

基于深度学习的胃癌病理图像分类方法

张泽中¹ 高敬阳¹ 吕纲^{2,3} 赵地⁴

(北京化工大学信息科学与技术学院 北京 100029)¹ (国家转化医学中心(上海) 上海 200025)²
(国家人类基因组南方中心 上海 201203)³ (中国科学院计算技术研究所 北京 100190)⁴

摘要 针对深度卷积神经网络能够有效提取图像深层特征的能力,选择在图像分类工作中表现优异的 GoogLeNet 和 AlexNet 模型对胃癌病理图像进行诊断。针对医学病理图像的特点,对 GoogLeNet 模型进行了优化,在保证诊断准确率的前提下降低了计算成本。在此基础上,提出模型融合的思想,通过综合不同结构和不同深度的网络模型,来学习更多的图像特征,以获取更有效的胃癌病理信息。实验结果表明,相比原始模型,多种结构的融合模型在胃癌病理图像的诊断上取得了更好的效果。

关键词 深度学习,卷积神经网络,胃癌病理图像,GoogLeNet 优化,模型融合

中图分类号 TP391.7 **文献标识码** A

Pathological Image Classification of Gastric Cancer Based on Depth Learning

ZHANG Ze-zhong¹ GAO Jing-yang¹ LV Gang^{2,3} ZHAO Di⁴

(College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, China)¹

(National Research Center of Translational Medicine(Shanghai), Shanghai 200025, China)²

(Chinese Human Genome Center at Shanghai, Shanghai 201203, China)³

(The Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China)⁴

Abstract Due to that CNN can effectively extract deep features of the image, this paper used GoogLeNet and AlexNet models which have excellent performance in image classification to diagnose the pathological image of gastric cancer. Firstly, according to the characteristics of medical pathological images, this paper optimized the GoogLeNet model to reduce the computational cost under the premise of ensuring the accuracy of diagnosis. On this basis, it proposed the idea of model fusion. By combining more images with different structures and different depths, more effective pathological information of gastric cancer can be acquired. The experimental results show that the fusion model with multiple structures has achieved better results than the original model in the diagnosis of pathological images for gastric cancer.

Keywords Deep learning, CNN, Gastric cancer pathology, GoogLeNet optimization, Model fusion

1 引言

胃癌是一种发病率和死亡率都较高的肿瘤,严重威胁着人们的健康和生命^[1]。全球每年新增胃癌病例达 90 多万例,有 70 多万人因此而死亡,死亡率仅次于肺癌,居第二位^[2]。癌症在欧美等发达国家的发病率和死亡率呈现逐年下降的趋势,但在发展中国家的发病率仍然较高。全球约 50% 的胃癌病例集中在东亚地区,2012 年我国新增胃癌患者 42.4 万,占全球总量的 40% 以上,每年因胃癌死亡 29.8 万人,发病率和死亡率分别位居全球第五位和第六位。在中国,胃癌已成为仅次于肺癌的第二大恶性肿瘤^[3]。

对胃癌的诊断主要依靠病理图像诊断和医学影像诊断两种方法,其中病理图像诊断是诊断胃癌最可靠的方法,相比医学影像诊断等其他方法而言,病理图像诊断更加准确,是确定诊断结果的依据^[4]。由于病理图像的特点,医生在长时间工

作后容易产生视觉疲劳,因此难免发生漏诊、误诊等情况。近年来随着计算机技术的迅速发展,深度学习在图像处理方面正不断取得突破性进展^[5],而医学图像作为一种结构化的数据,易于使用深度学习的方法对其进行处理^[6]。深度学习在乳腺癌、前列腺癌和脑部疾病等领域的应用^[7-9]已经有了广泛而深入的研究,其诊断的准确率都达到了 95% 以上,它不要求操作人员有辨别癌症组织的专业知识,其通过自动提取病理组织代替人工提取^[10]。在腹部疾病的研究中, Tirumala 等在前列腺癌的诊断中使用人工神经网络取得了 98.2% 的分类精度^[11];在宫颈癌的诊断中, Xu 等提出的深度学习框架也取得了 87.33% 的灵敏度和 90% 的特异性^[12],均达到了很高的准确率^[13]。同时,深度学习极大地减轻了医生的工作量,因此具有很高的实用价值^[14]。

本文将深度学习的方法应用于胃癌病理切片的识别中,经过大量的研究之后,以 GoogLeNet 和 AlexNet 两种深度学

本文受国家自然科学基金(61472026),国家重点研究发展计划(SQ2017ZX106047),北京市自然科学基金重点项目(4161004),北京市科技计划项目(Z171100000117001),北京市科技计划项目(Z161100000216143),北京市自然科学基金资助项目(5182018)资助。

