

卷积神经网络的原理及其在医学影像诊断中的应用

王霄, 朱恩照, 艾自胜

同济大学医学院医学统计学教研室, 上海 200092

【摘要】利用卷积神经网络快速高效地对医学影像数据进行分析和处理可以实现医学影像数据快速分类、定位等操作,提高医学诊疗的效率。本研究从卷积神经网络的背景和原理入手,介绍各种类型的卷积神经网络的应用场景和一些常用的卷积神经网络模型,包括残差卷积神经网络、U-net、循环卷积神经网络等及其在医学影像诊断中的应用,最后针对卷积神经网络和人工智能技术讨论了其未来的展望和挑战。

【关键词】人工智能;卷积神经网络;医学影像;医学诊断;综述

【中图分类号】R318;TP183

【文献标志码】A

【文章编号】1005-202X(2022)12-1485-05

Principle of convolutional neural network and its applications in medical imaging diagnosis

WANG Xiao, ZHU Enzhao, AI Zisheng

Department of Medical Statistics, School of Medicine, Tongji University, Shanghai 200092, China

Abstract: Convolutional neural network can realize the localization, classification and other tasks for the large amounts of medical imaging data by analyzing and processing the medical imaging data swiftly and efficiently, thereby improving the efficiency of medical diagnosis and treatment. Starting from the introduction on the background and principle of convolutional neural network, the application scenarios of various types of convolutional neural networks are presented, and the applications of some commonly-used neural network models, including residual convolutional neural network, U-net, recurrent convolutional neural network in medical imaging diagnosis are summarized. Finally, the future prospects and challenges of convolutional neural network and artificial intelligence are discussed.

Keywords: artificial intelligence; convolutional neural network; medical imaging; medical diagnosis; review

前言

根据 Intel 公司的估计,自人类起源开始到 2003 年,整个人类社会大概产生了 5 EB(Exabyte)的数据资料,而在 2012 年,总数据量已经增长到了 2.72 ZB(Zettabyte),大约是 2003 年的 500 倍,并且预计此后每两年数据量就会翻倍^[1]。在大数据时代背景下,单纯依靠人工对数据进行分析和处理会越发困难。随着人工智能(Artificial Intelligence, AI)技术的发展和成熟,AI 逐渐成为处理大量数据的高效手段。从 1950 年 Alan Turing 开始探讨能够思考的机器以来?,人们开始不断地追问和探索一个问题:是否存在可以独立思考的机器^[2]?AI 技术在经历了几次重大的进

化之后,在 2011 年,IBM 的 Watson 计算机在两次 Jeopardy 智力竞赛中战胜了人类选手^[3];与此同时,包括 Apple 公司在内的多家智能手机公司也陆续推出智能语音助手,甚至已经开发出针对个性化自动调整的 AI 来辅助完成智能化教学的系统^[4]。在智能驾驶领域,随着图像识别、分割和定位等 AI 技术的发展,辅助驾驶系统正快速地向自动驾驶系统进化^[5];在医学诊断领域,已经研发出可以用来预测心律不齐的可穿戴 AI 设备^[6]。AI 的发展正给人类的生活和生产带来前所未有的进步和无限的可能性。

在图像处理领域,如何对海量图像数据进行分析和处理是一个巨大的挑战,而图像的分类、定位和识别是最基本的图像处理问题之一。从简单的数字识别到复杂的动作识别,图像识别存在于生活中的方方面面。如何通过机器进行高效的图像识别,人们开发出很多模型和方法,包括计算机辅助图像分割技术和特征提取图像识别技术等^[7-8]。早在十余年前,就有研究者提出使用卷积神经网络的方法来进行计算网络的搭建,从而进行图像识别^[9]。卷积神经

【收稿日期】2022-06-13

【基金项目】国家自然科学基金(81872718)

【作者简介】王霄,博士研究生,研究方向:流行病与卫生统计, E-mail: naturalconan@163.com

【通信作者】艾自胜,博士,博士生导师,研究方向:流行病与卫生统计, E-mail: azs1966@126.com

网络通过模仿生物学中神经元的交互作用,进行了类似神经元作用机制的搭建。也有研究提出深层卷积神经网络的作用与大脑皮层的激活行为非常相似^[10]。ImageNet公司从2015年第一次将卷积神经网络技术应用到图像识别领域后,该技术在图像识别领域的应用越来越广泛^[11]。

