

# 基于深度学习的医学影像诊断综述

张巧丽 赵 地 迟学斌

(中国科学院计算机网络信息中心 北京 100190)

**摘 要** 目前各类医学影像数据积累迅速,给利用传统影像分析方法实现疾病诊断的医生带来了巨大挑战;计算机视觉领域的深度学习方法日渐成熟,为实现医学影像的自动分析及辅助医生实现疾病的高精度智能诊断提供了新的契机。文中综述了深度学习在医学影像领域的最新研究进展。首先,介绍了深度学习以及该类方法在医学影像领域的应用情况;然后,从应用深度学习主要研究的几大病症来分析具体的研究进展;最后,总结研究动向,预测研究趋势,并提出深度学习在医学影像研究中可能存在的问题以及建议。

**关键词** 深度学习,疾病诊断,医学影像

中图分类号 TP3-05 文献标识码 A

## Review for Deep Learning Based on Medical Imaging Diagnosis

ZHANG Qiao-li ZHAO Di CHI Xue-bin

(Computer Network Information Center, Chinese Academy of Sciences, Beijing 100190, China)

**Abstract** At present, the various modalities of medical image data accumulate rapidly, bringing great challenges to doctors who diagnose disease through traditional medical image analysis methods. Deep learning method has gained great success and become more and more popular in the computer vision field. All that case provides new chances for automatic medical image analysis and makes high precisely computer-aided disease diagnosis possible. In this paper, we reviewed state-of-the-art research progress of deep learning in the medical image field. Firstly, the method of deep learning and its application in the field of medical imaging are introduced. Then attention is focused on specific research progress of deep learning method in several typical and popular disease. Finally, the tendency of this research field is summarized, and then the existing problems and recommendations are put forward.

**Keywords** Deep learning, Disease diagnosis, Medical imaging

## 1 研究背景

自2006年起,深度学习就作为机器学习领域的一个分支出现在人们的视野中,它是一种使用多层复杂结构或由多重非线性变换构成的多个处理层进行数据处理的方法<sup>[1]</sup>。近年来,深度学习在计算机视觉、语音识别、自然语言处理、音频识别与生物信息学等领域取得了突破性进展<sup>[2]</sup>。深度学习由于在数据分析上具有很可观的应用前景,已经被誉为2013年来的十大技术突破之一。深度学习方法模拟人类神经网络,通过组合多个非线性处理层对原始数据进行逐层抽象,从数据中获得不同层面的抽象特征并用于分类预测。深度学习的优点在于用非监督式或半监督式的特征学习和分层特征提取高效算法来替代手工获取特征<sup>[3]</sup>。

医疗关乎着人民的生命健康。目前医疗数据量巨大,但能否利用好这庞大的医疗数据为医疗事业做出贡献至关重要。医疗数据量虽然巨大,但也存在很多问题:医疗数据多样,包括图、文、视频、磁等;由于使用的设备不同,数据的质量存在较大差异;数据呈现波动特征,随时间和特定事件变化;由于个体存在差异,病症的规律没有普遍适用性等。这些问题的存在有着很多不可抗因素,我们只能在现有的基础上利用现有技术,尽最大可能解决这些问题。在医疗数据中,医学

影像是一个十分重要的组成部分。

本文首先介绍深度学习中最重要几种模型;然后以具体的几大病症为例来分析当前深度学习在病例影像处理中的应用现状;最后总结可能遇到的问题,并对未来的应用前景进行预测。

## 2 模型介绍

深度神经网络的很多结构模型都被应用在医学影像中,例如卷积神经网络(Convolutional Neural Networks, CNN)、稀疏编码网络(Sparse Coding, HSC)、栈式自编码器(Stacked Auto-Encoders, SAE)、受限玻尔兹曼机(Restricted Boltzmann Machines, RBM)、深度信念网络(Deep Belief Networks, DBN)等。下文就对这些模型的原理和框架进行简单介绍。

### 2.1 卷积神经网络

Hubel和Wiesel<sup>[4]</sup>于1962年基于对猫视觉皮层细胞的研究,提出了感受野(receptive field)的概念。1984年日本学者Fukushima基于感受野<sup>[5]</sup>概念提出了神经认知机(neocognitron)模型,它可以看作是卷积神经网络的第一个实现网络,也是感受野概念在人工神经网络领域的首次应用。

卷积神经网络的关键思想在于多层堆叠、局部连接、权重共享和池化(pooling)。将单层的卷积神经网络进行多次堆

张巧丽(1990—),女,硕士生,主要研究方向为类脑计算;赵 地(1978—),男,副研究员,主要研究方向为类脑计算, E-mail: zhaodi@cnic.cn (通信作者);迟学斌(1963—),男,研究员,博士生导师,主要研究方向为高性能计算、脑科学。

叠,前一层的输出作为后一层的输入,便构成卷积神经网络;CNN 中的层与层之间不再是全连接的,而是局部连接的,这样可以大大简化模型的复杂度,减少参数的数量。

如图 1 所示,输入的数据为  $1000 \times 1000$  像素的图像,有 1 百万个隐层神经元,每个隐层神经元都连接图像的每一个像素点,这样便有  $1000 \times 1000 \times 1000000 = 10^{12}$  个连接,也即  $10^{12}$  个权值参数。而图(b)中的局部连接网络中,每一个节点与上层节点同位置附近  $10 \times 10$  的窗口相连接,则  $1000 \times 1000$  个隐层神经元就只有  $10^6 \times 100$  即  $10^8$  个参数。其权值连接个数比原来降低了 4 个数量级。

