

基于 Inception-ResNet-V2 的乳腺癌病理图像识别研究

刘靖雯, 黄理灿

(浙江理工大学 信息学院, 浙江 杭州 310018)

摘要: 乳腺癌严重威胁女性健康和生命, 及时诊断并提供治疗方案给医生带来了挑战, 病理图像分类结果是医生确诊的重要依据, 实现乳腺癌病理图像识别分类具有重要意义及临床应用价值。近年来, 大多数研究集中于良恶性分类, 而不同类型的乳腺肿瘤本身具有不同病因及治疗方法。采用 Inception-ResNet-V2 深度卷积神经网络模型, 实现对乳腺癌病理图像的八分类, 利用数据增强和迁移学习方法, 在 Matlab 上对数据集 BreaKHis 进行实验。结果表明, 该方法识别率基本达到 80% 以上, 比大部分已有研究成果效果更优。

关键词: Inception-ResNet-V2; 深度卷积神经网络; 数据增强; 迁移学习; 乳腺癌病理图像

DOI: 10. 11907/rjdk. 192019

开放科学(资源服务)标识码(OSID):



中图分类号: TP317.4

文献标识码: A

文章编号: 1672-7800(2020)005-0225-05

Pathological Image Recognition of Breast Cancer Based on Inception-ResNet-V2

LIU Jing-wen, HUANG Li-can

(School of Informatics Science and Technology, Zhejiang Sci-tech University, Hangzhou 310018, China)

Abstract: Breast cancer is a serious threat to women's health and lives, and it is a challenge for doctors to make timely diagnosis and treatment options. At the same time, because the classification results of pathological images become an important basis for the diagnosis of doctors, it is of great significance and clinical application value to realize the classification and classification of breast cancer pathological images. In recent years, most of the research has focused on benign and malignant classification, but different types of breast tumors should have different etiology and treatment methods. This paper uses the Inception-ResNet-V2 deep convolutional neural network model to achieve eight classifications of breast cancer pathological images. Using data augmentation and transfer learning methods, experiments were performed on the dataset Breakhis on Matlab, and the results show that the recognition rate of this method basically reaches over 80%, and the effect is better than most previous research results.

Key Words: Inception-ResNet-V2; deep convolutional neural network; data enhancement; transfer learning; breast cancer pathology image

0 引言

乳腺癌细胞丧失了正常细胞特性, 细胞之间连接松散, 容易脱落, 是目前威胁女性健康的常见肿瘤。相关统计数据表明, 乳腺癌在癌症死亡原因中位居第六, 且在中国女性中发病率最高^[1]。众所周知, 诊断乳腺癌的最可靠途径是通过病理诊断。临床上, 相对于 X-ray 及核磁共振图像, 医生最后确诊时依赖于对病理图像的分析。已有研

究多采用病理图像分类, 既费力又费时, 且采用人工提取方法有误诊的可能。计算机辅助诊断(Computer-Aided Diagnosis, CAD)可将病理图像自动分类, 既节省人力、提高效率, 诊断结果也更准确可靠^[2]。但实现较难: 一是病理图像存在微小差异, 容易出现细胞重叠现象、颜色分布出现不均匀等情况; 二是鉴于医学图像通常涉及病人隐私, 缺乏公开的、已标记的数据集, 使得因数据量少而造成分类困难。尽管如此, 仍有学者对乳腺癌病理图像进行大量研究, 并取得了重要研究成果。可大致分为以下两个部分:

收稿日期: 2019-07-30

作者简介: 刘靖雯(1995-), 女, 浙江理工大学信息学院硕士研究生, 研究方向为医学图像处理; 黄理灿(1962-), 男, 博士, 浙江理工大学信息学院教授、硕士生导师, 研究方向为深度学习、医学图像处理。本文通讯作者: 黄理灿。

(1)采用传统机器学习算法即人工方式提取特征。2013 年, Kowal 等^[3]利用 4 种不同的聚类算法在快速核分割的任务中进行了测试和比较。将 K 均值、模糊 C 均值、竞争学习神经网络和高斯混合模型相结合, 用于颜色空间中的聚类以及灰度中的自适应阈值处理。500 份病例被分类为良性或恶性, 识别率为 96%~100%。Zhang 等^[4]提出了一种基于单类核主成分分析(KPCA)模型集成的分类方案, 对 361 张病理图像进行分类, 识别率为 92%; Spanhol 等^[5]于 2015 年提出来源于 82 名患者的 7 909 例乳腺癌组织病理学图像数据集 BreakeHis, 并且可从提供的网址公开获得, 采用局部二值模式(LBP)等 6 种特征描述子, 利用支持向量机、随机森林等不同分类算法, 使得准确率达到 80%~85%。上述研究方法由于数据集不同导致实验结果不具有可比性, 在人工设计及特征提取时也存在过程复杂的问题。