张泽中(1994—),男,硕士生,CCF 会员,主要研究方向为深度学习、医学图像数据分析;高敬阳(1966—),女,博士,教授,CCF 高级会员,主要研究方向为机器学习、生物信息学,E-mail:gaojy@mail.buct.edu.cn(通信作者);吕纲(1976—),教授,主要研究方向为机器学习,E-mail:lug@chgc.sh.cn;赵地(1978—),男,博士,副教授,CCF 会员,主要研究方向为类脑计算,E-mail:zhaodi@esience.cn。

习网络模型^[15-16]为基础实现对胃癌的自动诊断。AlexNet 模型和 GoogLeNet 模型分别获得了 ILSVRC2012 和 ILSVRC2014 的冠军,均具有出色的分类能力。针对胃癌病理图像的特点,对两种模型中时间效率相对较低的 GoogLeNet 模型进行了优化,然后将两种模型相融合,综合两种不同的网络模型提取到的不同特征,来提高胃癌诊断的准确率。

本文使用的数据集包括了 700 张胃癌病理样本,其中带有标注的癌症样本有 560 张,非癌症样本有 140 张,数据来自 2017 中国大数据人工智能创新创业大赛首次开放的胃癌病理切片标注数据集,这是国内首次开放的胃癌病理切片数据,在此向大赛组委会表示感谢。该大赛由上海大数据联盟、大数据产业基地、中国信息通信研究院主办,旨在通过比赛的形式,实现技术上的创新,推进人工智能的技术在癌症的诊断与治疗中的应用。

2 深度学习方法的介绍

2.1 深度学习和卷积神经网络

深度学习是以人工神经网络为基础的一项技术,是机器学习领域兴起的一个新的研究方向。2006 年,加拿大多伦多大学的 Hinton 教授等第一次提出了深度学习的概念,此后 10 年左右的时间里,深度学习在计算机视觉等多个领域中不断取得突破性进展^[17]。深度学习的思想是建立有多层隐藏层的神经网络,通过组合低层特征,形成更为抽象的高层特征来表示数据的属性和特征等信息,进而发现数据的分布式特征表示^[18]。

相比传统机器学习的浅层网络,深度学习使用更深层次的网络和非线性相关^[19]映射结构模仿人类大脑的神经元工作原理,通过提取数据更高层次的特征,来发现数据在高层次特征中的关系^[20],解决传统机器学习方法存在的梯度消失和容易陷入局部最优解等问题^[21]。图 1 给出了一个简单的深度学习网络模型。

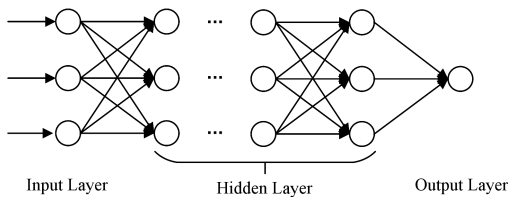


图 1 多层神经网络结构图

从图 1 可以看出,深度学习与传统的机器学习的不同点在于网络的层次结构更深。但当网络层次比较深时,误差传播到前面的层次时就会接近消失。Chen 等于 2006 年提出了一种基于无监督方法的深层神经网络^[7],通过自底向上的非监督学习和自顶向下的监督学习,使用 wake-sleep 算法对参数进行调优,有效地解决了网络深度上的问题^[22-23]。

卷积神经网络(CNN)是一种前馈型的神经网络,它的历史最早可以追溯到 1962 年。Hubel 和 Wiesel 两位生物学家通过研究猫的视觉发现,在猫的视觉皮层中一些细胞对视觉输入的部分区域十分敏感,据此提出了感受野的概念^[24]。Kunihiko Fukushima 于 1980 年根据 Hubel 和 Wiesel 的局部感受野理论,提出了神经认知机模型,这是人们最早实现的 CNN 网络模型^[25]。

卷积神经网络的优点在于局部连接和权值共享。在传统的人工神经网络中,一个神经元与其相邻层中的所有神经元

全部建立连接,这样随着网络规模的扩大,计算量将呈指数式增长。而卷积神经网络吸收了局部感受野的思想,在卷积层中,一个神经元只通过卷积核与其上一层感受野内的神经元建立连接^[16],从而有效地控制了参数规模和计算量。卷积核实际是一个由权值组成的矩阵^[26],根据实际情况不同,可以赋予卷积核不同的大小。卷积神经网络结构包括卷积层、池化层和全连接层,其中卷积层和池化层交替出现,全连接层一般设置在网络末尾。卷积神经网络使用局部连接、权值共享和下采样的方法^[5],能够有效地提取数据特征。图 2 为在胃癌病理切片上进行卷积和池化操作的示意图。

