文章编号：2096-1472(2022)-01-06-04
DOI:10.19644/j.cnki.issn2096-1472.2022.001.002

基于卷积神经网络的乳腺癌良恶性诊断

王 阳，陈薇伊，马军山

(上海理工大学光电信息与计算机工程学院，上海 200093)
✉wyusst2021@163.com; 18018592703@163.com; 1120157572@qq.com

摘 要：为了提高乳腺癌病理图像良恶性诊断的准确率，提出了一个基于卷积神经网络(CNN)对乳腺癌病理图像的诊断方法。利用这种方法，能够快速地对乳腺癌病理图像自动进行良恶性诊断。乳腺癌病理图像具有非常复杂的结构，利用VGG16架构的卷积神经网络对病理图像进行特征提取，利用数据增强的方法扩充数据集，使用迁移学习，将在ImageNet数据集上训练得到的权重作为该网络的初始化参数，该模型在乳腺癌数据集Breakhis上得到的准确率可以达到95%，而在经过解冻部分训练层、调整学习率等优化操作之后，分类准确率最高可以达到99%。实验结果表明，优化后的方法在乳腺癌良恶性诊断准确率方面有很大的提高。

关键词：乳腺癌；卷积神经网络；图像分类；数据增强；迁移学习
中图分类号：TP391 **文献标识码**：A

Diagnosis of Benign and Malignant Breast Cancer based on Convolutional Neural Network

WANG Yang, CHEN Weiyi, MA Junshan

(*School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai* 200093, *China*)
✉wyusst2021@163.com; 18018592703@163.com; 1120157572@qq.com

Abstract: In order to improve the accuracy of benign and malignant diagnosis for breast cancer pathological images, this paper proposes a diagnosis method for breast cancer pathological images based on convolutional neural network (CNN). This method makes a quick and automatic benign and malignant diagnosis for breast cancer pathology images. As breast cancer pathological images have very complex structures, VGG16 (Visual Geometry Group) architecture convolutional neural network is used to extract the features of pathological images, and data enhancement is used to expand the data set. By using transfer learning, weights trained on the ImageNet data set are used as the initialized parameters of the network. The model can achieve 95% accuracy on breast cancer data set Breakhis; the classification accuracy is as high as 99% after thawing some training layers and adjusting learning rate. The experimental results show that the optimized method can greatly improve the accuracy of breast cancer benign and malignant diagnosis.

Keywords: breast cancer; convolutional neural network; image classification; data enhancement; transfer learning

www.rjgczz.com

1 引 言(Introduction)

乳腺癌是目前全球女性发病率最高的癌症，对女性的身体健康造成非常大的伤害。世界卫生组织2018 年统计结果显示，乳腺癌在女性癌症新发病例中所占比例最高，并且死亡率也非常高[1]。

目前，诊断乳腺癌最常见方法的就是病理学诊断，并且大部分是由病理学专家借助显微镜观察完成的。我国目前正在进行乳房肿瘤系统的研究，大规模的临床应用有助于医生在诊断过程中大幅提高诊断效率[2]。但是，人为的主观意识存在偏差，长期的工作压力以及病理图像自身的复杂性都会导致误判。随着人工智能的发展，计算机辅助诊断治疗CAD(Computer Aided Diag-Nosis)已经成为医学领域研究的热门课题[3]。同人工诊断相比，计算机诊断更加客观，并且速度很快，可以同时完成多个病理图像的诊断，医生可以根