(2)用深度学习进行乳腺癌病理图像分类。人工智能的发展, 使得越来越多人了解卷积神经网络(Convolutional Neural Networks, CNN)。近几年, 它在语音识别、信号处理、机器翻译、计算机视觉等领域获得了广泛应用^[6], 使得 CNN 应用于医学图像成为可能^[7]。此外, 深度学习实现数据自主学习特征, 解决了人工设计提取特征造成的时间长及复杂性高等问题。

2016 年, Spanhol 等^[8]使用 AlexNet 网络, 提出一种基于图像块的提取方法训练 CNN, 并使用不同的融合策略分类, 其识别率比传统算法高 6%。Bayramoglu 等^[9]提出两个不同的架构: 单一任务用于预测恶性肿瘤, 多任务 CNN 用于预测恶性肿瘤并且同时放大图像倍数, 其识别率约 83%。2017 年, Wei 等^[10]提出一种 BICNN 的深度卷积神经网络模型, 增加了网络深度和宽度, 提出完全保留癌变区域的图像边缘特征的数据增强方法, 使得准确率达到 97%。2019 年, 王恒等^[11]以 ResNet50 为基础框架, 利用迁移学习设计了一个乳腺癌图像二分类模型, 正确率可达 97.4%, 并具有泛化好、深度强、精度高、收敛快等优点。

由上可知, 采用 BreakeHis 数据集作为研究乳腺癌病理图像的主要数据集, 一方面识别率虽有较大提升, 但针对 BreakeHis 数据集 8 种疾病的多样性, 简单的良性-恶性分类已经不能满足医生实际诊断需要。

针对上述问题, 本文采用深层次混合深度卷积神经网络, 它可以自动进行八分类, 不仅能判断良性恶性, 更能确定是哪一种良性或恶性肿瘤, 从而更好地为医生提供参考。同时, 采用数据增强和迁移学习防止过拟合, 提高图像识别率, 以适应高标准的实际临床需求。

1 研究方法

与目前常用的深度卷积神经网络模型 GoogLeNet^[12]、ResNet50^[13]等对比, 本文使用 Inception-ResNet-V2^[14]网络作为基础架构, 如图 1 所示。模型主要包括: 数据读取、数据

增强及标准化处理、加载预训练网络、替换最终层、训练网络、对结果进行预测统计等内容。

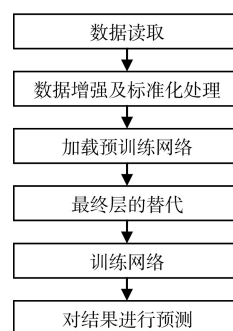


图 1 模型架构

首先在源数据上进行预训练得到模型的权重参数, 其次用得到的模型权重参数初始化新模型的残差网络层, 然后在目标数据集上进行微调训练得到最终模型, 最后在测试集上检验模型性能。

1.1 Inception-ResNet-V2 网络

2016 年, Google 团队发布 Inception-ResNet-V2 卷积神经网络, 在 ILSVRC 图像分类基准测试中拿下了最好成绩。它从微软的残差网络(ResNet)中得到启发, 由 Inception V3 模型变化而来。残差连接(Residual Connections)可以让 Shortcuts 存在于模型中, 从而简化 Inception 块, 可以让更深的神经网络训练成为可能。网络结构如图 2 所示。

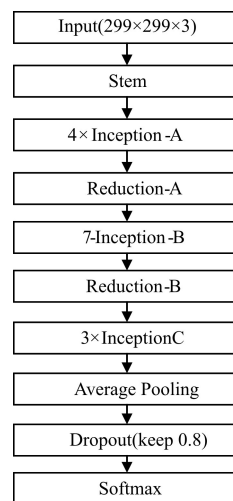


图 2 Inception-ResNet V2 网络结构

Inception-ResNet V2 设计思想如下:

由 Inception V4 和 ResNet 相结合; 3 种 Inception-ResNet 块加入直连, 使得通道多样; 比 Inception V4 参数少、收敛快; 比 Inception-ResNet-V1 块内卷积核的通道更多样化。核心结构为 Stem 块。