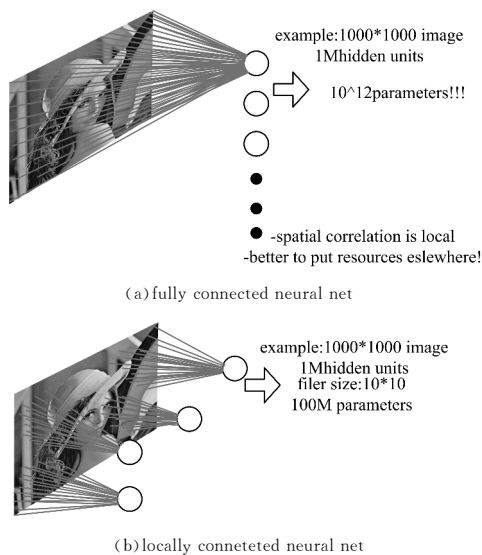


图 1 CNN 的全连接和局部连接

权值共享即设置每个神经元参数相同,用同一个卷积核去卷积图像,该方式能够避免算法过拟合,通过拓扑结构建立层与层间非全连接空间关系来减少训练参数的数目;池化即采样(subsampling)对每个特征图进行独立操作,通常采用平均池化(average pooling)或者最大池化(max pooling)的操作。平均池化依据定义的邻域窗口计算特定范围内像素的均值  $PA$ ,邻域窗口平移步长大于 1(小于等于池化窗口的大小);最大池化则将均值  $PA$  替换为最值  $PM$  输出到下一个阶段。池化操作后,输出特征图的分辨率降低,但能较好地保持高分辨率特征图描述的特征。

CNN 主要包括 3 种网络层:卷积层(Convolutional)、采样层(Subsampling)和全连接层(Fully connected)。图 2 给出了简要结构,包含两个卷积层、两个采样层和一个全连接层。

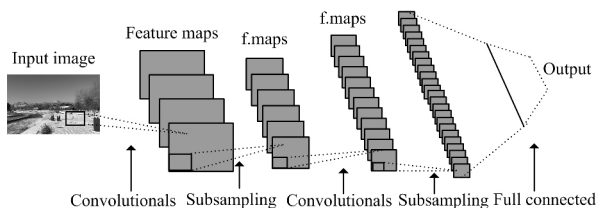


图 2 卷积神经网络架构图

## 2.2 稀疏编码网络

稀疏编码是一种模拟哺乳动物视觉系统主视皮层  $v1$  区简单细胞感受视野的人工神经网络方法。该方法具有空间的局部性、方向性和频域的带通性,是一种自适应的图像统计方法。而且,这种方法编码方式的实现仅依靠外界感知信息的统计特性,并不依赖于输入数据的性质,因而是一种自适应的

信号统计方法。稀疏编码有如下几个优点:编码方案存储能力大,具有联想记忆能力,并且计算简便;使自然信号的结构更加清晰;编码方案既符合生物进化普遍的能量最小经济策略,又满足电生理实验的结论。

稀疏编码网络的结构如图 3 所示。输入需要处理的图像数据,然后经中间的编码层的神经网络处理,相当于对原图进行任意的干扰,而解码层的神经网络处理又相当于将干扰后的图像进行了恢复。因为中间出现了干扰与解干扰的过程,输入等于了输出,多次变换后,保留下来的能识别到的原输入就是特征。然而,完整的神经网络处理过程需要大密度的运算,因此提出了稀疏编码算法来高效地表示样本数据。

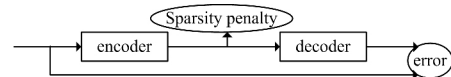


图 3 稀疏编码流程

## 2.3 栈自动编码器

Bengio 等<sup>[6]</sup>仿照 Stacked RBM 构成的 DBN(见图 4),提出 Stacked AutoEncoder,为非监督学习在深度网络的应用提供了更好的方法。

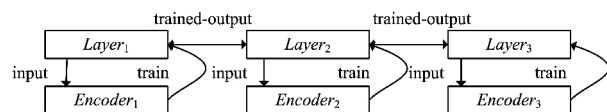


图 4 Stacked RBM 构成 DBN 图

非监督学习网络训练方式和监督学习网络的方式是相反的。在监督学习网络中,各个层的权值参数受制于输出层的误差函数,因而  $Layer_i$  参数的梯度依赖于  $Layer_{i+1}$  的梯度,形成了“一次迭代-更新全网络”的反向传播。但是在非监督学习中,各个 Encoder 的参数  $W$  只受制于当前层的输入,因而可以训练完  $Encoder_i$ ,把参数转给  $Layer_i$ ,利用优势参数传播到  $Layer_{i+1}$ ,再开始训练,形成“全部迭代-更新单层”的新训练方式。这样,  $Layer_{i+1}$  效益非常高,因为它吸收的是  $Layer_i$  完全训练奉献出的精华 Input。

## 2.4 受限玻尔兹曼机

受限玻尔兹曼机(Restricted Boltzmann Machine, RBM)是由 Hinton 和 Sejnowski<sup>[7]</sup>于 1986 年提出的一种通过输入数据集学习概率分布的随机生成式神经网络(Generative Stochastic Neural Network),该网络由一些可见单元(visible unit,对应可见变量,亦即数据样本)和一些隐藏单元(hidden unit,对应隐藏变量)构成,可见变量和隐藏变量都是二元变量,亦即其状态取  $\{0,1\}$ 。受限玻尔兹曼机是一种玻尔兹曼机的变体,但限定模型必须为二分图。模型中包含对应输入参数的输入(可见)单元和对应训练结果的隐单元,图中的每条边必须连接一个可见单元和一个隐单元。可见单元之间和隐藏单元之间都不会有边连接,如图 5 所示。