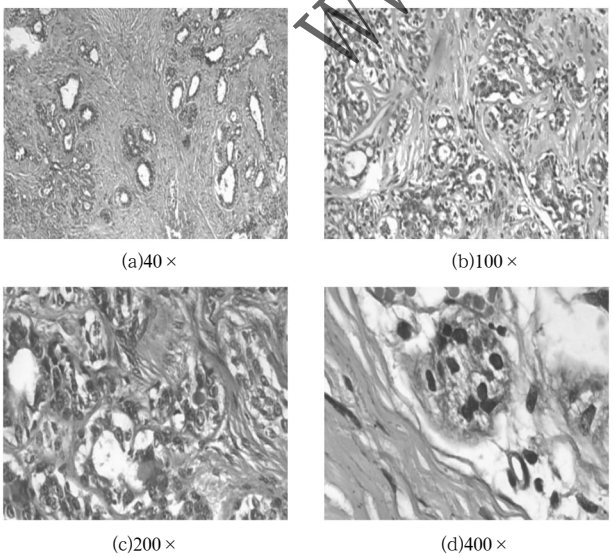
据计算机的诊断结果，结合自身的专业知识，做出更好的判断。目前基于计算机辅助诊断的主要有两种方式，一是人工提取特征和传统机器学习的乳腺癌病理图像诊断[4]；二是基于深度学习的乳腺癌病理图像诊断。

近些年，卷积神经网络(CNN)在图像分类方面取得了优异的成绩，在医疗图像等领域也有很大的进展[5]。2014 年，牛津大学的SIMONYAN[6]在“用于大规模图像识别得非常深卷积网络”的论文中提出了VGG卷积神经网络模型。该模型在ImageNet中实现了92.7%的top−5测试精度，这是一个1,000类的超过1,400万张图像的数据集。VGG卷积神经网络通过多个3×3大小的过滤器替换大型内核大小的过滤器(在第一个和第二个卷积层中分别为11和5)来改进AlexNet[7]。VGGNet证明了很小的卷积通过增加网络深度可以有效提高性能。目前主要的卷积神经网络模型有VGGNet、LeNet、GoogleNet、ResNet。通过卷积神经网络(CNN)构建的图像处理系统能够有效地减小过拟合，并且在对大像素图像内容识别方面也取得了不错的效果，特别是在融合GPU加速技术后，使得神经网络在实际训练中能够更好地拟合训练数据，最终更快更准确地识别大部分的图片。SPANHOL等在BreakHis数据集上采用卷积神经网络对乳腺癌病理图像进行良恶性诊断，准确度最高可达90%。

为了进一步提高乳腺癌病理图像诊断的准确度，本文以卷积神经网络VGG16网络为基础构造一个分类模型，并在公开的大型乳腺癌病理数据集BreakHis上进行研究。

2　数据集(Data set)

2016 年，SPANHOL等公布了BreakHis乳腺癌病理图像数据集。该数据集取自82 个病人，其中24 人为良性，58 人为恶性，共有7,909 张病理图片，分为四个放大倍数，即40×、100×、200×、400×。通过活体组织在不同放大倍数显微镜下成像得到如图1所示不同放大倍数下的良性癌病理图像。



(a)40×　(b)100×　(c)200×　(d)400×

图1 不同放大倍数的良性乳腺癌

Fig.1 Different magnifications of benign breast cancer

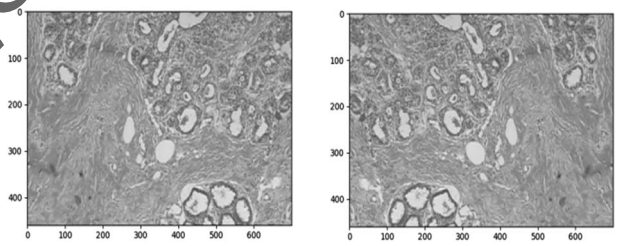
如表1所示是Breakhis数据集里的图像分布情况。

表1 图像类别和图像分布

Tab.1 Image categories and image distributions

放大倍数	良性	恶性	总计
40×	625	1,370	1,995
100×	644	1,437	2,081
200×	623	1,390	2,013
400×	588	1,232	1,820
总计	2480	5,429	7,909
患者数	24	58	82