3 种 Inception-ResNet 块: 卷积核通道数与 Inception V4 各模块内的数目相比减少了, 但整体宽度在模块升维后相同(相比之下参数更少); 它们都加入了直连结构使得其更深、收敛更快; 最后都用 1×1 拓宽维度。

1.1.1 Inception

这里存在一个重要思想,即让网络自己决定两个过程:一是决定使用过滤器,二是决定是否池化,而不再像从前那样人为确定参量。可人为在网络中添加所有参数的可能值,将它们的输出连接起来,让网络自主学习,得知它需要哪些过滤器的组合,具体需要什么样的参数。为了调整 Inception 网络防止过拟合发生,网络中设有 Softmax 分支,这样做的目的在于正常预测图片分类,即使在进行特征计算时出现其它层,比如隐藏其中的单元或中间层参与进来时,也不会耽误结果产生。

1.1.2 ResNet 的优势

ResNet 残差网络可认为是残差块的堆叠,这样可以将网络设计得很深。残差网络和普通网络的不同点在于,在进行非线性变化前,部分数据复制一份累加后进行了非线性变换。理想状态下,随着网络深度的增加,训练误差理应逐渐缩小,而 ResNet 能够做到随着网络深度的增加,训练产生的误差缩小。对于 CNN 而言,常用的优化算法为梯度下降等。网络深度逐渐增加,会呈现训练误差先减后增的情况。

1.1.3 1×1 卷积

当 1×1 卷积^[15]出现时,其主要作用是提高或降低通道数,也即厚度,却不改变图片的宽和高。改变的通道个数取决于使用的过滤器个数,将原来的通道数改为过滤器个数,以减少运算成本。比如,一张 227×227 且厚度为 50 的图片在 20 个过滤器上进行 1×1 的卷积,则结果自然变成 $227 \times 227 \times 20$ 。另一个作用是非线性也被加入其中。先经过 CNN 再经过激励层,在前一层上进行 1×1 的卷积,以提升网络表达能力,则称之为添加了非线性激励(Non-linear Activation)。提升非线性特性要注意保证一个前提:不损失分辨率,也即维持特征图尺度不变,最后结果是让网络更深。

1.2 数据增强

需要大量带有标记的样本,这是防止过拟合情况产生的有效途径,获得医学图像通常需要高昂的代价且很难实现,想要得到经过医生标注的医学图像更是难上加难。本文在数据集 BreKHis 基础上,采用数据增强^[16]加以实现。使用增强的图像数据存储可自动调整训练图像的大小^[17]。指定对训练图像额外执行的增强操作:沿垂直轴随机翻转训练图像,以及在水平和垂直方向上随机平移训练图像最多 30 个像素。

1.3 Softmax 分类器

Softmax 大部分在多分类过程中使用,多个神经细胞的输出映射到 $(0,1)$ 区间内,形成一个概率。在全连接层输入的特征矩阵接收到后,其输出的是输入所对应的每个类别上的概率,通过这一过程实现多分类^[18]。假设输入图像数目为 N ,即为 $\{x_i, y_i\}_{i=1}^m$, $y_i \in \{1, 2, \dots, m\}$ 标记为每个图像,模型中最后输出种类数记为 $m(m \geq 2)$ 。本文进行乳腺癌

八分类,因此 m 的值为 8。 x_i 为给定的输入图像,估计假定函数 $f_\theta(x_i)$ 对应类别的概率值 $P(y_i = j | x_i)$ 。则函数表示如式(1)所示。

$$f_\theta(x_i) = \begin{bmatrix} P(y_i = 1 | x_i; \theta) \\ P(y_i = 2 | x_i; \theta) \\ \vdots \\ P(y_i = m | x_i; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^m e^{\theta_j^T x_i}} \begin{bmatrix} e^{\theta_1^T x_i} \\ e^{\theta_2^T x_i} \\ \vdots \\ e^{\theta_m^T x_i} \end{bmatrix} \quad (1)$$

对概率分布进行归一化表示为 $\frac{1}{\sum_{j=1}^m e^{\theta_j^T x_i}}$, 即全部概率

为 1。参数用 θ 表示,则损失函数可表示为式(2)。

$$J(x, y, \theta) = -\frac{1}{N} \left[\sum_{j=1}^N \sum_{i=1}^m 1\{y_i = j\} \ln \frac{e^{\theta_j^T x_i}}{\sum_{j=1}^m e^{\theta_j^T x_i}} \right] \quad (2)$$