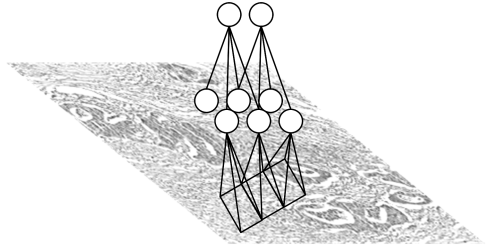


图 2 病理切片上的卷积和池化操作示意图

根据图像的统计学特性,我们可以使用一个共享的卷积核对所有不同的感受野执行卷积操作,抽象出图像的局部特征。对于多通道的特征图,卷积结果为上一层 N 个特征图中卷积核对应区域卷积操作的总和加上一个偏置值。若使用 W 表示目标特征图中要求的特征值, W^k 表示上一层的第 k 个特征图, i 和 j 表示特征图中当前像素的位置, x 表示像素值, b 表示偏置值,则目标特征图的卷积计算方法如式(1)所示:

$$W = \sum_{k=1}^N W_{ij}^k x + b \quad (1)$$

虽然卷积操作相比传统的全连接操作已经极大地减少了网络中的参数和连接数量,但在深度神经网络中,经过多次卷积操作后,随着特征图数量的增加,仍然会增加训练的难度,特征维度的增大也更容易产生过拟合现象,因此在深度神经网络中几乎都设置有池化层。池化操作能够在保留图像特征的情况下降低特征维度,一方面节约运算资源,加快训练速度,另一方面可以有效地避免过拟合。池化的方法主要有平均值法、最大值法、随机采样法等。

相比全连接网络,卷积神经网络极大地减少了参数和连接数量,降低了网络运算的复杂度,具有很强的健壮性^[27],特别是当图像数据较大时,卷积神经网络的优势尤为明显。

2.2 AlexNet 和 GoogLeNet 模型

AlexNet 是由 Alex Krizheviky 等设计的一种深度学习网络模型,它获得了 2012 年 ImageNet 竞赛的冠军,具有出色的图像分类能力。AlexNet 模型的结构如图 3 所示。

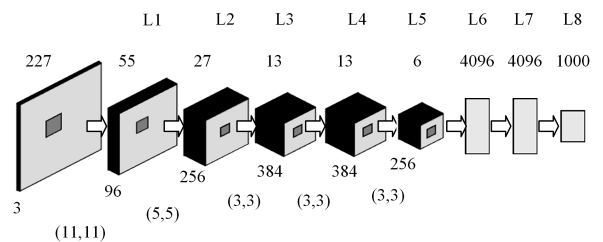


图 3 AlexNet 网络模型结构图

AlexNet 模型的输入为 $227 \times 227 \times 3$,本文使用的病理图像为三通道的 RGB 图像,维度恰好符合模型的要求。将病理图像调整为 $227 \times 227 \times 3$ 大小后送入网络,首先到达 conv

层,经过第一层卷积操作后,进入 ReLU 层激活,然后将得到的结果送到 pool 层池化,最后经过 norm 层标准化后,输入网络第二层^[28]。后续第二层到第五层的操作与第一层的操作类似,因此不再赘述。

网络第五层的输出结果将送入到后面的全连接层。第六层和第七层的输出均为长度为 4096 的向量,经过第八层后将得到一个长度为 1000 的向量。网络最后使用 Softmax 分类器^[29]进行二分类,得到最后的分类结果。Softmax 分类器的原理如式(2)所示:

$$y_c = \frac{e^{z_c}}{\sum_{d=1}^c e^{z_d}} \quad (2)$$

其中, z 为输入向量; c 为向量的维度; y 为分为某一类的可能性,取值范围为 0~1。

GoogLeNet 是 ILSVRC14 竞赛的冠军,该模型具有 22 层的网络结构。由于增加了网络的深度和宽度,如果使用原有的深度学习架构将会产生巨大的参数量,从而带来巨大的计算量,也容易造成过拟合。GoogLeNet 认为解决此问题的方法是采用稀疏连接操作^[30],它吸收了 NIN^[31]的思想,设计了一种称为 Inception^[17]的模块化网络结构,将稀疏矩阵聚类成为比较密集的子矩阵,极大地降低了参数量,提高了计算效率。Inception 的基本结构如图 4 所示。