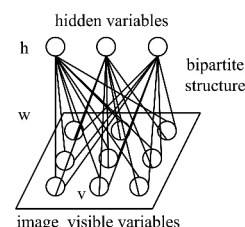


图 5 RBM 模型图

## 2.5 深度信念网络

深度信念网络(Deep Belief Network, DBN)由 Geoffrey Hinton<sup>[8]</sup>于 2006 年提出。它是一种训练产生的模型,通过训练其神经元间的权重,可以让整个神经网络按照最大概率来生成训练数据。DBN 既可以用于识别特征、分类数据,也可以用来生成数据。

DBN 由多层神经元构成,这些神经元又分为显性神经元和隐性神经元(以下简称显元和隐元)。显元用于接受输入,隐元用于提取特征,因此隐元也有一个别名,即特征检测器(feature detectors)。最顶上的两层间的连接是无向的,组成联合内存(associative memory)。较低的其他层之间是连接上下的有向连接。最底层代表了数据向量(data vectors),每一个神经元代表数据向量的一维。DBN 的组成元件是受限玻尔兹曼机。训练 DBN 的过程是逐层进行的。在每一层中,用数据向量来推断隐层,再把这一隐层当作下一层的数据向量。

经典的 DBN 网络结构是由若干层 RBM 和一层 BP 组成的一种深层神经网络,其结构如图 6 所示。

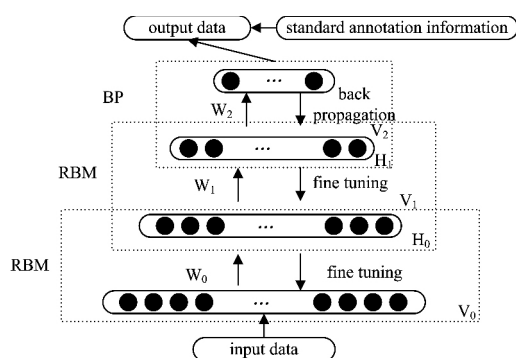


图 6 经典的 DBN 网络结构

DBN 在训练模型的过程中主要分为两步:

第一步 分别单独无监督地训练每一层 RBM 网络,确保特征向量映射到不同特征空间时都尽可能多地保留特征信息;

第二步 在 DBN 的最后一层设置 BP 网络,接收 RBM 的输出特征向量作为它的输入特征向量,有监督地训练实体关系分类器;而且每一层 RBM 网络只能确保自身层内的全值对该层特征向量映射达到最优,并不是对整个 DBN 的特征向量映射达到最优,因此反向传播网络还将错误信息自顶向下传播至每一层 RBM,微调整个 DBN 网络。RBM 网络训练模型的过程可以看作是对一个深层 BP 网络权值参数进行初始化,使 DBN 克服了 BP 网络因随机初始化权值参数而容易陷入局部最优和训练时间长的缺点。

## 3 研究现状分析

目前,在医学领域,医学病症的检查诊断多数以医学影像判断为主,然而医学影像有以下特点:高度依赖成像设备和成像环境;图像种类多差异,很难融合;图像像素大,信噪比低且图像分辨率低;生物个体存在差异性、易变性。这些影像问题很大程度地限制了医学病症的预测和诊断。从人体结构上看,人类疾病包括:脑(脑血管病,神经性疾病[阿尔茨海默病,帕金森病,癫痫],精神疾病[抑郁症,精神分裂症],脑瘤等)、胸(心脏疾病,肺结节/肺癌,乳腺结节/乳腺癌等)、颈(颈动脉

检测,甲状腺癌等)、眼(糖尿病眼病等)、腹部(胃癌)、男性盆骨(前列腺癌等)、女性盆骨(子宫颈癌等)等。深度学习由于基于的是对样例的学习,因此需要大量的医学图像作为基础。目前深度学习方法的研究集中在较大数据量的几大病症,如阿尔茨海默症、肺癌、乳腺癌等。只有在已有病症中获得较好的临床结果,才能拓展到其他病症领域。

### 3.1 脑部疾病

脑部疾病是颅内组织器官(脑膜血管、大脑、小脑、脑干、颅神经等)的炎症、血管病、肿瘤、变性、畸形、遗传病、免疫性疾病、营养代谢性疾病、中毒、外伤、寄生虫病等的总称。下面以老年人常见的阿尔茨海默症和帕金森病为例,介绍深度学习方法在这两种常见疾病中所取得的进展。

#### 3.1.1 阿尔茨海默病(Alzheimer's Disease, AD)

AD 是一种起病隐匿的进行性发展的神经系统退行性疾病,最常见的临床特征是患者在晚年陷入痴呆状态。在接下来的 20 年里,估计患病人数将增加一倍,到 2050 年大概每 85 个人中将有一人患此病。由于患病人数激增,AD 患者的护理费用急剧上升,早期的计算机辅助治疗和准确的 AD 诊断成为关键<sup>[9]</sup>。在这方面,深度学习方法为 AD 病症的研究做出了巨大贡献。

