

卷积神经网络的研究进展综述

杨 斌, 钟金英

(南华大学 电气工程学院, 湖南 衡阳 421001)

摘 要:深度学习(deep learning, DL)强大的建模和表征能力很好地解决了特征表达能力不足和维数灾难等模式识别方向的关键问题,受到各国学者的广泛关注.而仿生视觉系统的卷积神经网络(convolutional neural network, CNN)是 DL 中最先成功的案例,其局部感受野、权值共享和降采样三个特点使之成为智能机器视觉领域的研究热点.对此,本文综述 CNN 最新研究成果,介绍其发展历程、最新理论模型及其在语音、图像和视频中的应用,并对 CNN 未来的发展潜力和发展方向进行了展望和总结.

关键词:深度学习;卷积神经网络;特征提取;智能识别

中图分类号:TP412.14

文献标识码:A

Review of Convolution Neural Network

YANG Bin, ZHONG Jin-ying

(School of Electrical Engineering, University of South China, Hengyang, Hunan 421001, China)

Abstract: Deep learning theory has received extensive attention of scholars all over the world because of its powerful modeling and high representational abilities. It solved the key problems of pattern recognition, such as the insufficiency of expression ability and dimensionality curse. Convolutional neural network (CNN) is a successful component of deep learning, which imitates the biological vision system. Local receptive field, sharing weights and down sampling are three important characteristics of CNN which lead it to be the hot-spot in the field of intelligent machine vision. Therefore, this paper summarizes the latest research works of CNN. Firstly, the history of CNN is introduced. Secondly, state-of-the-art modified models of CNN are reviewed. Then, the applications of CNN in speech, image and video processing are illustrated. Finally, the development trends of CNN are concluded.

收稿日期:2016-04-05

基金项目:南华大学青年英才支持计划基金资助项目(聘字 2014-004 号);国家自然科学基金资助项目(61102108);南华大学校内博士启动基金资助项目(2011XQD29);湖南省优秀博士学位论文基金资助项目(YB2013B039)

作者简介:杨 斌(1980-),男,河南平顶山人,南华大学电气工程学院副教授,博士.主要研究方向:多传感器信息融合理论、多分辨率表示理论、压缩传感理论、深度学习理论和模式识别理论.

key words: deep learning; convolutional neural network; feature extraction; intelligent recognition

1 卷积神经网络模型

20 世纪 40 年代末提出 M-P 神经元模型, 人工神经网络的研究由此拉开序幕. 1986 年, Rumelhart 等学者提出反向传播网络学习算法 (back propagation, BP)^[1], 神经网络开始在各领域得到广泛应用. 继 BP 神经网络之后, 一系列分类器如支持向量机、Boosting 和最大熵等方法相继提出. 这些算法通常只含一个隐藏层, 属于浅层学习, 与人类大脑的深度层次结构相差较大^[2]. 且全连接的神经网络训练时间长, 参数依靠经验和技巧设定, 容易过拟合和陷入局部最优解. 由于缺乏深度

不能高效地学习数据深层次结构信息, 导致神经网络技术的发展停滞不前. CNN 打破了原有全连接神经网络的技术瓶颈, 获得了全球研究者的广泛关注. CNN 受启发于猫的视觉系统结构, 即以局部感受野来平铺覆盖整个视野区域. 如图 1 所示, 若以 1000×1000 的输入图像为例, 卷积层的节点数为 10^6 , 则全连接方式 a 需要 10^{12} 个权值. 假设图 1b 采用 10×10 的局部感受野, 则其权值减少为 10^8 个. 若系统有 100 个滤波器, 在局部感受野的基础上, CNN 采用权值共享 c 可进一步将权值个数减少到 10^4 . 由此可见, CNN 的权值共享和局部感受野可以大幅度的减少网络参数.

