针」的softmax概率分布来解决表征变量字典的问题。

## 6. 深度生成模型

在本节中，我们将简要讨论其他深度架构，它们使用与深度神经网络类似的多个抽象层和表示层，也称为深度生成模型(deep generate Models, DGM)。Bengio(2009)解释了深层架构，例如Boltzmann machine(BM)和Restricted Boltzmann Machines(RBM)等及其变体。

Goodfellow等人(2016)详细解释了深度生成模型，如受限和非受限的玻尔兹曼机及其变种、深度玻尔兹曼机、深度信念网络(DBN)、定向生成网络和生成随机网络等。

Maaløe等人（2016）提出了辅助的深层生成模型（Auxiliary Deep Generative Models），在这些模型中，他们扩展了具有辅助变量的深层生成模型。辅助变量利用随机层和跳过连接生成变分分布。

Rezende等人(2016)开发了一种深度生成模型的单次泛化。

### 6.1 玻尔兹曼机

玻尔兹曼机是学习任意概率分布的连接主义方法，使用最大似然原则进行学习。

### 6.2 受限玻尔兹曼机

受限玻尔兹曼机(Restricted Boltzmann Machines, RBM)是马尔可夫随机场的一种特殊类型，包含一层随机隐藏单元，即潜变量和一层可观测变量。

Hinton和Salakhutdinov(2011)提出了一种利用受限玻尔兹曼机(RBM)进行文档处理的深度生成模型。

### 6.3 深度信念网络

深度信念网络(Deep Belief Networks, DBN)是具有多个潜在二元或真实变量层的生成模型。

Ranzato等人(2011)利用深度信念网络(deep Belief Network, DBN)建立了深度生成模型进行图像识别。

### 6.4 深度朗伯网络

Tang等人(2012)提出了深度朗伯网络(Deep Lambertian Networks，DLN)，它是一个多层次的生成模型，其中潜在的变量是反照率、表面法线和光源。DLNis是朗伯反射率与高斯受限玻尔兹曼机和深度信念网络的结合。

### 6.5 生成对抗网络

Goodfellow等人(2014)提出了生成对抗网络(generate Adversarial Nets, GAN)，用于通过对抗过程来评估生成模型。GAN架构是由一个针对对手（即一个学习模型或数据分布的判别模型）的生成模型组成。Mao等人(2016)、Kim等人(2017)对GAN提出了更多的改进。

Salimans等人(2016)提出了几种训练GANs的方法。

#### 6.5.1 拉普拉斯生成对抗网络

Denton等人(2015)提出了一种深度生成模型(DGM)，叫做拉普拉斯生成对抗网络(LAPGAN)，使用生成对抗网络(GAN)方法。该模型还在拉普拉斯金字塔框架中使用卷积网络。

### 6.6 循环支持向量机

Shi等人(2016a)提出了循环支持向量机(RSVM)，利用循环神经网络(RNN)从输入序列中提取特征，用标准支持向量机(SVM)进行序列级目标识别。

## 7. 训练和优化技术

在本节中，我们将简要概述一些主要的技术，用于正则化和优化深度神经网络(DNN)。

### 7.1 Dropout

Srivastava等人(2014)提出Dropout，以防止神经网络过拟合。Dropout是一种神经网络模型平均正则化方法，通过增加噪声到其隐藏单元。在训练过程中，它会从神经网络中随机抽取出单元和连接。Dropout可以用于像RBM(Srivastava et al.，2014)这样的图形模型中，也可以用于任何类型的神经网络。最近提出的一个关于Dropout的改进是Fraternal Dropout，用于循环神经网络(RNN)。

### 7.2 Maxout

Goodfellow等人(2013)提出Maxout，一种新的激活函数，用于Dropout。Maxout的输出是一组输入的最大值，有利于Dropout的模型平均。

### 7.3 Zoneout

Krueger等人(2016)提出了循环神经网络(RNN)的正则化方法Zoneout。Zoneout在训练中随机使用噪音，类似于Dropout，但保留了隐藏的单元而不是丢弃。