1 基础原理

1.1 神经网络的基础原理

在图像分类研究开始之前,曾采用感知机^[12]等办法对图像进行分类识别。对输入信息进行加权和计算后得到一个输出信息,通过设定相应的阈值,对输出信息进行判断得到相应的分类结果。每一个感知机都可以作为一个神经元,也就是网络中的一个节点。神经网络可以整合多个神经元,构成输入层,最后负责输出的神经元则为输出层,中间的各个层级的神经元称为隐藏层,可有效对图像进行分类。

根据上述原理,假设需要识别一个 64×64 像素的黑白图像,在实际操作中可将有颜色的像素识别为输入信息为1,而没有颜色的像素识别为输入信息为0,则此图像的输入信息为 $64 \times 64 = 4096$ 个神经元,在每个神经元中,都会产生相应的结果并输入到隐藏层中。经隐藏层处理后,令最后的输出层为一个神经元,并限制这个神经元的结果为0~1,得到的图像识别输出结果为0.5以上,则判断这个结果为真值,反之则为假值。

在已经建立了一个神经网络的基础上,如何对这个神经网络进行训练得到合适的结果就是下一个需要考虑的问题。对神经网络进行训练,需首先定义一个损失函数来判断神经网络选择的相关参数结果的好坏,寻找将损失函数的值更接近于0的参数设置。一般来说,在建立神经网络的时候会采用梯度下降法对其进行预估。梯度下降法的原理是通过损失函数在某一个输入数据情况下的梯度值乘以一个预先设定的步长的负值,对损失函数进行移动后计算移动后的损失函数的值,并一直重复这个过程,求得最小的损失函数值^[13]。采用不同的步长设置,函数可达到不同的局部位置的最低点,即损失函数的局部最小值^[14]。在求解某一函数的梯度时,一般可以通过求解其相应的偏微分得到最后的梯度值;而在非常复杂的神经网络中,对每一个神经元进行偏微分的求导将产生巨大的计算量,耗费大量计算时间。为解决这个问题,一般采用微积分的链式法则和向后传播原理,将损失函数进行向后的传递,并通过每一步得到的结果利用链式法则对其进行处理,最后得到其梯度值^[15]。

1.2 卷积神经网络的基础原理

使用卷积的神经网络通常包括卷积层、池化层、全连接层和输出层,其架构如图1所示。卷积层的作用主要是对一个范围内的特征进行特征值的提取,使用不同的过滤器进行卷积操作之后能对同一个图像进行多次特征值提取而不会导致数据量的过度增加,能更好地对图像的特征值进行分解和利用,在此基础上结合非线性激活函数能更好地去除数据的冗余。池化层的作用是利用池化函数将邻近的特征进行总体的输出,得到近似不变的结果同时进一步提高网络的效率。全连接层可以将前面各层得到的抽象化的特征进行整合和归一化,得到高度提纯的特征分类概率。输出层一般是根据实际需要进行的分类任务而设置的输出,可以直接根据输出的结果进行分类任务的结果确认。在进行卷积神经网络的学习过程中,由于数据量巨大,如果每次都完整地进行数据的计算和运行会导致计算速度缓慢^[16]。因此,一般采用多批次的方式进行梯度运算来提升计算速度。使用卷积神经网络对医学图像进行高效的处理,可以完成对图像进行分类和定位等任务。本研究在此基础上,针对于卷积神经网络在医学影像中的应用场景进行介绍。