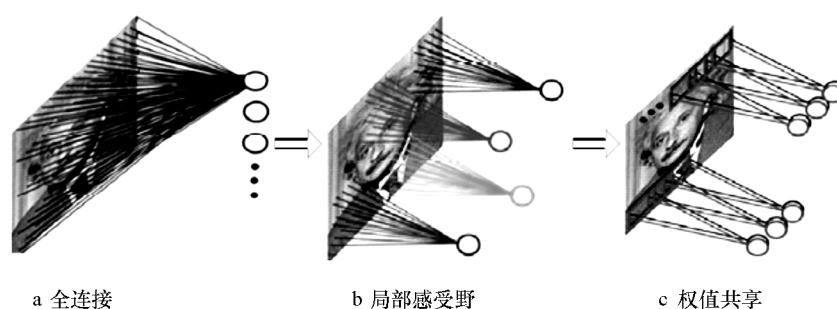


图 1 卷积神经网络减少参数的过程

Fig.1 The process of reducing the parameters of convolutional neural network

与 BP 类似, CNN 采用前向传播计算出输出值, 反向传播调整权重和偏置. 如图 2 所示, 经典的 CNN 包括卷积层 (C)、降采样层 (S) 和全连接层 (F) 三部分^[3]. 卷积层的卷积核本质就是特征提取器, 配合深度网络模型可以自动提取输入信号的深层信息. 降采样层即池化层, 实现特征图的

采样处理, 在减少数据量的同时保留有用的信息, 也使 CNN 具有抗畸变的能力. 全连接层 (如 softmax 层) 一般位于网络尾端, 对前面逐层变换和映射提取的特征进行回归分类等处理, 也可作输出层, 把训练得到的特征图汇总成特征向量, 为输入图像的另一信息表示.

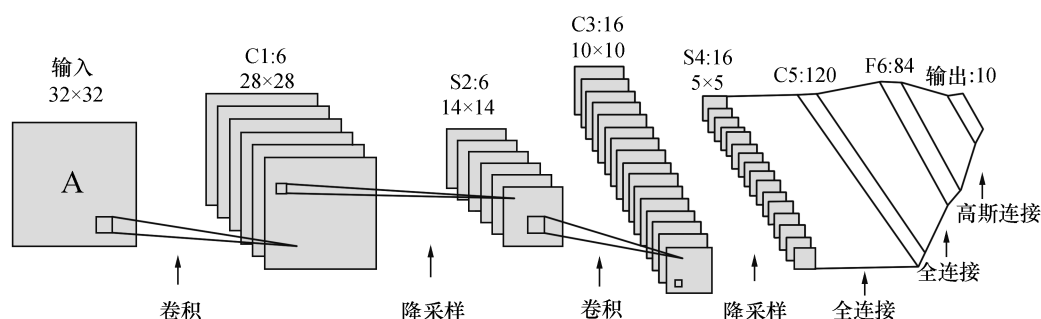


图 2 LeNet5 模型

Fig.2 The model of LeNet5

CNN 可能是机器学习领域近十年最成功的研究方向,因此本文拟对 CNN 最新研究成果进行介绍,并对其预期研究方向进行展望。

2 卷积神经网络的发展历程

人类身处大量的数据包围之中,却总能以一种灵巧的方式来获取自己想要的信息。对于人工智能领域来说,模仿人脑高效准确地从大量感知数据中获得本质信息,一直是模式识别理论研究的核心挑战。2006年,机器学习领域的泰斗 Hinton 指出多层的神经网络结构能够学习到更深刻更本质的特征,并通过“逐层初始化”的训练办法克服了传统神经网络训练瓶颈问题,引领了 DL 的发展方向^[4]。从此,深度学习受到了各国学者的广泛关注,谷歌、微软、IBM、百度等拥有大数据的高科技公司相继投入大量资源进行 DL 理论研究。2011年以来,微软研究院和谷歌的语音识别人员采用深度神经网络(deep neural network, DNN)技术,使语音识别有了突破性进展,识别误差可降低 20% ~ 30%。2013 年百度宣布成立百度研究院,而其第一个研究重点就是 DL。同年 10 月,Facebook 试图通过启用 DL 这一新的人工智能技术来帮助理解它的用户和相关数据,DL 甚至可以仅仅基于那些并未有明确暗示的帖子来理解用户的心情或者预测将要发生的事。