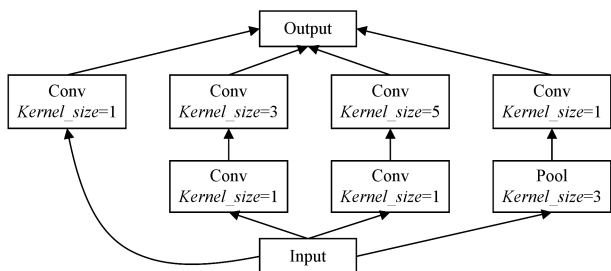


图 4 Inception 结构图

Inception 是一种稀疏性的结构,它采用了 3 种不同大小的卷积核学习图像的特征。由于采用了不同大小的卷积核,Inception 既能够捕捉到小范围内的相关性单元的特征,也能捕捉到距离较远的相关性特征。同时,由于较大的卷积核会带来参数量增加的问题,Inception 在卷积层前增加了 1×1 的卷积核对数据进行降维,从而提高了计算效率。

GoogLeNet 模型中使用了 9 个 Inception 模块,其结构如图 5 所示。随着网络层数的增加,图片的视野将会增大,因此在更深的网络层次上,模型增加了卷积核的个数以更有效地把握图像特征。在卷积操作完成后执行全连接操作,GoogLeNet 的全连接操作采用与 AlexNet 相似的结构,最后使用 Softmax 分类器进行分类,得到最终的分类结果。

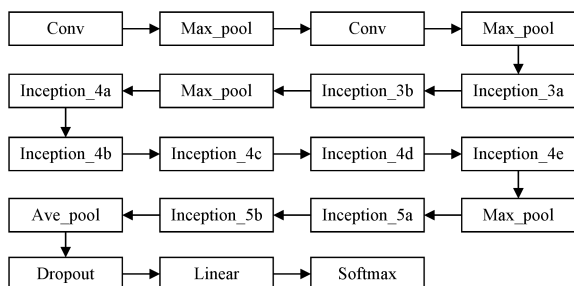


图 5 GoogLeNet 结构图

3 方法

3.1 实验工具

首先将病理图像进行叠加、裁剪等预处理,得到可以用于训练深度神经网络的数据集,然后将数据送入网络模型中进行训练。训练工具使用 NVIDIA 公司推出的深度学习软件 DIGITS 6.0,深度学习框架采用了 Caffe^[32]。DIGITS 是一款可视化的深度学习软件,它能够帮助用户更方便地管理自己的工作项目,并极大地简化了模型训练中的操作。Caffe 是一种基于 C++ 的开源深度学习框架,用户可以直接调用它的底层算法构建网络,而不需要自己实现网络的具体操作。这给深度学习的应用带来了极大的便捷。

3.2 数据集和预处理

胃癌病理图像是取自患者胃部的组织样本,经过脱水、浸蜡、包埋、切片以及染色等一系列技术处理后在显微镜下观察取得的显微图像。病理图像的精度达到细胞级,可以帮助医生从更本质的角度对胃癌做出判断。本文使用的胃癌病理图像来自于 2017 中国大数据人工智能创新创业大赛,该病理图像数据采用的是已经非常成熟的 HE 制片技术,放大倍数为 20 倍,图像为 RGB 三通道位图图片,并经过了标准化处理。样本分为有癌症和无癌症两类,每张癌症样本图像对应一张 label 图片,标注出了图像上典型的癌症区域,但并非全部癌症区域。胃癌病理图像和对应癌巢区域的 label 图像如图 6 所示。

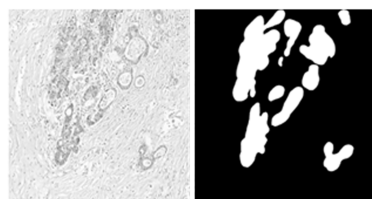


图 6 胃癌病理切片与对应的 label

该数据集包含了 700 张带有标注的胃癌病理切片图像。我们需要从中划分出一部分用作测试集,以验证模型的泛化表现。原始的 700 张病理图像数据虽然具有一定的规模,但远无法满足深度学习的要求。深度学习的方法要求训练集必须具有一定大的规模,因此我们首先对数据进行预处理,以达到训练深度学习神经网络对数据的需求。

在数据的预处理中,考虑到癌症样本的 label 并没有包括样本中的全部癌症区域,因此首先将癌症图像与对应的 label 文件进行了叠加处理,筛选出了病理图像上的癌巢部分,以排除未标注区域可能产生的干扰,处理后的病理图像如图 7 所示。



图 7 使用 label 进行过滤后的病理图像

将一张病理图像按照 128×128 像素的尺寸分割成了更小的图像,并按照一定的癌症区域占比对子图像进行了筛选。

为了保证数据对于模型训练的有效性,实验中取癌巢区域占比在50%以上的图像作为有效数据。实验通过这个方法大大增加了样本的容量,有效地解决了数据量不足的问题。分割的子图像如图8所示。