早期诊断在预防和治疗 AD 中起着至关重要的作用。早期诊断是基于提取脑图像特征进行分类的。这些特征必须准确捕获解剖大脑的主要 AD 相关变异结构,如心室大小、海马状、皮质厚度和脑体积。这些特征的获取则来源于几种流行的非侵入性神经成像工具,如结构 MRI(sMRI)、功能 MRI(fMRI)和正电子发射断层扫描(PET)。很多科学家利用这些成像工具为各类病症的诊断和研究做出了重大贡献。例如, Liu 等<sup>[10]</sup>从 MRI 和 PET 扫描中提取 83 个感兴趣区域,并使用多模态融合创建一组特征来训练堆叠的去噪自动编码器,文中提出的数据融合策略遵循训练自动编码器的去噪方式,理论上增加了特征学习的难度,但控制了过拟合。 Suk 等<sup>[11]</sup>用堆叠自编码从 MRI、PET 和脑脊液(CSF)图像中独立提取特征;充分考虑了它们的简易精神状态检查(MMSE)和 AD 评估表比较组合这些特征;并根据选择的 3 个 MRI、PET 和 CSF 特征与多核 SVM 进行 AD 分类。后来,其又采用多模式深度玻尔兹曼机从每个选定的 MRI 和 PET 扫描贴片中提取一个特征,并用 SVM 的集合预测 AD。 Payan 等<sup>[12]</sup>基于 SAE 的预训练,提出了一种用于 AD 诊断的 3D 卷积神经网络。随机选择 3D MRI 扫描片来预训练 SAE,训练的 SAE 权重用于 3D-CNN 的卷积过滤预训练。最后,对 3D-CNN 的全连接层进行微调分类,但是微调需要在训练阶段以牺牲大量的计算复杂度为代价。然而, Hosseini-Asl<sup>[13]</sup>等分析基于 sMRI 的特征提取技术限制了 AD 分类的准确性,因为从脑 sMRI 获得的体素特征非常嘈杂,这些数据需要经过平滑和聚类后才能用于分类,这样经过处理的数据本身就缺乏准确性。为了改进上述传统的特征提取方法,他们提出了一种新的深度监督自适应 3D-CNN 网络,在该网络中,3D-CAE 学习并自动提取识别 AD 特征,捕获由 AD 引起的变化。将 3D-CAE 预训练的卷积过滤器进一步应用于另一个数据域集,如 CAD Dementia 预训练后的 AD 神经影像学(简称 AD-ND)数据集。

图 7 给出 CAD Dementia 脑 sMRI 在堆叠 3D-CAE 的 3 个层:皮质厚度和体积(i),脑大小(j),心室大小(k)和海马模型(l)。通过 max-pooling 在每个层对特征图进行下采样,以减小它们的大小,并检测较高级特征。

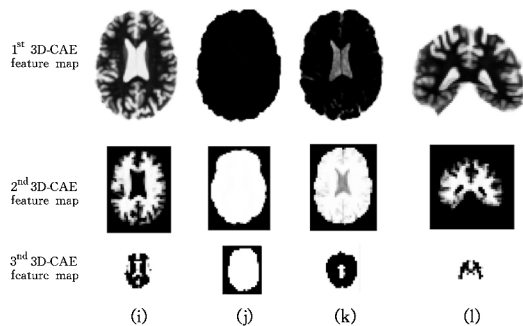


图 7 轴向(i,j,k)和从源域提取的 3D 特征图层次矢状切片<sup>[13]</sup>

除此之外, Sarraf 等<sup>[14]</sup>则分别针对 sMRI 和 fMRI 使用卷积神经网络中有名的 LeNet-5 框架对 AD 样本进行训练, 分别得到了 98.84% 和 96.85% 的准确率, 这是第一次将 fMRI 数据用于培养基于深度学习的管道。除此之外, 该项研究还表明, 开发的管道可以作为表征多模式 MRI 生物标志的有效算法。LeNet-5 的框架结构如图 8 所示, 其中包括两个卷积层、两个池化层、两个全连接层以及在该分类问题中输出二进制判决的 softmax 层。

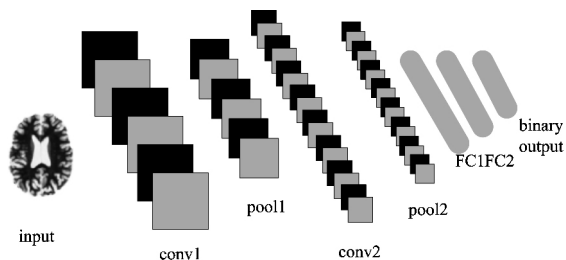


图 8 LeNet-5 的框架结构

在 sMRI 和 fMRI 分开通过深度学习方法进行 AD 分类的研究基础上, Sarraf 等<sup>[15]</sup>进一步提出了先进的以深度学习为基础的管道, 用于区分 AD 病症与相同年龄组的正常对照组中的 MRI 图像和 fMRI。

在 AD 的早期诊断中, 我们课题组也进行了相关实验, 提出了基于增强 AlexNet 的深度学习方法。实验根据 AD 的特点提出改进原始模型的 4 种算法, 将分类效果最杰出的算法命名为增强的 AlexNet 模型<sup>[16]</sup>; 分别使用原始的 AlexNet 模型和增强的 AlexNet 模型进行训练, 然后将其与同行的实验结果进行比较; 采用并行方式计算, 使用曙光 W780-G20 服务器, 另用 8 块 NVIDIA Tesla K80 进行 GPU 计算, 图像总数量超过 7 万张, 耗时不超过 30 分钟; 最后对敏感度、特异度、精准度等测试结果进行评估。最终, 实验展示出比同行出色的测试结果。

### 3.1.2 帕金森(Parkinson's Disease, PD)

帕金森病是一种常见的神经系统变性疾病, 常见于老年人, 平均发病年龄为 60 岁左右, 40 岁以下起病的青年帕金森病较少见。我国 65 岁以上人群中 PD 的患病率大约是 1.7%。大部分帕金森病患者为散发病例, 仅有不到 10% 的患者有家族史。受 PD 影响的人显示出各种运动特征, 其严重

程度随着疾病的发展而增加, 包括僵硬、运动的迟缓、抖动和步态问题等。这些运动特征的严重性和性质在一天中变化, 严重影响着 PD 患者的生活质量。

帕金森疾病的诊断主要依赖于对肢体状态的检测。将深度学习应用检测诊断过程中取得了可观的成果。AL-Fatlawi A H, Jabardi M H 等<sup>[17]</sup>提出使用深度信念网络(DBN)作为有效的诊断技术, 这种诊断是基于患者的语音信号建立的。DBN 用于分类帕金森病时包括两个堆叠的限制玻尔兹曼机(RBM)和一个输出层。需要应用两个阶段的学习来优化网络的参数: 第一阶段是无监督学习, 其使用 RBM 来克服可能由于初始权重的随机值而导致的问题; 第二阶段, 反向传播算法被用作微调的监督学习。该系统的总体测试精度为 94%。深度学习在语音和图像分类中取得了突破性进展, 但在分析可穿戴传感器数据中还没有得到研究。鉴于此, Shamir R R, Dolber T 等<sup>[18]</sup>提出将深度学习作为可穿戴传感器用于评估运动障碍的检测方法, 他们使用惯性测量单位从 10 个特发性帕金森病患者中收集数据, 设定了几个专家标记的运动任务用于分类, 检测运动的迟缓程度。比较标准机器学习管道和基于卷积神经网络的深度学习, 结果深度学习在分类率方面优于其他最先进的机器学习算法至少 4.6%。