DL 的基本思想是建立模拟人脑的神经元连接结构,在处理具有复杂的数据内部关系时,通过多个变换分层对数据特征进行描述,进而得到数据的本质表示^[5]。常见的 DL 模型有 CNN^[6]、深度信任网络(deep belief networks, DBN)^[7]和栈式自动编码器(stacked auto-encoder, SAE)^[8]。其中, CNN 是 DL 中最成功的一个分支,是对著名的 Hubel-Wiesel 生物视觉模型的模拟,通过多网络结构、卷积运算和降采样来实现对信号深层信息提取。CNN 把特征提取归入模型学习,把特征学习和分类学习有机结合起来,更有效地实现对图像的识别。特别是近两年,卷积神经网络受到了更为广泛地关注。2014 年的大型视觉辨识挑战赛(ImageNet Large Scale Visual Recognition Challenge, ILSVRC)中,几乎所有的参赛队伍都采用了卷积神经网络及其变形方法。2015 年国际机器学习大会上来自工业界和学术界的各位专家对 DL 展开激烈讨论,指出 DL 在人类擅长的领域已经取得了令人振奋的成功,未来的探索方向是人类并不擅长的任务、数据集。2015 年国际计算机视觉与模

式识别会议上关于 DL 和 CNN 的研究成果较往年有大幅的提升。2015 年 9 月,微软亚洲研究院的“深层残差网络”(deep residual networks)在 ILSVRC 中获得图像分类、定位以及检测全部项目的冠军,如分类方面将错误率降低至 4.94%,比人眼辨识(错误率 5.1%)更精确^[9]。

在大数据时代,各领域争先占领 DL 的技术制高点,希望找出更强大的模型来揭示海量数据所承载的丰富信息,对未知事件更精准地预测。

3 卷积神经网络模型的改进

近年,围绕 CNN 的卷积层、降采样层及全连接层三部分,各国学者不断探索研究,提出了多种形式的改进。

3.1 CNN 卷积层的改进

在增加模型深度可以有效增加网络性能的理论基础上, Krizhevsky 等把 CNN 的卷积层数加至 5 层,构成了深度卷积神经网络(deep convolutional neural network, DCNN)^[10]。Szegedy 等组成的 GoogLeNet 小组构建了 22 层深度网络^[11],该网络比 2012 ILSVRC 获胜者的网络参数少了 12 倍,却能更精确的分类和检测,在 2014 年的 ILSVRC 一举夺冠。如何构建更深网络一直是 CNN 的研究热点,2015 ILSVRC 的 152 层“深层残差网络”^[9]采用了一个全新的“残差学习”原则指导学习,并重新定向了深层 CNN 中的信息流,高效地解决了深层神经网络准确度与层数之间的矛盾。

王冠皓提出了多级金字塔卷积神经网络(multi-level pyramid CNN)^[12], CNN 每层先使用稀疏自编码器预训练,再使用金字塔卷积神经网络对整个网络训练,将低层的滤波权值共享到高层,保证训练时的卷积滤波只在小尺寸图像上进行,大尺度图像只进行一次向前传播,加快训练速度。此外, Zhang 在 ImageNet 上训练具有 7 层的快速非线性卷积神经网络^[13],对每层加速以达到对整个网络加速的目的,速度可达到与“AlexNet”^[14]一致的同时,分类精度也提高了 4.7%。为了保证卷积特征不变性, Springenberg 等构造了 stride 卷积层^[15],省去最大池化层,该网络结构相对简单,但其识别效果比常见的卷积层和池化层组合模型更好。

3.2 CNN 降采样层的改进

常规 DCNN 只能对尺度固定的图像进行处理,对尺度不同的输入图像需人工裁剪或拉伸之后才输入网络,会造成图像失真而影响分类的准

确率.He 等在 DCNN 的卷积层尾端引入空间金字塔降采样层去除固定尺度的约束,降采样获得与全连接层相匹配的特征.多尺度训练得到的空间金字塔降采样网络 (spatial pyramid pooling network, SPP-net)^[16] 可处理任意尺度的图像. Rippel 等提出了频谱池化的方法^[17],将 CNN 进行傅里叶变换来把耗时的卷积操作替换为乘积操作可加速网络训练,通过截取信号在频域特征表示来降低维度,不仅可以保留更多的信息,还比其它的池化方法更灵活,提高了利用傅里叶变换计算卷积效率.