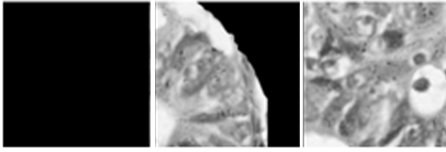


图8 分割后的病理图像

由于实验数据存在正负样本量不均衡的问题,通过分割方法对数据进行处理后该问题仍然存在,因此通过旋转、镜像等操作对数据进行了扩增,以解决样本不平衡的问题。经过预处理后,癌症样本和非癌症样本各自接近10万张,能够满足深度学习的训练要求。

实验取训练数据的25%用作验证集,训练集与验证集的数据分布如表1所列。最后对训练数据进行resize操作,采用拉伸的方式将图像数据调整为 224×224 像素,以适应神经网络模型的输入标准。

表1 训练数据的分布

类别	训练集	验证集
有癌症	73710	24570
无癌症	72576	24192

3.3 GoogLeNet 模型优化

GoogLeNet 在图像分类中的表现优异,但它对计算资源的需求比 AlexNet 高更多。GoogLeNet 是一个22层的深层次神经网络,针对本文中使用的病理图像,GoogLeNet 的网络层数完全能够满足诊断的需求。而由于网络层级较深,模型在提取到更深层次特征的同时,感受野的扩大反而会带来特征消失的问题,从而对识别精度造成影响。因此,针对实验所使用的病理图像的特点,对 GoogLeNet 的结构进行了优化。

GoogLeNet 是一个模块化的深度学习网络,它的每个 Inception 模块具有相同的结构。从 GoogLeNet 的训练表现来看,网络在第6个 Inception 后产生的中间结果与最后结果的准确率几乎相同。针对这一现象,实验以 Inception 为单位对网络结构进行优化。经过多次实验,最终确定网络在第7个 Inception 后的输出便达到了最优水平。因此,实验保留了 GoogLeNet 的前7个 Inception 结构,而在第7个 Inception 模块后直接使用 FCN 对特征图进行分类输出。优化后的网络节省了较多的计算资源,其与原网络在 Tesla K20c GPU 上运行的性能对比情况如表2所列。

表2 优化的 GoogLeNet 与原网络的对比

网络模型	显存/GB	GPU 使用率/%	训练时间/h
GoogLeNet	3.52	65	21
优化的 GoogLeNet	2.93	43	17

3.4 AlexNet 与 GoogLeNet 的模型融合

AlexNet 和经过针对性优化的 GoogLeNet 在对胃癌病理图像的诊断中均可取得了较好的效果,而两者在网络结构上的差异很大。AlexNet 是8层结构,能够有效提取到图像的局部特征;GoogLeNet 有22层,对图像整体特征的把握优于 AlexNet。我们可以看到,在 GoogLeNet 中吸收了 NIN 的思想而设计的 Inception 结构取得了很好的效果,证明不同大小

卷积核的并联有助于对数据特征的提取。而如果将这一思想扩展到更深层次即网络模型级的融合,是否能够实现不同网络的优势互补,取得较好的效果呢?因此,实验提出将 AlexNet 与 GoogLeNet 两者相融合,综合两者所提取到的不同病理特征,以求在对胃癌病理图像的诊断中取得更好的效果。本实验对不同的模型采用加权的方法,通过多次实验寻找不同模型的最优权重,使融合的网络模型达到最好的诊断效果。使用的加权公式如下:

$$p = \sum_{i=1}^n \omega_i p_i \quad (3)$$

AlexNet 网络有8层结构,优化后的 GoogLeNet 网络有18层结构,两种网络结构之间的主要差异在于卷积层的操作:AlexNet 的每一层都使用单一大小的卷积核,进行了5次卷积操作;而 GoogLeNet 的 Inception 结构中并联了不同大小的卷积核。在全连接层,两种网络所采用的结构基本相同。因此,实验将两种网络结合的重点放在卷积层的融合使用上。融合的模型使用同一个训练集进行训练,因此需要对输入和输出进行标准化处理。实验通过对 AlexNet 网络的卷积核的大小进行调整来修改其输入输出标准,将网络的输入统一为 $224 \times 224 \times 3$ 的三通道图像,卷积层的特征图输出统一为 7×7 大小。数据分别经过 AlexNet 和 GoogLeNet 的卷积层得到两种不同的特征图像,将得到的两种不同特征图按照不同的权值进行分配,通过尝试不同的权值,使网络达到最好的诊断效果。该网络算法的结构如图9所示。