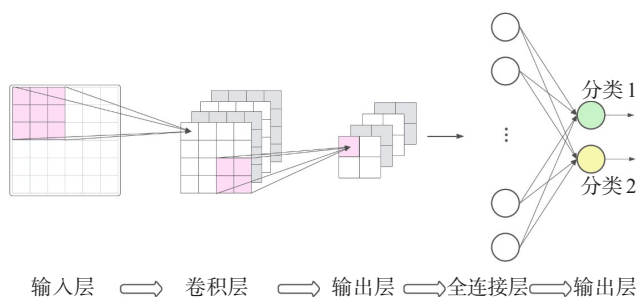


图1 卷积神经网络示意图

Figure 1 Diagram of convolution neural network

2 卷积神经网络的应用场景

医学影像诊断是现代医学中一个极其重要的诊断手段,随着医学影像学的发展,影像学资料越来越丰富、复杂,利用AI来加速对影像学数据的分析也成为了现行的趋势。既有基于机器学习中的预设特征值的算法进行的传统AI方法,也有利用类似于卷积神经网络这种深层学习方法来进行的图像识别技术。使用AI的方法已经使机器对影像资料的识别、判断水平与人类医生的表现达到基本接近的程度^[17]。

常用的医学影像数据包括从计算机断层扫描(CT)、正电子发射型计算机断层扫描(PET)和核磁

共振成像(MRI)等各种手段获得的影像数据。这些影像数据都可以通过数据转换进行算法代入,以此进行神经网络训练并将其应用于其他同类工作中;而对于不同的医学影像应用场景,可以采用不同的更为合适的卷积神经网络进行训练并进行应用。

2.1 卷积神经网络针对医学影像中分类问题的应用场景

使用卷积神经网络对影像资料进行处理的优点主要在于能使用已有的图像数据对神经网络进行分类学习的训练,然后使用训练后的神经网络进行图像分类。如前所述,卷积的操作是可以通过不同的过滤器对原有的图像数据进行特征值的提取,在卷积处理之后,也经常使用ReLU(Rectified Linear Unit)等非线性激活函数进行处理,并采用池化的方法进行特征值的压缩,通过降低数据量的大小来加速训练的速度。常用的神经网络模型有 AlexNet 和 VGG。在角膜溃疡的影像学检查中,有研究者基于 VGG 进行不同类别的角膜溃疡的分类测试,结果表明 VGG 在角膜溃疡分类任务中的准确率接近 90%^[18]。

在 AlexNet 和 VGG 等神经网络模型的广泛应用中,神经网络的层数堆叠到一定程度之后就会出现退化现象,无法降低预测的错误率^[19]。这种现象限制了深层网络的应用,随着网络层数增多,研究者无法判断能否进一步提升准确率。在这样的背景下,He 等^[19]提出残差神经网络的架构,如图 2 所示,其基本原理是在原有卷积神经网络的基础上,加上一个残差学习的通路;在原有网络的分支上,加入一个原有输入的镜像,使得进行过卷积处理的输出结果保留原有输入的结果,通过比较,使输出的结果不会比输入的结果更差。在数据维度发生变化时,也可以采用不同的方法,例如采用 1×1 卷积的方法进行增维的操作来保证维度的正常变化。使用残差网络之后,持续增加神经网络层数,错误率在一定的程度下是会持续下降的^[19]。残差网络在医学影像资料的分类上有着广泛的应用。Yoo 等^[20]采用残差神经网络的框架对前列腺癌的发现进行了神经网络模型的搭建,利用 175 名前列腺癌病人的 MRI 图像和 252 名没有患前列腺癌的受试 MRI 图像作为训练数据,对不同的残差网络进行测试,最终分别得到了在切片图像级别下判断切片图像是否有前列腺癌病灶时得到了 0.87 的受试者工作特征曲线下面积(AUC),在病人级别下判断病人是否患有前列腺癌时得到了 0.84 的 AUC,显示出了比较好的结果。在胸部 X 光图像数据的疾病诊断分类中,有研究使用不同的残差网络进行分类,在 14 个分类中,在 5 个疾病分类中能够得到较高的表现水平^[21]。