尽管BreakHis数据集有7,000多张图片，但是对于神经网络来说这是不够的，因此我们需要用到数据增强技术。数据增强主要是可以起到减少网络过拟合的作用，一方面通过增加数据量提高模型的泛化能力，另一方面通过添加噪声提高模型的鲁棒性[8]。数据增强的方法主要有随机旋转、随机裁剪、添加噪声等[9]。如图2所示，图2(b)是由图2(a)经过反转得到的。对于肉眼来说，这是两张一样的图片，对图片进行翻转之后还是原来的图片，但是在神经网络看来，它们是完全不同的。



(a)原图　(b)翻转图

图2 翻转前后图

Fig.2 Diagrams before and after flipping

3　模型与方法(Model and methods)

3.1　VGG16架构

卷积神经网络的主要结构包括输入层、卷积层、池化层、全连接层和输出层。本文采用的VGGNet16模型包括五个卷积层和三个全连接层，之所以选择VGG16作为基础网络，是因为VGG16的分类性能很好，网络结构非常规整，修改起来比较容易。VGGNet用3×3的卷积核替代AlexNet里面11×11的卷积核，因为对于给定的感受野，采用多个小的卷积核效果要比大的卷积核好，多层的非线性层能够增加网络的深度，从而保证网络能够学到更深层的东西，并且可以更好地保持图像性质。另外，用多个小卷积核代替大卷积核也能够减少参数的作用。VGG16模型结构如图3所示。

www.rigcz.com

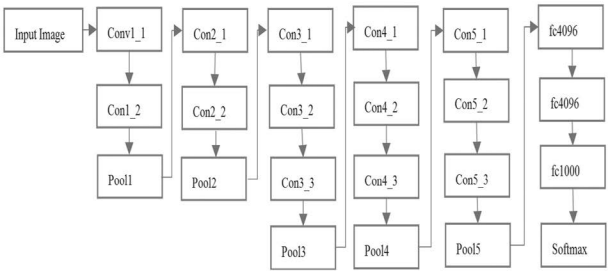


图3 VGG16结构图

Fig.3 VGG16 structure map

输入层负责载入图像，本文的输入图像就是经过数据增强后的BreakHis数据集。输入图像的大小为224×224×3，3代表通道数，产生的输出向量作为卷积层的输入。

卷积层主要负责提取图像特征，通过卷积核实现，卷积核就是给定一个输入图像，输入图像中的一块区域像素经过加权平均后成为输出图像的对应像素，权值是一个给定函数，这个函数就叫做卷积核。对于输入的乳腺癌图像，经过多个卷积层之后提取到每张图像的特征。

VGG16模型共有五个卷积层和三个全连接层，16指的是它的网络深度。每个卷积层中都用到Relu激活函数，前两个卷积层使用两次大小为3×3、步长为1的卷积核进行卷积操作，后三层用了三次同样的卷积核。为了提高模型的训练速度，在卷积层之后添加了Batch Normalization层。

池化层也叫下采样层，一般放在连续的卷积层之间，用于压缩数据和参数。池化层的作用主要有两个，一是特征不变性，池化操作就是图像的缩放，压缩去掉一些无关紧要的信息，留下的信息依然能够保留原图的特征；二是特征降维，一幅图像里的特征有很多，但有些信息没有太大的作用或者重复，就可以去掉这些信息，保留最重要的特征。常见的下采样有最大池化(max-pooling)和均值池化(mean-pooling)，这里使用的是最大池化，池化单元尺寸大小为2×2，可以对输入图像尺寸减半。

卷积取的是局部特征，全连接指的就是把这些局部特征重新通过权值矩阵组装成完整的图，全连接层的每一个节点都是跟上一层的所有节点相连的，用来把前面提取到的特征综合起来。VGG16网络有三个全连接层，前两层都用了4,096个神经元，第三层使用了1,000 个神经元，常用的激活函数有Sigmoid、Tanh和Relu[10]，本实验使用的是Relu函数，最后使用Softmax对图像进行分类，输出1,000 个预测结果。利用Softmax函数可以很好地处理多分类问题，因为本文只需要分类乳腺癌是良性还是恶性，所以分类标签为2。在全连接层之前还添加了Dropout层[11]，为的是防止过拟合，提高模型的泛化能力。