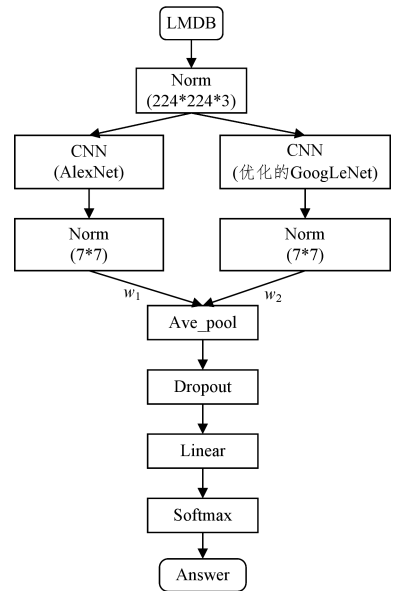


图9 AlexNet 与 GoogLeNet 相融合的模型结构

经过多次实验发现,将两种不同的网络融合后,对胃癌病理图像诊断的准确率得到了较为明显的提高。当 ω_1 与 ω_2 之比为1:3时,融合的网络模型取得了最佳的诊断结果,证明实验对 NIN 应用的深度进行拓展的思想是有效的。

实验通过30次迭代训练得到最终的模型。训练中学习率采用梯度下降的方法,训练的初始阶段较大的学习率避免了陷入局部最优解的情况,后期学习率的下降使模型能够收敛得到最优解。对于实验提出的融合模型,由图10可以看到,其在训练后期已经收敛到了稳定值,对分割处理后的病理切片诊断具有极高的精度。

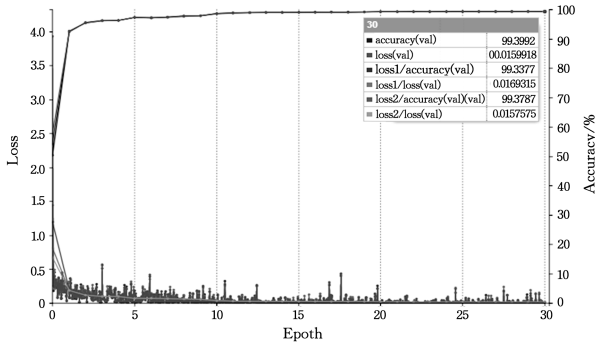


图 10 融合模型的训练折线图

4 实验结果和分析

4.1 实验过程及结果

本文将训练数据以有无癌症为标签建立训练集。首先对 AlexNet 和 GoogLeNet 两种网络模型分别进行实验,根据训练结果,针对 GoogLeNet 占用计算资源较大的缺点,对其进行了改进以得到优化的 GoogLeNet 模型。两种网络模型在对病理图像的诊断中均取得了较好的效果,因此将两种网络模型相融合,以综合不同结构的网络模型,从而提取到不同特征,获得更好的诊断结果。因此,实验一共设置了 AlexNet 模型、优化的 GoogLeNet 模型和融合网络模型 3 组神经网络模型。通过多次实验找到模型融合的最优点后,对比 3 种模型在训练过程中的表现,最后通过测试数据验证模型的泛化能力,对实验结果进行分析和总结。

本实验分别对 3 种网络模型进行训练,得到它们在验证集上的诊断精度,如表 3 所列。

表 3 训练集上的诊断准确率与损失

网络模型	准确率/%	损失
AlexNet	99.14	0.024
优化的 GoogLeNet	99.28	0.018
融合网络模型	99.40	0.016

可以看到,本文提出的融合模型相比于 AlexNet 和优化的 GoogLeNet 两种网络模型,在对验证集诊断的准确率上获得了一定的提高。若要验证模型在实际应用中是否也能取得较好的效果,则需要测试和比较 3 种模型的泛化能力。测试在 DIGITS 平台上进行,DIGITS 提供了非常强大的可视化功能,我们能够通过它实现模型各层计算结果的可视化。以 GoogLeNet 的第一层为例,病理图像经过卷积操作、ReLU 激活和池化操作后,将输出进行可视化显示,如图 11 所示。

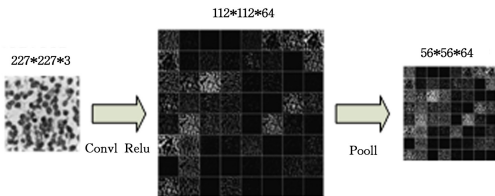


图 11 GoogLeNet 第一层的可视化过程示意图

因为网络中每一次的卷积和池化除了特征图的尺寸不同外,在操作上均是相似的,因此这里对每一次的特征提取操作不再赘述。在卷积操作之后,网络通过全连接层将得到的最终特征图转化为 1024 维度的特征向量,通过 Softmax 分类器进行分类,输出最终的 label 值。

在测试模型的泛化能力时,测试集使用与训练集相似的

处理方法,因此在此不再赘述。根据有癌症病例图像的癌症标注范围,从 52 个癌症病理图像中获得了 2208 张癌症切片,从 13 张非癌症病理图像中获得了 3328 张非癌症切片。