### 3.2 胸部疾病

胸部疾病主要分布在肺、乳腺、食管、气管、支气管等部位。其中, 肺癌和乳腺癌是当前世界上发病率比较高的两大疾病, 医学界也为研究这些疾病投资了巨大的人力和物力。在治疗疾病的同时, 积累了大量临床数据和影像资料。下面以这两种病症为例来介绍深度学习方法在这两类疾病中的研究进展。

#### 3.2.1 肺癌(Lung Cancer)

肺癌是发病率和死亡率增长最快、对人群健康和生命威胁最大的恶性肿瘤之一。近 50 年来, 许多国家都报道肺癌的发病率和死亡率均明显增高, 男性肺癌发病率和死亡率均占所有恶性肿瘤的第一位, 女性发病率和死亡率占第二位。在美国和中国, 肺癌都是致死率最高的癌症之一。美国癌症协会于 2013 年发布的报告显示, 美国肺癌 5 年整体生存率为 17%; 随着早期筛查在美国的开展, 早期肺癌 5 年生存率已达到 60%~80%。由此可见, 早期预防和诊断在挽救病人生命中起着至关重要的作用。在肺癌早期诊断方面, 应用深度学习方法取得了很大进展。

传统的肺癌筛查方法是依靠传统的 X 线胸片、痰脱落细胞及纤维支气管镜等, 但这些方法并没有降低被筛选人群肺癌的死亡率。近年来, 肺癌的早期诊断已取得了许多新的进展, 如螺旋 CT(Spiral Computed Tomography, SCT)和正电子发射断层成像(Positron Emission Tomography, PET)、荧光支气管镜检(Fluorescence Bronchoscopy)、基因检测、与生物技术相结合的痰液检测、呼吸气体检测、胸部 X 线计算机辅助图像分析(Computer-assisted Image Analysis of Chest Radiographs)等。深度学习方法的应用则是结合其中的图像进行肺癌的检测和诊断。Sun 等<sup>[19]</sup>采用深度学习算法对肺癌图像数据库联盟(Lung Image Database Consortium, LIDC)数据库的肺癌诊断案件进行可行性测试, 选择了 3 种深度学习算法, 包括卷积神经网络(CNN)、深度信念网络(DBN)、堆叠

去噪自编码(SDAE)。CNN,DBNs 和 SDAE 的准确率分别是 0.7976,0.8119 和 0.7929,而设计的传统 CADx 的准确率是 0.7940,比 CNN 和 DBNs 低一些。Pan H<sup>[20]</sup>等提出了一种基于深度卷积神经网络的自动肺癌细胞检测方法。在该方法中,需要弱注释的图像作为训练集,然后通过训练卷积神经网络进行特征提取。该实验中采用的是数百名的肺癌组织病理学图像,在准确率和效率方面取得了很好的性能。

### 3.2.2 乳腺癌(Breast Cancer)

乳腺癌是发生在乳腺腺上皮组织的恶性肿瘤。全球乳腺癌发病率自 20 世纪 70 年代末开始一直呈上升趋势。中国乳腺癌发病率增长很快,为每年 3%~4%,5 年生存率为 73%。美国乳腺癌在 I 期阶段即诊断出的概率超过 80%,而中国在 I 期阶段诊断出乳腺癌的概率不到 20%,一旦查出,多数已经转移或扩散。对于乳腺癌的治疗,早期诊断特别重要。在乳腺癌的诊断上,很多学者应用深度学习的方法做了相关研究并取得了不错的成果。

近年来,图像处理的发展和深度学习技术的不断成熟,使得计算机辅助检测/诊断(Computer-Aided Detection/Diagnosis,CAD/CADx)可以帮助病理医师进行更加客观、有效的判断。Spanhol 等<sup>[21]</sup>提出了一种基于图像块的提取方法,其将预训练 CNN 和图像块的组合用于最终的分类。这些图像则来源于 BreaKHis(一个公开的数据库)中的乳腺癌组织病理学图。AlexNet 被用于对对象的彩色图像进行分类,在该样例中对其进行调节,使其适应乳腺癌组织病理学图像的分类。在乳腺癌中,还会出现癌细胞转移的情况。针对这种情况,Wang 等<sup>[22]</sup>提出使用深度学习的方法识别转移性乳腺癌,他们使用数以百万计的训练补丁来训练深层卷积神经网络,以进行补丁水平预测,从而区分肿瘤补丁与正常补丁。将该系统的预测与人类病理学家的诊断相结合,将病理学家的 AUC 增加到 0.995,人为的准确率降低了大约 85%。Ciresan 等<sup>[23]</sup>还使用了最大池化深度卷积神经网络来检测乳房组织学图像中的有丝分裂。利用该训练网络对图像中的每个像素进行分类,然后将简单的后期处理应用于网络输出。该方法在当年 International conference pattern recognition 竞赛中取得了很高的准确率。

### 3.3 腹部疾病

腹部包括肝、胰、脾、肾等多个器官,腹部疾病也主要集中在这些器官的损伤上。常见的腹部疾病包括:肝癌、肝炎、前列腺癌、前列腺炎、急性阑尾炎、子宫癌等。下面以前列腺癌和子宫癌为例,介绍深度学习的方法在这两类疾病的诊断中所取得的成果。