Zeiler 等提出了反卷积神经网络 (deconvolutional networks, DN)^[18],采用非监督方法对自然图像自顶向下进行卷积分解,再组合滤波器组学习到的卷积特征,获得一个特征丰富具有鲁棒性的图像表征.但大量卷积操作也使 DN 比自底向上的编码模型耗时,因而 Zeiler 在 DN 每层之间采用 3D 最大池化层,引入一组潜在开关变量,使得模型可以根据多种输入模式自适应调整,构建了由卷积稀疏编码层和最大降采样层交替的自适应反卷积神经网络 (adaptive deconvolutional networks, ADN)^[19],在中间层和高层特征学习都能取得不错的效果.

3.3 CNN 全连接层的改进

为了克服过拟合问题, Krizhevsky 在 CNN 两个全连接层采用了“Dropout”策略^[10],每次迭代以 0.5 的概率使一些随机分布的节点不工作但权值保留,之后又可能参与训练.该策略可以防止某些特征仅在其它特殊特征出现下才产生作用的情况,从而有效地避免了网络中特征之间的相互依赖性.孙艳丰提出了基于改进 Fisher 准则的深度 CNN 识别算法 (fisher-based convolution neural network, FCNN)^[20],反向传播采用 Fisher 的约束准则,加入判别准则的能量函数,在迭代求解权值的搜索空间受到判别条件约束影响,从全局搜索缩小到更有利于分类的局部空间,从而使权值更快地逼近便于分类的最优值.

Girshick 等提出基于区域卷积神经网络 (Regions with CNN features, RCNN)^[21],先找兴趣区域再进行检测和分割.由于 RCNN 兴趣区域中心点的运算量大而耗时,提出 fast R-CNN (FRCN)^[22],把全连接层的矩阵进行 SVD 分解,速度提高了 30%.对于给定的图像和目标框方案, FRCN 进行目标检测耗时 60~330 ms,忽略了在区域方案提取所需的时间 (约 1.51 s).因此, Ren 等

人进一步提出了 faster R-CNN^①,构建一个耗时极少且精确度高的全卷积层区域提取方案网络 (region proposal network, RPN),从提区域方案到目标检测总共耗时 198 ms,在 VOC2007 上 mAP 也从 FRCN 的 70% 提高到 73.2%.

无论卷积层的改进,还是降采样层及全连接层的改进,都是针对网络的学习能力及学习效率两方面的改进.因为 CNN 的学习能力可以通过调节网络的深度来增强,增加模型深度可以有效增加网络性能.为了使 CNN 更好地模拟大脑复杂的深层次神经网络,增加网络深度逐渐成为一种趋势.另外, CNN 的卷积核相当于特征提取器,可以有效地提取输入信号的有用信息,但是耗时是卷积操作必须重视的问题,针对这个问题,提出了一系列的加速方法,在一定程度上降低了训练和测试的时间.但整体上,与传统的非神经网络方法相比, CNN 的耗时还是比较大,需要进一步改善.如何在提高网络学习能力和减少耗时之间找到一个平衡点,依旧是 CNN 面临的问题.

4 卷积神经网络算法的应用

在 LeNet5 成功的用于银行手写支票识别之后,关于 CNN 模型在不同的数据集和实际生产应用的研究越来越多,网络多样化决定其应用对象各式各样.

在语音识别方面,主流自动语音识别系统通常使用隐形马尔科夫模型和混合高斯模型 GMM-HMM,而 DNN 可以建模相关性复杂的语音特征,用 DNN-HMM 模型^[23]比 GMM-HMM 对说话方式和环境具有更好的鲁棒性. Ossama 等在 DNN-HMM 模型后面加上局部滤波层和最大采样层^[24],获得更高性能的多个说话者语音识别系统.随后发现在处理语音特征时有限的权值共享更具有优势,对于不同的频率通道使用单独的权值,学习不同的特征更适合语音识别^[25].