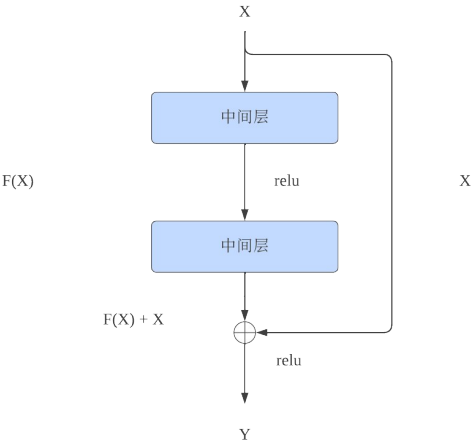


图 2 残差神经网络示意图
Figure 2 Diagram of residual neural network

2.2 卷积神经网络针对医学影像中定位问题的应用场景

针对医学影像的定位问题,最近几年也提出了使用 U-net 模型进行任务的处理,其架构如图 3 所示^[11]。定位问题就是对每一个像素点执行分类任务,U-net 模型是基于基本的卷积神经网络加入几个不同的操作,即上采样卷积、复制和剪切等,从而保证能更好地在小样本的图像数据的基础上完成更精确的分割目的。其主要的过程是一边进行编码,针对图像进行卷积操作,提取特征值;同时在另一边进行解码,进行上采样并合并两边等大小的数据,使得前后两边的特征值得到合并,用于增强最后对于每一个像素点分类的准确性。使用 U-net 可以更好地实现对图像进行分割的操作。Zeleznik 等^[22]使用 U-net 模型针对 CT 图像和心血管风险预估进行了实验。在实验中,考虑到冠状动脉中的钙含量是心血管事件的一个重要预测指标,而钙含量是可以通过胸部 CT 观察得到的,但是此过程费时费力,采用人工手段进行观察并不常规。因此,Zeleznik 等^[22]利用 U-net 对受试者数据进行分析,搭建用于分割图像中冠状动脉中的钙的神经网络,并用分割后的图像中的比例来计算得出钙含量,结果发现自动计算得出的钙含量和人工测量的结果之间几乎没有差异。针对使用 U-net 时会出现的合并低级和高级卷积特征导致模糊的特征图的限制,Zunair 等^[23]提出 Sharp U-net,用于克服 U-net 的这些局限性,结果也表明在针对医学影像的分割任务上,Sharp U-net 的性能表现优于 U-net 等其他一些神经网络。

2.3 卷积神经网络针对其他医学问题的应用场景

随着卷积神经网络应用的发展,也陆续出现了包括循环卷积神经网络(Recurrent Convolutional Neural Network, RCNN)和三维卷积神经网络(3D

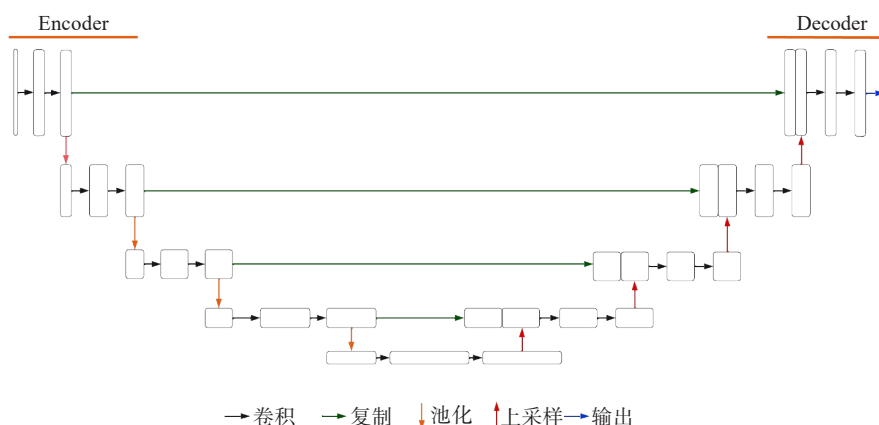


图3 U-net示意图

Figure 3 Diagram of U-net

Convolutional Neural Network, 3D-CNN)等卷积神经网络模型。其中RCNN的使用对于自然语言的解读,如医学病历信息有着重要意义^[24],而3D-CNN则可以用于医学影像和视频资料动态识别等工作中^[25]。在目标识别上,Salman等^[26]采用YOLO目标识别卷积神经网络进行前列腺癌分级诊断系统的构建,主要利用YOLO网络强大的目标识别功能进行了其中重要区域识别。对于新冠肺炎的诊断,在针对CT和X光片的AI分类的解决方案上也有非常多的实践,但仅仅限于分类结果,不足以满足临床需求。因此,Kumar^[27]提出使用基于YOLOv4的残差YOLO(RYOLO)网络进行肺炎感染部位的识别,并得到了相比YOLOv4网络更好的结果。而在与YOLO网络的比较上,Zhu等^[28]利用Faster-RCNN进行龋齿的识别,相比YOLOv5网络能更好地识别龋齿的位置和数量。