其中, $1\{y_i = j\}$ 为指示性函数,其取值与括号内的真值相同,即 $y_i = j$ 成立时函数值为 1, 否则为 0。进而通过随机梯度下降法最小化误差函数。

1.4 迁移学习

迁移学习^[19]思想类似于人们常说的举一反三,本文使用已有机器学习方法解决相关实际问题。这里所说的相关是指不同领域但富有关联性,目的是打通知识迁移性,解决实际问题中标签样本少甚至由于现实条件下的各种情况导致标签样本不存在的情况。其基本思路是使用预训练,将现有大型数据集上训练好的权重值作为初始化值,将它迁移到实际问题的数据集上进行再次训练并且微调参数。相当于在预训练中进行学习而得到一些基本特征,可能是颜色,也可以是边框特征,总之这种方法可以提高准确率并节省训练时间^[20]。本文使用 ImageNet 数据集将预训练好的 Inception-ResNetV2 网络模型参数迁移到 BreKHis 上进行微调训练。

2 实验

2.1 数据集

乳腺癌活检是医生确诊的重要步骤,医疗上常采取肿物切活检获得最终诊断。针对这项技术,常见方法分为 4 类:细针穿刺、核心穿刺活检、真空辅助和外科(开放)活组织检查(SOB)。通过 SOB 方法收集数据集中存在的样品,也称为部分乳房切除术或切除活组织检查。本文所用的数据集 BreKHis^[5]由 SPANHOL 与巴西巴拉那 P&D 实验室(病理解剖学和细胞病理学)合作构建。从 82 名患者(良性患者 24 人,恶性患者 58 人)收集的乳腺肿瘤组织显微图像,共 7 909 张,包含 4 种不同的放大因子(40X、100X、200X 和 400X)。内含 2 480 个良性乳腺肿瘤图像和 5 429 个恶性乳腺癌症图像(3 通道 RGB,每个通道 8 位深度,700×460 像素,PNG 格式),详情如表 1 所示。

表 1 BreakHis 数据集图像分布

放大倍数	良性	恶性	总计
40X	625	1 370	1 995
100X	644	1 437	2 081
200X	623	1 390	2 013
400X	588	1 232	1 820

根据肿瘤细胞在显微镜下的外观,可以将良性和恶性乳腺肿瘤分为不同类型。各种类型的乳腺肿瘤可具有不同的预后和治疗意义。该数据集目前包含 4 种组织学上不同良性乳腺肿瘤,一般有较好机会治疗。它们分别为:纤维腺瘤(F)、腺病(A)、管状腺瘤(TA)、叶状肿瘤(PT)。4 种恶性肿瘤(乳腺癌),均以癌症被人们所知,它们分别为:癌(DC)、粘液癌(MC)、乳头状癌(PC)、小叶癌(LC)。

每个图像文件名存储关于图像本身的信息:程序活组织检查方法、肿瘤类别、肿瘤类型、患者识别和放大因子。例如,SOB_B_TA-14-4659-40-001.png 是放大因子 40X 的图像 1,其是管状腺瘤的良性肿瘤,原始来自载玻片 14-4659,其通过程序 SOB 收集。

2.2 训练策略

从本实验中将数据集按倍数分为 4 组完成对 BreakHis 数据集的八分类,将它随机划分为 3 个部分,即验证集(20%)、测试集(20%)和训练集(60%),并且采取相同的预处理操作,划分为网络可以接受的输入 $299 \times 299 \times 3$ 大小,以保证验证和测试结果的准确性,并保存在 data_set_split_8.m 文件中。对网络进行模型训练和参数学习并在训练集中完成,以 40X 为例:学习率为 0.001,最大迭代为 2 390,验证集频率为 100,最小的批大小为 5,训练时间达 3 小时左右。在验证集中主要完成以下步骤:优化模型,在训练过程中对模型进行测试,根据测试结果进行微调;提升测试模型的识别和泛化能力并在测试集中完成。本实验对训练集进行了数据增强,验证集和测试集用于测试实验效果;加载 Inception-ResNet-V2 网络,并替换最终网络层,使得网络分类个数为 8;开始训练,并微调参数,使得实验结果最优。训练效果如图 3 和图 4 所示(彩图扫描 OSID 码可见,下文同),在迭代 2 390 次后,Loss 值为 1.31。