在图像识别领域, Sun 等设计三层 CNN 用于估计脸部关键点的位置^[26],利用了脸部的纹理环境信息来锁定关键点,训练网络时考虑了关键点的几何位置分布,从而克服了传统人脸识别受遮挡、姿势变换和光照等影响大的缺点.闫鹏使用级

① Ren S, He K, Girshick R, et al. Faster R-CNN: towards real-time object detection with region proposal Networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016. (已接收还未发表)

联 CNN 实现车牌检测^[27],通过引入级联结构,同时提高检测准确率和检测效率.Zhang 等用基于贝叶斯优化的搜索策略,目标检测框依次通过候选区域,把具有明显定位不准确的目标训练 CNN,处理定位问题^[28].

CNN 在视频识别还处于起步阶段,不过仍有很多卓有成效的识别方法.Ji 等将 3D 的视频看成空域上的 2D 信号,学习时空域相关性来获得帧之间的关联特征,从而获得多个相邻帧间的运动信息,实现行为识别^[29].Karpathy 等用来自 487 类的 100 万段 YouTube 视频构建了 Sports⁻¹M 数据集,并训练了基于 CNN 的大规模视频分类的经验评估模型^[30],其分类准确率可达 63.3%,比基于人工特征 55.3%的识别率有明显提高.Liu 等基于滤波器响应提出了对运动敏感的时空域深度卷积神经网络^[31],可以很好的检测到视频的运动特征.

在低层图像处理方面 CNN 同样显示了其巨大的潜力.Dong 等提出的超分辨率卷积神经网络(super-resolution CNN, SRCNN)^[32],在低分辨率图像和高分辨图像之间学习一个端对端映射,网络结构简单却能提供优越的处理结果.此外, CNN 在图像去噪^[33]、图像反卷积^[34]和图像修复^[35]等方向也展示了广阔的应用前景.

5 卷积神经网络研究展望

深度卷积神经网络在图像处理和机器学习领域的应用,取得了突破性的进展,体现了深度 CNN 处理图像数据的优势.未来卷积神经网络将会在以下五个方面进一步发展.

1) 多输入卷积神经网络

CNN 主要用来实现传统的模式识别任务,性能较传统模式有很大改进,但依然没有突破单个图像输入+特征提取+分类器的结构.对于图像清晰度识别,视频异常判断等应用,需要多幅输入图像综合分析才能得到识别结果.越来越多的视觉处理系统都采用多种具有互补特性的图像传感器来提高系统的可靠性和鲁棒性.因此构造多图像输入的深度 CNN 具有极高的理论和应用价值.另外不同卷积层能够代表不同层次的图像特征信息,通过构造不同层次的信息融合分层,可以方便地实现多传感视觉系统像素级、特征级和决策级的信息融合.

2) 提高卷积神经网络的泛化能力

实践中,具有类标号的数据较少,为每个数据

人工设置类标号是非常费时和枯燥的.然而,为了使得监督卷积神经网络通过训练具有较高的泛化能力,需要大量具有类标号的训练样本,这也是制约监督卷积神经网络在实践中应用的主要因素.因此,研究非监督或者半监督卷积神经网络是进一步拓宽卷积神经网络应用的重要途径.另外还可以利用核函数等结构最优化方法提高网络全连接层的泛化能力.

3) 神经网络的结构先验约束

CNN 是仿生物视觉模型,其卷积核模拟视觉细胞的功能.然而现有 CNN 的结构与真实视觉结构仍然有很大的不同.现存的 CNN 模型的单个卷积核内部各元素在训练的时候视为相互独立的,如果能够将卷积核各像素之间的关系考虑进来,充分发挥卷积核的优势,则隐层可以捕获更多的特征信息,应该可以更好地表达和理解图像,在图像重建及识别上可取得更好的效果.论文^[33]也提到,卷积层可视为具有非线性约束的特殊的稀疏表示,利用 CNN 的卷积层能对图像局部凸信息进行特征提取,可以很好地解决图像融合的关键问题.