灵敏度和特异度分别是用来量化实验或模型,能够判断出有病者和无病者能力的指标,是在医学领域中广泛使用的一种衡量标准。灵敏度又称真阳性率(TPR),指患者诊断结果为阳性的概率,灵敏度越高,则越不容易漏诊;特异度又称真阴性率(TNR),指将非癌症个体诊断为阴性的概率,特异度越高,越不容易误诊。本文中 3 种模型的灵敏度和特异度如表 4 所列。

表 4 3 种不同模型的灵敏度和特异度

(单位:%)

网络模型	灵敏度	特异度
AlexNet	96.20	99.61
优化的 GoogLeNet	96.65	99.73
融合网络模型	97.60	99.49

4.2 分析

由表 4 可见,在对胃癌病理图像的诊断中,3 种深度学习模型均取得了很高的诊断准确率,超过了传统人工神经网络在乳腺癌诊断上的效果^[7],也超过了其在前列腺癌诊断中的表现^[12]。本实验提出的融合网络模型的效果更佳,在测试集上的灵敏度相对于 AlexNet 和优化的 GoogLeNet 模型分别提高了 1.40% 和 0.95%。灵敏度指标意义重大,灵敏度的提升意味着漏诊率的下降,而癌症的治疗贵在早期发现,少漏诊并及时确诊对于癌症患者来说争取了宝贵的早期治疗时间,因此灵敏度的提升对于实际应用有着非常重要的意义。表 4 中,本实验特异度指标相比 AlexNet 和优化的 GoogLeNet 有 0.12% 和 0.24% 的劣势。由于灵敏度和特异度两个参数存在相互制约的关系,因此不难理解为何融合模型的特异度会稍低于其他两种网络模型。特异度与误诊率相关,特异度越高意味着误诊越少,虽然本文方法的特异度略有下降,但相较于灵敏度的提升是完全可以接受的。同时由于癌症对人类的健康和生命有着巨大的威胁,诊断的最主要目的是尽可能多地找出癌症患者,做到早发现早治疗,意义非常重大,因此灵敏度相比特异度是更为关键的指标。虽然 AlexNet 和 GoogLeNet 两种模型在对胃癌的诊断上已经取得了较好的效果,而融合模型相比两者在灵敏度上获得了明显的提高,相应地在特异度上只付出了很小的代价,说明融合模型在对胃癌的诊断上比另外两种模型更具优势。实验还发现 3 种不同模型在非癌症图像上的准确率均明显高于在癌症图像上的准确率,分析认为这一差异是实验数据因手工标注产生的误差造成的,说明实验使用的方法依赖于医学指导,同时对训练数据的精确性有着较高的要求,这将是后续的研究中需要解决的问题。

对测试结果进行整合之后发现,在 3 种模型对 52 例癌症病例和 13 例非癌症病例的诊断中,非癌症病例都得到了正确的诊断,对癌症病例的诊断中,GoogLeNet 和 AlexNet 分别有 3 例癌症病例未被检出,而本文提出的融合网络模型有 2 例未被检出。从诊断结果上看,本文提出的融合网络对病理切片的诊断准确率有所提高,由此可以说明,融合网络模型诊断的准确率高于其他网络,尤其对癌症组织的病理特征更为敏感。同时网络具有良好的泛化能力,说明实验提出的融合网络模型更有效地提取到了胃癌病理特征。

结束语 本文使用深度卷积神经网络实现对胃癌病理图像的自动分类,采用了 AlexNet 和 GoogLeNet 两种在结构上

差异较大的模型。GoogLeNet 在网络层次和结构上都具有更高的复杂度,其参数量相比 AlexNet 模型更少,但它在计算资源占用上相比 AlexNet 模型而言有较大劣势。本文针对胃癌病理图像的特点,首先对 GoogLeNet 模型的结构进行了优化,优化后的 GoogLeNet 显著降低了计算成本。进一步将 GoogLeNet 和 AlexNet 模型进行融合,综合两种不同结构的模型学习到的不同特征,形成了一种融合的网络模型。改进后的模型同时具有两种网络结构的特点,且更加针对胃癌病理切片,提高了胃癌病理切片识别的灵敏度,因此对胃癌的诊断具有积极的意义。可以看到,模型取得了很高的诊断准确率,其灵敏度达到了 97.60%,特异度更是达到了 99.49%。但实验使用的方法比较依赖人工选取的胃癌病理特征,因此人工筛选中出现的失误不可避免地会对模型训练的结果造成影响,在训练数据精度很低的情况下,训练出的深度学习模型可能无法达到实际的要求。如何做到使算法不需要依靠人工选取的特征便能够对胃癌病理图像进行正确分类,将是下一步的研究方向。