### 7.4 深度残差学习

He等人(2015)提出了深度残差学习框架，该框架被称为低训练误差的ResNet。

### 7.5 批归一化

Ioffe和Szegedy(2015)提出了批归一化，通过减少内部协变量移位来加速深度神经网络训练的方法。Ioffe(2017)提出批重归一化，扩展了以前的方法。

### 7.6 Distillation

Hinton等人(2015)提出了将知识从高度正则化模型的集合(即神经网络)转化为压缩小模型的方法。

### 7.7 层归一化

Ba等人(2016)提出了层归一化，特别是针对RNN的深度神经网络加速训练，解决了批归一化的局限性。

## 8. 深度学习框架

有大量的开源库和框架可供深度学习使用。它们大多数是为Python编程语言构建的。如Theano、Tensorflow、PyTorch、PyBrain、Caffe、Blocks and Fuel、CuDNN、Honk、ChainerCV、PyLearn2、Chainer,、torch等。

## 9. 深度学习的应用

在本节中，我们将简要地讨论一些最近在深度学习方面的杰出应用。自深度学习(DL)开始以来，**DL方法以监督、非监督、半监督或强化学习的形式被广泛应用于各个领域**。从分类和检测任务开始，DL应用正在迅速扩展到每一个领域。

例如：

图像分类与识别

视频分类

序列生成

缺陷分类

文本、语音、图像和视频处理

文本分类

语音处理

语音识别和口语理解

文本到语音生成

查询分类

句子分类

句子建模

词汇处理

预选择

文档和句子处理

生成图像文字说明

照片风格迁移

自然图像流形

图像着色

图像问答

生成纹理和风格化图像

视觉和文本问答

视觉识别和描述

目标识别

文档处理

人物动作合成和编辑

歌曲合成

身份识别

人脸识别和验证

视频动作识别

人类动作识别

动作识别

分类和可视化动作捕捉序列

手写生成和预测

自动化和机器翻译

命名实体识别

移动视觉

对话智能体

调用遗传变异

癌症检测

X射线CT重建

癫痫发作预测

硬件加速

机器人等。

Deng和Yu(2014)提供了DL在语音处理、信息检索、目标识别、计算机视觉、多模态、多任务学习等领域应用的详细列表。

使用深度强化学习(Deep Reinforcement Learning, DRL)来掌握游戏已经成为当今的一个热门话题。每到现在，人工智能机器人都是用DNN和DRL创建的，它们在战略和其他游戏中击败了人类世界冠军和象棋大师，从几个小时的训练开始。例如围棋的AlphaGo和AlphaGo Zero。

## 10. 讨论

尽管深度学习在许多领域取得了巨大的成功，但它还有很长的路要走。还有很多地方有待改进。至于局限性，例子也是相当多的。例如：Nguyen等人表明深度神经网络（DNN）在识别图像时容易被欺骗。还有其他问题，如Yosinski等人提出的学习的特征可迁移性。Huang等人提出了一种神经网络攻击防御的体系结构，认为未来的工作需要防御这些攻击。Zhang等人则提出了一个理解深度学习模型的实验框架，他们认为理解深度学习需要重新思考和概括。

Marcus在2018年对深度学习(Deep Learning, DL)的作用、局限性和本质进行了重要的回顾。他强烈指出了DL方法的局限性，即需要更多的数据，容量有限，不能处理层次结构，无法进行开放式推理，不能充分透明，不能与先验知识集成，不能区分因果关系。他还提到，DL假设了一个稳定的世界，以近似方法实现，工程化很困难，并且存在着过度炒作的潜在风险。Marcus认为DL需要重新概念化，并在非监督学习、符号操作和混合模型中寻找可能性，从认知科学和心理学中获得见解，并迎接更大胆的挑战。

## 11. 结论

尽管深度学习（DL）比以往任何时候都更快地推进了世界的发展，但仍有许多方面值得我们去研究。我们仍然无法完全地理解深度学习，我们如何让机器变得更聪明，更接近或比人类更聪明，或者像人类一样学习。DL一直在解决许多问题，同时将技术应用到方方面面。但是人类仍然面临着许多难题，例如仍有人死于饥饿和粮食危机，癌症和其他致命的疾病等。我们希望深度学习和人工智能将更加致力于改善人类的生活质量，通过开展最困难的科学研究。最后但也是最重要的，愿我们的世界变得更加美好。