4) 高层输出对低层输出的反馈

CNN 的网络结构由多个不同作用的节点层构成,每个层次的输出特征图表示不同层次的源图像特征.现有 CNN 的应用往往更注重最终的识别结果,而没有充分用到不同层次抽象级别的图像特征.对于低层图像处理任务,综合利用不同层次的图像特征可以有效改善处理效果,而利用高层信息对低层信息进行反馈也是低层图像处理的一个潮流.CNN 的纵向多层结构为利用高层特征指导低层图像融合提供了方便的途径.因此,研究深度 CNN 逐层特征之间的关系,制定高层信息对低层信息的反馈指导是进一步提升图像处理质量的关键.

5) 卷积神经网络应用的扩展

近几年, CNN 在目标检测、图像识别及处理等领域已经显示了巨大的优势.然而, CNN 的应用领域还可以大大扩展.正如 2015 年国际机器学习大会上的各国专家所指出的, CNN 应用研究的中心未来将转移到模式识别以外的更广泛领域,甚至是探索一些人类难以完成的任务,继续研究新的理论和方法,突破存在的限制,更好的服务于各行各业.

6 结 论

目前 CNN 已成为机器学习领域中的一个研究热点,其强大的建模和特征学习能力为模式识别的瓶颈问题提供了高效的途径和思路,成功应用于多种模式分类问题.其本质是构建含有多隐层的网络模型,训练之后可以有效地捕获隐藏的数据内部结构,得到大量更具代表性的特征信息,从而对样本更高精度的分类和预测.本文选择深度学习中最接近人类视觉系统的卷积神经网络,介绍其基本原理,然后综述 CNN 最新的算法及应用,最后结合自己的研究方向展望 CNN 用于多聚焦图像融合、多光谱图像融合等应用.虽然 CNN 还存在系统复杂欠稀疏性、训练时间长及依赖标签数据等缺点,其强大的学习能力和巨大潜力表明,如果能突破这一系列的难题,将大大加速推进人工智能向前发展,更加广泛地应用在各个领域.

参考文献:

- [1] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors [J]. *Nature*, 1986, 323(6088): 533-536.
- [2] Bengio Y. Deep Learning of Representations: looking forward [J]. *Lecture Notes in Computer Science*, 2013, 7978: 1-37.
- [3] Bottou L, Bengio Y, Cun Y L. Global training of document processing systems using graph transformer networks [C]// *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Juan, Puerto Rico: IEEE, 1997: 489-494.
- [4] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks [J]. *Science*, 2006, 313(5786): 504-507.
- [5] Bengio Y. Learning deep architectures for AI [M]. Hanoi: The Association for Computing Machinery, 2009.
- [6] Lecun Y, Bottou L, Bengio Y, et al. Gradient based learning applied to document recognition [J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2324.
- [7] Hinton G E, Osindero S, Yw T. A fast learning algorithm for deep belief nets [J]. *Neural Computation*, 2006, 18(7): 1527-1554.
- [8] Deng L, Seltzer M, Yu D, et al. Binary coding of speech spectrograms using a deep auto-encoder [C]// *Proceedings of the 11th Annual Conference on International Speech Communication Association*. Chiba, Japan: Makuhari, 2010: 1692-1695.
- [9] He K M, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]// *Proceedings of the International Conference on Computer Vision and Pattern Recognition*. Las Vegas, Nevada: IEEE, 2016: 770-778.
- [10] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks [C]// *Proceeding of 26th Annual Conference on Neural Information Processing System*. Lake Tahoe, USA: MIT Press, 2012: 1097-1105.
- [11] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions [C]// *Proceedings of International Conference on Computer Vision and Pattern Recognition*. Boston, USA: IEEE Computer Society, 2015: 1-9.
- [12] 王冠皓, 徐军. 基于多级金字塔卷积神经网络 (MLPCNN) 的快速特征表示方法 [J]. *计算机应用研究*, 2015, 32(8): 2492-2495.
- [13] Zhang X, Zou J, Ming X, et al. Efficient and accurate approximations of nonlinear convolutional networks [C]// *Proceedings of the Conference on Computer Vision and Pattern Recognition*. Columbus, USA: IEEE, 2014: 1984-1992.
- [14] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]// *Proceedings of the Conference on Computer Vision and Pattern Recognition*. Columbus, USA: IEEE, 2014: 580-587.
- [15] Springenberg J T, Dosovitskiy A, Brox T, et al. Striving for simplicity: the all convolutional net [C]// *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, San Diego, USA: IEEE, 2015.
- [16] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition [J]. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2015, 37(9): 1904-1916.
- [17] Rippel O, Snoek J, Adams R P. Spectral representations for convolutional neural networks [J]. *Advances in Neural Information Processing Systems*, 2015, 28(1): 2440-2448.
- [18] Zeiler M D, Krishnan D, Taylor G W, et al. Deconvolutional networks [C]// *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. California, USA: IEEE, 2010: 2528-2535.
- [19] Zeiler M D, Krishnan D, Taylor G W, et al. Adaptive deconvolutional networks for mid and high level feature learning [C]// *Proceedings of the International Conference on Computer Vision*. Colorado Springs, USA: IEEE, 2011: 2018-2025.
- [20] 孙艳丰, 齐光磊, 胡永利. 基于改进 Fisher 准则的深度卷积神经网络识别算法 [J]. *北京工业大学学报*,