损失函数Loss描述的是预测结果和实际结果之间的差别，模型的每一次预测的好坏用损失函数来度量，损失越小，模型越好。本文所用的损失函数是sparse_categorical_crossentropy。

3.2 迁移学习及改进

迁移学习[12]，顾名思义就是把已经训练好的模型参数迁移到新的模型上，从而帮助新模型进行训练。神经网络在训练的时候能够随机初始化参数，从头训练，由于大部分数据或者任务都是存在相关性的，因此通过迁移学习可以将已经学到的模型参数通过某种方式传递给新的模型，从而加快训练速度并优化模型的学习效率。

在卷积神经网络中，很多模型都是在ImageNet数据集上进行训练的，训练得到的参数通过迁移学习使用在其他网络上。ImageNet数据集由超过1,400万张图像组成，每张图像都包含注释，标明物体属于哪种类别。其实不管是哪种图像，它们的边缘特征都是相似的，所以考虑到了迁移学习。迁移学习可以将已经训练好的模型参数直接用在自己的数据集上，但是训练结果不一定会非常好。

本文所用的迁移学习的权重来自VGG16网络在ImageNet数据集上训练得到的参数，这里将其应用到乳腺癌病理组织的良恶性分类诊断问题上。针对网络训练结果的Loss偏大，对模型进行微调，之前是冻结预训练模型的所有卷积层，只训练自己要求的全连接层；现在只冻结指定预训练模型的部分卷积层，这里设置的是解冻卷积层的后五层，训练剩下的卷积层和自己的全连接层，原学习率为0.001,这里将学习率调整为原来的1/10，并添加了Dropout层，将丢弃率设置为0.5，又添加了Batch Normalization层。

3.3 实验结果

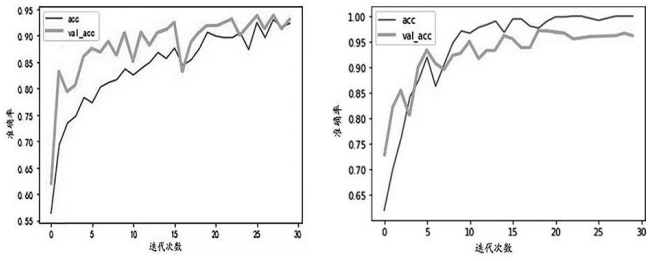
本文使用Thinkpad T490、Intel i5，基于Python语言的tensorflow框架进行训练和测试。

本文采用识别分类准确率对实验效果进行评估，以7∶2∶1划分训练集、验证集和测试集。用N_r表示正确分类出来的图片数量，用N_a表示输入病理图片的数量，图像水平的识别准确率(Image_level_Rate)表达式如下式所示：

$$Image_level_Rate = \frac{N_r}{N_a}$$

本文设置了两组对比实验，将迁移学习后的网络直接进行训练得到的准确率和微调后的模型准确率进行对比，图4(a)是微调前训练集和验证集的准确率，图4(b)是微调后的训练集和验证集的准确率。

www.rjgcz.com



(a)微调前准确率　　(b)微调后准确率

图4 微调前后准确率

Fig.4 Accuracy before and after fine-tuning

从图4可以明显看出，微调之后的网络在准确率性能上表现得更好。首先从图4(a)可以看出微调前的模型，在经过大约25 次迭代后，准确率已经达到了95%，这说明在模型中增加了数据增强和迁移学习模块能够使得模型快速地处理乳腺癌病理图像，使模型具有更好的鲁棒性和泛化能力。然后在经过微调后，模型能够更加快速地收敛，只需要经过20 次迭代，训练集准确率便达到了99%，同时验证集也达到了95%以上。最后利用测试集对模型进行测试，微调前测试集准确率是90%，而微调后测试集准确率达到了98%。表2是微调前后测试集的Loss对比，可以看出微调后的Loss下降更快，这进一步说明微调后的模型更适合使用BreakHis数据集进行良恶性诊断。