#### 3.3.1 前列腺癌(Prostate Cancer)

前列腺癌是发生于男性前列腺组织中的恶性肿瘤,是前列腺腺泡细胞异常无序生长的结果。前列腺癌的发病率具有明显的地理和种族差异。Tirumala S S 等<sup>[24]</sup>提出一种基于人工神经网络的扩展方法,用表达基因的数据对前列腺癌进行分类和预测。前列腺癌的数据集包含 12600 个属性,但只有 102 个样本。他们使用了 4 种方法从这 12600 种属性中识别出最有影响力的属性,然后使用人工神经网络和朴素贝叶斯分别分类完整的属性集。人工神经网络取得了 98.2% 的分类精度,优于朴素贝叶斯方法。在前列腺癌的检测中,Azizi S 等<sup>[25]</sup>还提出了一种用于分析前列腺组织的超声信号的自动

特征选择框架。该框架包括 3 个步骤:1)减少非监督特征;2)监督微调,使用组织样品的组织病理学来进一步优化深度信念网络;3)使用支持向量机进行分类,使用深度信念网络的激活量作为诊断前列腺癌的依据。

#### 3.3.2 子宫颈癌(Cervical Cancer)

子宫颈癌是女性常见恶性肿瘤之一,发病率位于女性肿瘤的第二位。在宫颈癌研究中,Xu 和 Zhang 等<sup>[26]</sup>提出了利用多模态信息设计一个用于宫颈发育不良诊断的深度学习框架。首先,使用卷积神经网络将低级图像数据转化为与其他非图像模态融合的特征向量;然后,共同学习了神经网络中所有模态之间的非线性相关性。该方法在大数据集上取得了 87.83% 的灵敏度和 90% 的特异性。Devi M A<sup>[27]</sup>等还使用人工神经网络对宫颈癌细胞进行检测,并对正常和异常细胞进行分类;同时还对人工神经网络中不同类型架构及其精度结果和性能进行了分析讨论。除此之外,还讨论了基于神经网络检测子宫颈癌的不同类型方法及其结构,并提出了一种具有前馈神经网络组合遗传算法的基因前馈神经网络用于检测子宫颈癌。

### 3.4 眼部疾病

常见的眼科疾病有:中心浆液性视网膜病变、干眼症、交感性眼炎、夜盲症、失明眼部结构、弱视、散光、沙眼、白内障、糖尿病视网膜病变、结膜炎、老花眼、色盲、虹膜异色症、视网膜色素变性、视网膜中央动脉阻塞、视网膜脱落、近视、远视、针眼、雪盲症、霰粒肿、青光眼、飞蚊症等。下面以糖尿病视网膜病变为例来介绍应用当前深度学习的方法取得的具体成果。

在全球,糖尿病视网膜病变(Diabetic Retinopathy)是一种主要的致盲疾病。糖尿病性视网膜病发生在糖尿病对视网膜造成损伤的时候。这是一种系统疾病,它影响了 80% 的患有 20 年以上糖尿病的患者。尽管统计的数据很惊人,但研究表明,如果对眼睛进行适当警惕的治疗和监测,则可以减少 90% 的新病例<sup>[28]</sup>。目前,糖尿病视网膜病变患者数量增长很大,主要是由于早期不能诊断正常眼睛视力。对于年轻患者而言,尽早发现糖尿病视网膜病变并阻止视力丧失非常必要。全球的学者们应用深度学习的方法在这方面做了深入的研究。

糖尿病性视网膜病变阶段图如图 9 所示。

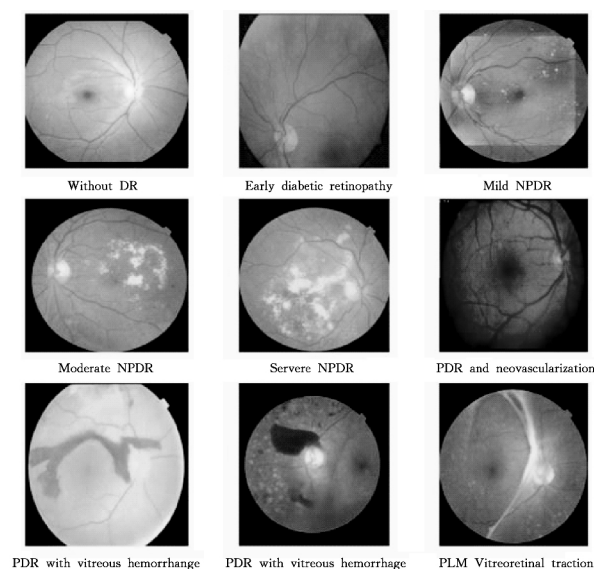


图 9 糖尿病性视网膜病变阶段图<sup>[28]</sup>

专家将糖尿病性视网膜病变分为 5 个阶段:正常,轻度,中度,重度非增殖性(Non-Proliferative Diabetic Retinopathy, NPDR)或增生性糖尿病性视网膜病变(Proliferative Diabetic Retinopathy, PDR)。

Mrinal Haloi 等<sup>[29]</sup>提出了一种基于深度学习的计算机辅助诊断新颖的微动肿瘤(MicroAneurysm, MA)方式来检测早期糖尿病视网膜病变。他们以 MA 作为糖尿病性视网膜病变的指标,因此需要准确检测 MA 并进行治疗。该深度神经网络不需要预处理或手动提取特征,为工作人员带来了很大的便利。Wang 等<sup>[30]</sup>提出了一种基于特征和集合学习的监督方法来解决视网膜血管分割的问题,该方法结合了两个分类器:CNN 和随机森林(Random Forest, RF)。在该方法中, CNN 作为训练层的特征提取器,避免了人工提取特征的复杂和误差;RF 作为训练分类器,可以很好地降低误差,提升了模型的泛化能力。Chandrakumar 等<sup>[28]</sup>提出使用深度学习架构进行糖尿病性视网膜病变分类,该学习架构包括卷积层、池化层、修正线性单元、防过拟合层(Dropout layer)、全连接层和分类层,如图 10 所示。