- 2015, 41(6): 835-841.
- [21] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition, Columbus, Ohio, USA; IEEE, 2014: 580-587.
- [22] Girshick R, Girshick R. Fast R-CNN[C]//Proceedings of the International Conference on Computer Vision. Copenhagen, Denmark; IEEE, 2015: 1440-1448.
- [23] Mohamed A R, Dahl G E, Hinton G. Acoustic modeling using deep belief networks[J]. IEEE Transactions on Audio Speech and Language Processing, 2012, 20(1): 14-22.
- [24] Ossama A H, Mohamedz A R, Jiang Hui, et al. Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition[C]//Processings of International Conference on Acoustics, Speech and Signal Processing. Kyoto, Japan; IEEE, 2012: 4277-4280.
- [25] Ossama A H, Mohamedz A R, Jiang Hui, et al. Convolutional neural networks for speech recognition[J]. IEEE Transactions on Audio, Speech, and Language processing, 2014, 22(10): 1533-1545.
- [26] Sun Y, Wang X, Tang X. Deep convolutional network cascade for facial point detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Portland, USA; IEEE, 2013: 3476-3483.
- [27] 闫鹏, 牛常勇, 范明. 基于级联卷积网络的自然场景下的车牌检测[J]. 计算机工程与设计, 2014, 35(12): 4296-4301.
- [28] Zhang Y, Sohn K, Villegas R, et al. Improving object detection with deep convolutional networks via Bayesian optimization and structured prediction[J]. Computer Science, 2015, 8(1): 132-132.
- [29] Ji S, Yang M, Yu K. 3D convolutional neural networks for human action recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(1): 221-231.
- [30] Karpathy A, Toderici G, Shetty S, et al. Large-scale video classification with convolutional neural Networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA; IEEE, 2014: 1725-1732.
- [31] 刘琮, 许维胜, 吴启迪. 时空域深度卷积神经网络及其在行为识别上的应用[J]. 计算机科学, 2015, 42(7): 245-249.
- [32] Dong C, Chen C L, He K, et al. Learning a deep convolutional network for image super-resolution[C]//Proceedings of the European Conference on Computer Vision. Zurich; Springer, 2014: 184-199.
- [33] Jain V, Seung H S. Natural image denoising with convolutional networks[C]//Proceedings of the 28th Annual Conference on Neural Information Processing Systems. Montreal Canada; MIT Press, 2015.
- [34] Xu L, Ren S J, Liu Ce, et al. Deep convolutional neural network for image deconvolution[C]//Proceedings of the 27th Annual Conference on Neural Information Processing Systems. Montreal, Canada; MIT, 2014: 1790-1798.
- [35] Ren S J, Xu L, Yan Q, et al. Shepard convolutional neural networks[C]//Proceedings of the 28th Annual Conference on Neural Information Processing Systems. Montreal, Canada; MIT, 2015.