除此之外,由于医学影像来源的有限性,在建立卷积神经网络时也会存在数据量不足的情况。在卷积神经网络数据处理过程中,可采用两类常规方法进行数据的增加,分别为数据增强和转化学习^[29]。数据增强的方法一般包括对图像资料进行旋转、剪切、对比度增强等,通过对图像的预处理来增加图像数据的数量,以更好地完成训练;而转化学习的方法是可以使用之前已经训练完成的神经网络进行进一步的训练,节省所需的训练数据。在生成数据方面,生成对抗网络也被用于生成额外的图像数据来帮助训练,有时甚至可以大幅度提高神经网络的准确性^[30]。

3 展望和挑战

利用卷积神经网络的技术,结合医学场景的需求,可以加速和优化医学影像学的诊疗过程,不但可

以节省更多的人力,提高效率,甚至还能提高准确率,相信随着技术的进步和发展,利用AI的手段能够进一步促进医学影像学的效能。而与此同时,不能忽视使用卷积神经网络技术过程中面临的一系列挑战。首先,虽然卷积神经网络是近来非常有效的方法,但是其网络的搭建和参数的调节都还是无法完全被解释,需要靠不断的尝试来搭建合适的模型。其次,即使是现在已经有了非常多的数据增强的方法,对于一些罕见病的影像学评价,还是会因为数据量的问题导致无法很好地进行学习和训练。最后,在AI使用场景的伦理上也面临着各种各样的问题,包括对于AI在实用场景下的法规认定,出现判断问题下的法律认定等讨论也在持续进行中^[31]。

【参考文献】

- [1] Intel. A vision for big data[J/OL]. 2013. <https://www.intel.com/content/dam/www/public/us/en/documents/reports/intel-corp-big-data-policy-position-paper.pdf>.
- [2] Castelfranchi C. Alan Turing's "Computing machinery and intelligence"[J]. Topoi, 2013, 32(2): 293-299.
- [3] Rogers J, Jeopardy F. As artificial intelligence rapidly approaches what scientists call "technical singularity," the implications for humanity are beginning to come into focus[J/OL]. 2017. <https://businessfacilities.com/2017/12/artificial-intelligence-final-jeopardy>.
- [4] Conati C, Barral O, Putnam V, et al. Toward personalized XAI: a case study in intelligent tutoring systems[J]. Artif Intell, 2021, 298. <https://doi.org/10.1016/j.artint.2021.103503>.
- [5] Gupta A, Anpalagan A, Guan L, et al. Deep learning for object detection and scene perception in self-driving cars: survey, challenges, and open issues[J]. Array, 2021, 10: 100057.
- [6] Kodera S, Akazawa H, Morita H, et al. Prospects for cardiovascular medicine using artificial intelligence[J]. J Cardiol, 2022, 79(3): 319-325.
- [7] Malhotra S. How-to guide: deep learning for image recognition applications[J/OL]. 2020. <https://artificialintelligence.oodles.io/blogs/deep-learning-for-image-recognition>.
- [8] Lisowski E. Using artificial intelligence (AI) for image recognition[J/OL]. 2021. <https://addepto.com/using-artificial-intelligence-ai-for-image-recognition>.
- [9] Jaswal D, Sowmya, Soman KP. Image classification using convolutional neural networks[J]. Int J Adv Res Technol, 2014, 3 (6). DOI:10.14299/IJSER.2014.06.002.