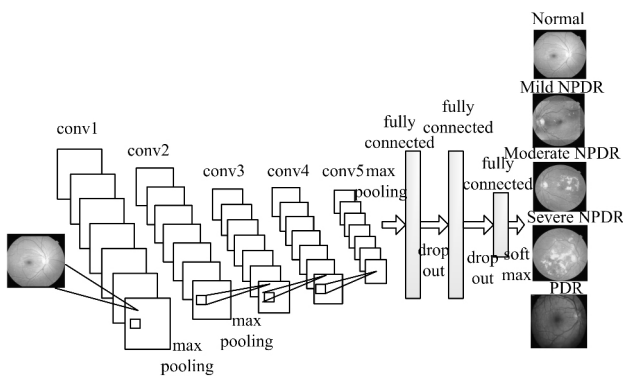


图 10 深度卷积神经网络架构<sup>[28]</sup>

该深度卷积神经网络,对于糖尿病性视网膜病变的各个阶段均适用,不用手动提取特征。通过应用过拟合技术,使得该网络架构分类更加精准,召回率也得到了提升。但是由于网络架构的复杂和计算密集度,需要高级图像处理单元来处理多层图像。Gulshan<sup>[31]</sup>于 2016 年提出了一种基于深度学习的算法,该算法能够在视网膜造影中对糖尿病视网膜病变的迹象做出解释,帮助医生克服资源短缺困难,为更多的病人做出更专业的诊断。该算法使用的是专为图像分类而优化过的深度神经网络模型,该网络使用 128175 张视网膜图像的可追溯的开发数据集进行了训练。对算法使用 2016 年 1 月和 2 月的两个互相独立的数据集进行了验证,其中的每张图像都至少经过了 7 位美国认证的眼科医生的高 intragrader 一致性的评估。同年,Wong<sup>[32]</sup>也提出了用深度学习算法开发糖尿病视网膜病变筛查软件的研究,并取得了可喜的成果,但是同时他也指出,将软件用于糖尿病视网膜病变筛查计划中的临床转化和效用确实存在挑战。

### 3.5 其他病症

除了以上提及的一系列疾病之外,还有很多疾病的预测和诊断都用到了深度学习中的方法或者应用这些方法中的某几个方法训练出新的学习框架,在此不一一详述。

## 4 总结

总体来说,深度学习在医学影像中的研究主要包含以下几个方向:1)提供临床诊断辅助系统服务,应用于早期的病症筛查、诊断、康复等场景;2)进行医学影像识别,帮助医生更快、更准确地读取影响;3)健康管理,通过一些可穿戴设备监控用户的健康指数等。

通过对深度学习在以上病症的医学图像中的诊断分析可以看出,目前在使用深度学习的模型或者方法上存在相似性,大多采用 CNN 或者其他常用的深度学习算法,或者几种算法相融合的方式来进行图像分类检测;而且现在大多数的模型都在理论阶段,还没有应用到临床中。采用无监督学习方法确实为图像分类提供了很多便利,例如可以不用人工提取特征,避免由于医生的个人因素带来的误诊情况等,但是目前该方法的研究还主要集中在发病率比较高即病人数据量比较多的病症研究中,在一些稀有病症中的研究则很少。这可能是由于深度学习本身的自学习、自适应性决定的,它需要大量的样本数据学习才能更加准确地获取病症图像特征,以进一步作出病症诊断。但是,如果深度学习能用比较少量的数据就可以准确地获取图像特征,即强化学习过程,相信无论对于疑难杂症还是大众疾病来说都将是一个福音。

当然,目前的深度学习方法也正在发展期,每一项技术的成熟和突破都是人们不断尝试、创新和完善的过程。当前,无论是计算机硬件技术还是医学摄影技术,都为深度学习处理医学影像问题提供了很好的基础条件。总之,深度学习的自学习优势,使得在提取特征的同时,提升了分类效果。相信随着该方法的不断成熟和完善,一定可以在医疗影像分析中为病症诊断提供更好的帮助。

## 5 趋向预测

综上所述,深度学习在医疗影像领域将会有更广阔的应用,现在研究进展成果仅仅是冰山一角。当然,除了机遇之外,还会面临来自各方面的挑战。首先,应用深度学习的诊断方式能够代替医生吗?结合当前的发展趋势,完全取代是不太可能的。虽然深度学习随着基于影像诊断的发展可以更好地辅助医生诊断诸多疾病,可以在诊断的环节帮助医生快速判断,为医生提供建议,但是深度学习并没有实现全智能,它没有人类的思维,因此在应用上很难代替人类医生,毕竟深度学习的研究依赖的是大量的已有数据样本,而对于疾病而言,它的变异性往往是不可控的。其次,尽管深度学习方式的辅助诊断已经取得了不错的成果,解决了很多医学上的难题,但在落地时依然会面临很多问题。尽管困难重重,但是结合当前医疗改革的趋势,可以预想深度学习在医疗方面的应用也将会有更大的需求;并且随着深度学习方法的更加精准完善,相信基于深度学习的医疗诊断技术完全可以克服这些困难。

## 参 考 文 献

- [1] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.