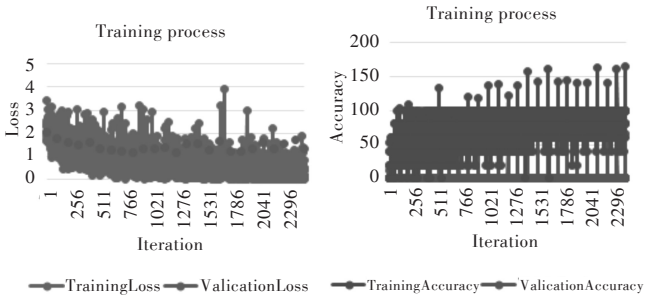


图 3 训练结果 Loss 值

图 4 训练结果 Accuracy 值

2.3 硬件及软件平台

实验中所用硬件配置为:64 位 Windows 操作系统,16GB 内存 CPU,GEFORCE GTX1050ti 显卡。软件方面:使用 2019 版 Matlab 以及图像处理工具箱。

2.4 实验结果及分析

本实验中使用正确率 ACC 为评价指标,ACC 的值越大,代表模型的预测和真实情况越接近,模型性能越好。假定 N 代表所有测试集图像数目, N_q 代表网络判断正确次数。则 ACC 的表达式为式(3)。

$$ACC = \frac{N_q}{N} \quad (3)$$

以方法倍数是 40X 的情况为例,由实验结果绘制而成的混淆矩阵如图 5 所示。

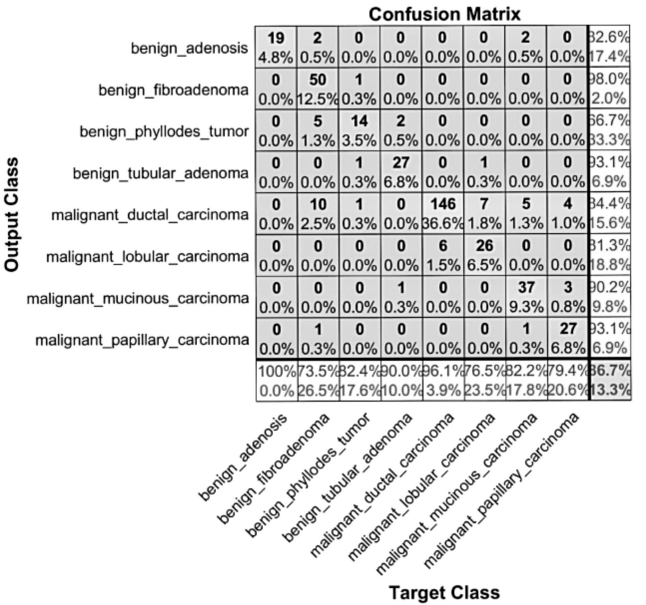


图 5 训练集数据格式

由图 5 可知,在 40 倍的实验测试数据中,绿色对角线上显示为正确分类的图像数目,而旁边红色的数据显示为网络判断错误的图像数目,混淆矩阵边缘数据为根据测试数据计算的准确率。以良性病 benign_adenosis 为例:测试集中共 23 张图片,网络正确分类图像数目为 19 张,错误分类成 benign_fibroadenoma 为 2 张,占测试集总数的 0.5%,被错分成 malignant_lobular_carcinoma 的数目为 2 张,占总测试集的 0.5%;正确分类占真正实例的 82.6%,错误分类比例占 17.4%。同理可知其余 7 类结果,得出最后准确率为 86.7%。同理,在其它倍数上也进行了实验汇总,得到如表 2 所示结果。

表 2 本文实验结果

内容	数据			
倍数(X)	40	100	200	400
ACC(%)	86.7	80.3	83.5	68.5

由此可见,倍数对于实验结果有很大影响。总体上讲,医学图像倍数越高,细胞组织越大,越难以区分内部环境,导致结果越差。近几年来,关于良恶性二分类的研究偏多,对比性不高,因此选取本文对比其研究者用过的二分类卷积神经网络进行实验,得到实验结果如表 3 所示。

表 3 对比实验结果				
内容	数据			
倍数(X)	40	100	200	400
ACC	86.7	80.3	83.5	68.5
网络结构(X)	40	100	200	400
GoogleNet(%)	68.7	65.9	69.1	62.8
ResNet50(%)	82.5	78.8	84.3	81
本文方法(%)	86.7	80.3	83.5	68.5