参考文献

- [1] PARKIN D M, BRAY M F, FERLAY M J, et al. Global Cancer Statistics, 2002 [J]. *Ca A Cancer Journal for Clinicians*, 2015, 55(2): 74.
- [2] FERLAY J, SOERJOMATARAM I, DIKSHIT R, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012 [J]. *International Journal of Cancer*, 2015, 136(5): E359-E386.
- [3] 左婷婷, 郑荣寿, 曾红梅, 等. 中国胃癌流行病学现状 [J]. *中国肿瘤临床*, 2017, 44(1): 52-58.
- [4] 孟焱, 张丹, 李燕东, 等. 特殊类型乳腺恶性肿瘤的超声与病理图像分析 [J]. *中国医学影像学杂志*, 2015(3): 188-191.
- [5] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. *Nature*, 2015, 521(7553): 436.
- [6] GURCAN M N, BOUCHERON L E, CAN A, et al. Histopathological Image Analysis: A Review [J]. *IEEE Reviews in Biomedical Engineering*, 2009, 2: 147.
- [7] CHEN H, DOU Q, WANG X, et al. Mitosis detection in breast cancer histology images via deep cascaded networks [C] // Thirtieth AAAI Conference on Artificial Intelligence. AAAI Press, 2016: 1160-1166.
- [8] KWAK J T, HEWITT S M. Multiview boosting digital pathology analysis of prostate cancer [J]. *Comput Methods Programs Biomed*, 2017, 142: 91-99.
- [9] SARRAF S, TOFIGHI G. Deep learning-based pipeline to recognize Alzheimer's disease using fMRI data [C] // Future Technologies Conference. IEEE, 2017.
- [10] 张巧丽, 赵地, 迟学斌. 基于深度学习的医学影像诊断综述 [J]. *计算机科学*, 2017, 44(S2): 1-7.
- [11] TIRUMALA S 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. Springer, Cham, 2016: 26-34.
- [12] 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. Springer, Cham, 2016: 115-123.
- [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] ZHENG Q, MILTHORPE B K, JONES A S. Direct neural network application for automated cell recognition [J]. *Cytometry Part A*, 2004, 57A(1): 1-9.
- [15] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [C] // Computer Vision and Pattern Recognition. IEEE, 2015: 1-9.
- [16] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [J]. *Communications of the Acm*, 2012, 60(2): 2012.
- [17] HINTON G E, OSINDERO S, TEH Y W. A Fast Learning Algorithm for Deep Belief Nets [J]. *Neural Computation*, 2006, 18(7): 1527.
- [18] BENGIO Y, DELALLEAU O. On the expressive power of deep architectures [C] // International Conference on Algorithmic Learning Theory. Springer-Verlag, 2011: 18-36.
- [19] 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. Springer, Cham, 2016: 115-123.
- [20] ERHAN D, BENGIO Y, COURVILLE A, et al. Why Does Unsupervised Pre-training Help Deep Learning? [J]. *Journal of Machine Learning Research*, 2010, 11(3): 625-660.
- [21] 陈小燕. 机器学习算法在数据挖掘中的应用 [J]. *现代电子技术*, 2015, 45(20): 11-14.
- [22] SCHÖLKOPF B, PLATT J, HOFMANN T. Greedy Layer-Wise Training of Deep Networks [C] // International Conference on Neural Information Processing Systems. MIT Press, 2007: 153-160.
- [23] HINTON G E, DAYAN P, FREY B J, et al. The "wake-sleep" algorithm for unsupervised neural networks [J]. *Science*, 1995, 268(5214): 1158.
- [24] KANDEL E R. An introduction to the work of David Hubel and Torsten Wiesel [J]. *Journal of Physiology*, 2009, 587(12): 2733-2741.
- [25] FUKUSHIMA K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position [J]. *Biological Cybernetics*, 1980, 36(4): 193-202.
- [26] GAO L, CHEN P Y, YU S. Demonstration of Convolution Kernel Operation on Resistive Cross-Point Array [J]. *IEEE Electron Device Letters*, 2016, 37(7): 870-873.
- [27] 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.
- [28] 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.
- [29] LIU W, WEN Y, YU Z, et al. Large-margin softmax loss for convolutional neural networks [C] // International Conference on International Conference on Machine Learning. JMLR., 2016: 507-516.
- [30] WRIGHT J, GANESH A, ZHOU Z, et al. Demo: Robust face recognition via sparse representation [C] // IEEE International Conference on Automatic Face & Gesture Recognition. IEEE, 2009: 1-2.
- [31] LIN M, CHEN Q, YAN S. Network In Network [J]. *Computer Science*, arXiv: 1312. 4400.
- [32] JIA Y, SHELHAMER E, DONAHUE J, et al. Caffe: Convolutional Architecture for Fast Feature Embedding [C] // Acm International Conference on Multimedia. ACM, 2014: 675-678.