表2 微调前后测试集Loss值

Tab.2 Test set Loss values before and after fine-tuning

迭代次数	5	10	15	20	25	30
微调前的Loss	0.4806	0.3488	0.3388	0.2673	0.2689	0.2182
微调后的Loss	0.1702	0.0037	0.0007	0.0006	0.0002	0.0049

4 结论(Conclusion)

本实验在公开乳腺癌数据集BreakHis上进行，同时使用卷积神经网络模型VGG16，并且为了减少训练时间，以及避免训练过程中可能存在的误差，这里使用了迁移学习，将VGG16在大型数据集ImageNet上训练得到的权重用在本实验上。为进一步提高模型性能，又进行了模型微调，解冻了模型训练层的后五层进行训练，试验结果表明，微调后的Loss明显小于微调前，准确率也从90%提高到了98%，充分证实了本实验方法的可行性和实用性。未来的工作包括：(1)在乳腺癌二分类诊断的基础上，进一步细分是哪种癌；(2)将本实验的模型用在其他医学疾病上，验证该模型在其他疾病上是否具有通用性。

参考文献(References)

[1] BRAY F, FERLAY J. Global cancer statistics 2018: Globocan for estimates of incidence and mortality worldwide 36 cancers in 185 countries[J]. A Cancer Journal for Clinicians, 2018, 68(6):394-424.

[2] 陈诗慧,刘维湘,秦璟.基于深度学习和医学图像的癌症计算机辅助诊断研究进展[J].生物医学工程杂志,2017,34(02):314-319.

[3] ERICKSON B J, KORFIATIS P, AKKUS Z, et al. Machine learning for medical imaging[J]. RadioGraphics, 2017, 37(2):505-515.

[4] 龚磊,徐军,王冠皓.基于多特征描述的乳腺癌肿瘤病理自动分级[J].计算机应用,2015,35(12):3570-3575.

[5] SPANHOL F A, OLIVEIRA L S, PETITJEAN C, et al. A dataset for breast cancer histopathological image classification[J]. IEEE Transactions on Bio-medical Engineering, 2015, 63(7):1455-1462.

[6] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[DB]. Eprint ArXiv, 2014:1409-1556.

[7] RAJKOMAR A, LINGAM S, TAYLO A G, et al. High-throughput classification of radiographs using deep convolutional neral networks[J]. Journal of Digital Imaging, 2017, 30(1):95-101.

[8] LI X, TANG W L, YANG B. Semantic segmentation of high resolution remote sensing image using depth residual network[J]. Journal of Applied Sciences, 2019, 37(2):282-290.

[9] YU X M, HONG S, YU J X, et al. Research on the enhancement method of visible light remote sensing image ship target data[J]. Journal of Instrumentation, 2020, 41(11):261-269.

[10] ZHANG Y D, MUHAMMAD K, TANG C. Twelve-layer deep convolutional neural network with stochastic pooling for tea category classification GPU platform[J]. Mutimedia Tools and Applications, 2018, 77(17):22821-22839.

[11] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: A simple way to prevent neural networks from overfitting[J]. The Journal of Machine Learning Research, 2014, 15(1):1929-1958.

[12] PAN S J, YANG Q. A survey on transfer learning[J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10):1345-1359.

作者简介：

王　阳(1995-)，男，硕士生.研究领域：计算机视觉，深度学习.

陈薇伊(1997-)，女，硕士生.研究领域：计算机视觉.

马军山(1967-)，男，博士，教授.研究领域：光学精密仪器，图像处理.

www.rigczz.com