在 4 种放大倍数上对数据集进行分类,实验结果表明,本文使用的 Inception-ResNet-V2 深度卷积神经网络实验效果在 40X~200X 时效果优于其它两种近几年使用较多的卷积神经网络。但 400X 环境下深层次网络较复杂且层数较多,导致实验结果不如 ResNet50 效果好,可见放大倍数对实验结果的影响。

4 结语

本文使用以 Inception-ResNet-V2 为架构的深度卷积神经网络,将乳腺癌病理图像数据集分成 8 种不同类型的良恶性乳腺肿瘤。利用数据增强和迁移学习方法扩充数据量,从而解决因标记样本过少带来的过拟合问题。从实验结果可以看出,本文提出的方法具有较高的识别率,较其它良恶性肿瘤二分类能够更好地为临床医疗提供帮助。同时本文也存在不足,主要是公开可用的实验数据量少。因此,今后的改进方向是使用更多的医学真实数据进行测试,并采用更好的卷积神经网络以提高识别率。

参考文献:

[1] STEWART B, WILD C P. World cancer report 2014[M]. Lyon: World Health Organization, 2014.

[2] 何雪英,韩忠义,魏本征. 基于深度学习的乳腺癌病理图像自动分类[J]. 计算机工程与应用, 2018, 54(12): 126-130.

[3] KOWAL M, FILIPCZUK P, OBUCHOWICZ A, et al. Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images [J]. Computers in Biology and Medicine, 2013, 43 (10) : 1563-1572.

[4] ZHANG Y, ZHANG B, COENEN F, et al. One-class kernel sub-space ensemble for medical image classification[J]. EUR-ASIP Journal on Advances in Signal Processing, 2014(1): 17.

[5] SPANHOL F A, OLIVEIRA L S, PETITJEAN C, et al. A dataset for breast cancer histopathological image classification[J]. IEEE Transactions on Biomedical Engineering, 2016, 63(7): 1455-1462.

[6] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolutional neural networks[C]. Advances in Neural

Information Processing Systems, 2012: 1097-1105.

[7] CHEN H, DOU Q, WANG X, et al. Mitosis detection in breast cancer histology images via deep cascaded networks [C]. Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, 2016: 1160-1166.

[8] SPANHOLF A, OLIVEIRA L S, PETITJEAN C, et al. Breast cancer histopathological image classification using convolutional neural networks [C]. Vancouver: IEEE International Joint Conference on Neural Networks, 2016.

[9] BAYRAMOGLU N, KANNALA J, HEIKKILA J. Deep learning for magnification independent breast cancer histopathology image classification [C]. International Conference on Pattern Recognition (ICPR), 2016: 2441-2446.

[10] WEI B, HAN Z, HE X, et al. Deep learning model based breast cancer histopathological image classification [C]. Chengdu: IEEE International Conference on Cloud Computing & Big Data Analysis, 2017.

[11] 王恒,李霞,刘晓芳,等. 基于 Resnet50 网络的乳腺癌病理图像分类研究[J]. 中国计量大学学报, 2019, 30(1): 72-77.

[12] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1-9.

[13] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.

[14] SZEGEDY C, IOFFE S, VANHOUCKE V, et al. Inception-v4, Inception-ResNet and the impact of residual connections on learning [J]. ARXIV, 2016: 1-12.

[15] SZEGEDY C, VANHOUCKE V, IOFFE S, et al. Rethinking the inception architecture for computer vision [C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 2818-2826.

[16] SHEN D, WU G, SUK H I. Deep learning in medical image analysis [J]. Annual Review of Biomedical Engineering, 2017, 19 (1) : 221-248.

[17] 刘栋,李素,曹志冬. 深度学习及其在图像物体分类与检测中的应用综述[J]. 计算机科学, 2016(12): 20-30.

[18] 薛迪秀. 基于卷积神经网络的医学图像癌变识别研究[D]. 合肥: 中国科学技术大学, 2017.

[19] SHIN H C, ROTH H R, GAO M, et al. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning [J]. IEEE Transactions on Medical Imaging, 2016, 35(5): 1285-1298.

[20] BECKER A S, MUELLER M, STOFFEL E, et al. Classification of breast cancer in ultrasound imaging using a generic deep learning analysis software: a pilot study [J]. Br J Radiol, 2018, 91: 20170576.

(责任编辑:孙 娟)