- [2] DENG L, YU D. Deep Learning[J]. Signal Processing, 2014, 7: 3-4.
- [3] SONG H A, LEE S Y. Hierarchical Representation Using NMF [C] // International Conference on Neural Information Processing. Springer Berlin Heidelberg, 2013: 466-473.
- [4] HUBEL D H, WIESEL T N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex[J]. The Journal of physiology, 1962, 160(1): 106-154.
- [5] FUKUSHIMA K. A hierarchical neural network model for associative memory[J]. Biological cybernetics, 1984, 50(2): 105-113.
- [6] BENGIO Y, LAMBLIN P, POPOVICI D, et al. Greedy layer-wise training of deep networks[J]. Advances in neural information processing systems, 2007, 19: 153.
- [7] HINTON G E, SEJNOWSKI T J. Learning and relearning in Boltzmann machines[J]. Parallel distributed processing: Explorations in the Microstructure of Cognition, 1986, 1: 282-317.
- [8] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554.
- [9] ALZHEIMER'S A. Alzheimer's disease facts and figures[J]. Alzheimer's & Dementia, 2014, 10(2): e47-e92.
- [10] LIU S, CAI W, et al. Multimodal Neuroimaging Feature Learning for Multiclass Diagnosis of Alzheimer's Disease[J]. IEEE Transactions on Biomedical Engineering, 2015, 62(4): 1132-1140.
- [11] SUK H I, SHEN D. Deep learning-based feature representation for AD/MCI classification[C] // MICCAI. 2013: 583-590.
- [12] PAYAN A, MONTANA G. Predicting Alzheimer's disease: a neuroimaging study with 3D convolutional neural networks [OL]. <http://arxiv.org/pdf/1502.02506.pdf>.
- [13] HOSSEINI-ASL E, GIMEL'FARB G, EL-BAZ A. Alzheimer's Disease Diagnostics by a Deeply Supervised Adaptable 3D Convolutional Network[OL]. <https://arxiv.org/pdf/1607.00556.pdf>.
- [14] SARRAF S, TOFIGHI G. DeepAD: Alzheimer's Disease Classification via Deep Convolutional Neural Networks using MRI and fMRI [OL]. <https://www.biorxiv.org/content/biorxiv/early/2016/08/21/070441.full.pdf>.
- [15] SARRAF S, TOFIGHI G. Deep Learning-based Pipeline to Recognize Alzheimer's Disease using fMRI Data[OL]. <https://arxiv.org/pdf/1610.07231.pdf>.
- [16] 吕鸿蒙, 赵地, 迟学斌. 基于增强 AlexNet 的深度学习阿尔茨海默病的早期诊断[J]. 计算机科学, 2017, 44(s1): 50-60.
- [17] AL-FATLAWI A H, JABARDI M H, LING S H. Efficient diagnosis system for Parkinson's disease using deep belief network [C] // IEEE Evolutionary Computation. 2016: 1324-1330.
- [18] SHAMIR R R, DOLBER T, NOECKER A M, et al. Machine learning approach to optimizing combined stimulation and medication therapies for Parkinson's disease[J]. Brain Stimulation, 2015, 8(6): 1025-1032.
- [19] SUN W, ZHENG B, QIAN W. Computer aided lung cancer diagnosis with deep learning algorithms[C] // SPIE Medical Imaging. 2016.
- [20] PAN H, XU Z, HANG J. An effective approach for robust lung cancer detection[C] // International Workshop on Patch-based Techniques in Medical Imaging. 2015: 87-94.
- [21] SPANHOL F A, OLIVERIRA L S, PETITJEAN C, et al. Breast Cancer Histopathological Image Classification using Convolutional Neural Networks[C] // International Joint Conference on Neural Networks. 2016: 2560-2567.
- [22] WANG D, KHOSLA A, GARGEYA R, et al. Deep learning for identifying metastatic breast cancer[OL]. <https://arxiv.org/pdf/1606.05718.pdf>.
- [23] CIREŞAN D C, GIUSTI A, GAMBARDELLA L M, et al. Mitosis detection in breast cancer histology images with deep neural networks[C] // International Conference on Medical Image Computing and Computer-assisted Intervention. 2013: 411-418.
- [24] TIRUMALA S, NARAYANAN A. Attribute Selection and Classification of Prostate Cancer Gene Expression Data Using Artificial Neural Networks[C] // Pacific-Asia Conference on Knowledge Discovery and Data Mining. 2016: 26-34.
- [25] AZIZI S, IMANI F, ZHUANG B, et al. Ultrasound-Based Detection of Prostate Cancer Using Automatic Feature Selection with Deep Belief Networks[C] // International Conference on Medical Image Computing and Computer-Assisted Intervention. 2015: 70-77.
- [26] XU T, ZHANG H, HUANG X, et al. Multimodal Deep Learning for Cervical Dysplasia Diagnosis[C] // International Conference on Medical Image Computing and Computer-Assisted Intervention. 2016: 115-123.
- [27] DEVI M A, RAVI S, VAISHNAVI J, et al. Classification of Cervical Cancer Using Artificial Neural Networks[J]. Procedia Computer Science, 2016, 89: 465-472.
- [28] CHANDRAKUMAR T, KATHIRVEL R. Classifying Diabetic Retinopathy using Deep Learning Architecture[J]. International Journal of Engineering Research and Technology, 2016, 5(6): 19-24.
- [29] HALOI M. Improved Microaneurysm Detection using Deep Neural Networks[OL]. <https://arxiv.org/pdf/1505.04424.pdf>.
- [30] WANG S, YIN Y, CAO G, et al. Hierarchical retinal blood vessel segmentation based on feature and ensemble learning[J]. Neurocomputing, 2015, 149: 708-717.
- [31] GULSHAN V, PENG L, CORAM M, et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs[J]. JAMA, 2016, 316(22): 2402-2410.
- [32] WONG T Y, BRESSLER N M. Artificial Intelligence With Deep Learning Technology Looks Into Diabetic Retinopathy Screening[J]. JAMA, 2016, 316(22): 2366-2367.