- [10] Kuzovkin I, Vicente R, Petton M, et al. Activations of deep convolutional neural networks are aligned with gamma band activity of human visual cortex[J]. *Commun Biol*, 2018, 1: 107.
- [11] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015. 2015: 234-241.
- [12] Nielsen M. Using neural nets to recognize handwritten digits[J/OL]. 2015. <http://neuralnetworksanddeeplearning.com/chap1.html>.
- [13] 史加荣, 王丹, 尚凡华, 等. 随机梯度下降算法研究进展[J]. *自动化学报*, 2021, 47(9): 2103-2119.
Shi JR, Wang D, Shang FH, et al. Research advances on stochastic gradient descent algorithms[J]. *Acta Automatica Sinica*, 2021, 47(9): 2103-2119.
- [14] 郭田德, 韩丛英. 从数值最优化方法到学习最优化方法[J]. *运筹学学报*, 2019, 23(4): 1-8.
Guo TD, Han CY. From numerical optimization method to learning optimization method[J]. *Operations Research Transactions*, 2019, 23(4): 1-8.
- [15] Rumelhart DE, Hinton GE, Williams RJ. Learning representations by back-propagating errors[J]. *Nature*, 1986, 323: 533-536.
- [16] Brownlee J. A gentle introduction to mini-batch gradient descent and how to configure batch size [J/OL]. 2017. <https://machinelearningmastery.com/gentle-introduction-mini-batch-gradient-descent-configure-batch-size/>.
- [17] Hosny A, Parmar C, Quackenbush J, et al. Artificial intelligence in radiology[J]. *Nat Rev Cancer*, 2018, 18(8): 500-510.
- [18] Tang N, Liu H, Yue K, et al. Automatic classification for corneal ulcer using a modified VGG network[C]//2020 International Conference on Artificial Intelligence and Computer Engineering (ICAICE). 2020: 120-123.
- [19] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016: 770-778.
- [20] Yoo S, Gujrathi I, Haider MA, et al. Prostate cancer detection using deep convolutional neural networks[J]. *Sci Rep*, 2019, 9(1): 19518.
- [21] Baltruschat IM, Nickisch H, Grass M, et al. Comparison of deep learning approaches for multi-label chest X-ray classification[J]. *Sci Rep*, 2019, 9(1): 6381.
- [22] Zeleznik R, Foldyna B, Eslami P, et al. Deep convolutional neural networks to predict cardiovascular risk from computed tomography [J]. *Nat Commun*, 2021, 12(1): 715.
- [23] Zunair H, Ben Hamza A. Sharp U-Net: depthwise convolutional network for biomedical image segmentation[J]. *Comput Biol Med*, 2021, 136: 104699.
- [24] Lai S, Xu L, Liu K, et al. Recurrent convolutional neural networks for text classification[C]//Twenty-Ninth AAAI Conference on Artificial Intelligence. 2015: 2268-2273.
- [25] Ji S, Xu W, Yang M, et al. 3D convolutional neural networks for human action recognition[J]. *IEEE Trans Pattern Ana*, 2013, 35(1): 221-231.
- [26] Salman ME, Çakırsöy Çakar G, Azimjonov J, et al. Automated prostate cancer grading and diagnosis system using deep learning-based Yolo object detection algorithm[J]. *Expert Syst Appl*, 2022, 201: 117148.1-117148.10.
- [27] Kumar A. RYOLO v4-tiny: a deep learning based detector for detection of COVID and non-COVID Pneumonia in CT scans and X-ray images[J]. *Optik*, 2022: 268.
- [28] Zhu Y, Xu T, Peng L, et al. Faster-RCNN based intelligent detection and localization of dental caries[J]. *Displays*, 2022: 74.
- [29] Fujita H. AI-based computer-aided diagnosis (AI-CAD): the latest review to read first[J]. *Radiol Phys Technol*, 2020, 13(1): 6-19.
- [30] Onishi Y, Teramoto A, Tsujimoto M, et al. Automated pulmonary nodule classification in computed tomography images using a deep convolutional neural network trained by generative adversarial networks[J]. *Biomed Res Int*, 2019, 2019: 6051939.
- [31] Irina Evgrafova LE, Marchenko A, Tarasov A. Ethical problems of practical interaction between strong artificial intelligence and humans in the urban transport environment and legal proceedings [J]. *Transportation Research Procedia*, 2022, 63: 2094-2098.

(编辑:谭斯允)