

DOI:10.3969/j.issn.1005-202X.2022.06.013

医学影像物理

基于改进阈值的VGG网络的新冠肺炎CT图像自动诊断算法

翁羽洁,李忠贤,姬宇程,薄素玲,梁莹
内蒙古医科大学计算机信息学院,内蒙古呼和浩特 010110

【摘要】肺部CT能够较准确地鉴定新冠肺炎病例,但医生工作量较大,本研究提出一种基于改进阈值的VGG网络的新冠肺炎CT图像自动诊断算法,通过该模型可快速准确地完成新冠肺炎病例的自动识别,为进一步控制其传播提供帮助。通过比较卷积神经网络VGG中的VGG-11、VGG-13、VGG-16,获得准确率较高的新冠肺炎CT图像自动诊断模型VGG-13,并在此基础上通过改进阈值的方式使准确率由86%提高到了89%,进一步提高诊断的准确性。

【关键词】新冠肺炎;VGG;CT图像;卷积神经网络;自动诊断

【中图分类号】R318;R563.1

【文献标志码】A

【文章编号】1005-202X(2022)06-0731-06

Automatic diagnosis algorithm for COVID-19 CT images using improved threshold-based VGG network

WENG Yujie, LI Zhongxian, JI Yucheng, BO Suling, LIANG Ying

College of Computer and Information, Inner Mongolia Medical University, Hohhot 010110, China

Abstract: Lung CT can accurately identify COVID-19, but the workload of doctors is relatively large. An automatic diagnosis algorithm for COVID-19 CT image using improved threshold-based VGG network is proposed. The model can quickly and accurately complete the automatic identification of COVID-19 cases, and provide help for further control of its spread. By comparing VGG-11, VGG-13 and VGG-16 in convolutional neural network VGG, the automatic diagnosis model for COVID-19 CT image with high accuracy is obtained. On this basis, the accuracy is enhanced from 86% to 89% by modifying the threshold, further improving the accuracy of diagnosis.

Keywords: COVID-19; VGG; CT image; convolutional neural network; automatic diagnosis

前言

新冠肺炎潜伏期长、传染率高且传播速度快^[1-2]。核酸检测是新冠肺炎检测的主要手段,但依然有一定的局限性,而肺部CT检查对病毒性肺炎具有较高的检出率及准确性,可作为新冠肺炎早期的重要检查手段^[3-4]。

人工智能是研究可以模拟、延伸和扩展人的智能的一门科学^[5-6]。人工智能与医学的结合,尤其与医学影像的结合是现阶段最具有发展前景的领域之一,其中自动诊断的发展尤其迅速^[7-8]。目前,基于人

工智能的自动诊断技术已经应用到很多医疗领域,例如乳腺癌、肺癌的自动诊断等,如果将人工智能技术应用于新冠肺炎的诊断中,可以帮助医生迅速鉴别病情,及时进行隔离和阻断,为阻止疫情的传播起到积极作用。

目前国内李欢等^[9]、Liu等^[10]利用计算机辅助分析软件,研究了新冠肺炎的3个进展阶段,而未对分析算法做进一步的研究。冯毅博等^[11]对比VGG16、ResNet18等网络,提出一种更优的DWSDenseNET网络,此网络对均衡化后的肺部X线图像进行训练,使得模型参数明显下降,但分类效果却更加精确。李文哲等^[12]对比高原肺水肿和新冠肺炎的肺部CT图像,找到两者在图像中的差别,并总结出两者在不同进展期的影像特征。国外Belfiore等^[13]通过研究证实了人工智能技术在新冠肺炎诊断中的必要性,但其未进行具体的人工智能诊断算法研究。Ozturk等^[14]使用DarkNet模型作为分类器,实现二分类(新冠肺炎和正常)和多分类(新冠肺炎、正常和肺炎)的

【收稿日期】2021-10-26

【基金项目】内蒙古自治区自然科学基金(2020LH01012);内蒙古医科大学青年创新基金(YKD2018QNCX014)

【作者简介】翁羽洁,讲师,硕士,研究方向:医学图像处理、人工智能, E-mail: wengyujie@immu.edu.cn

【通信作者】李忠贤,教授,硕士,研究方向:图像处理、生物力学, E-mail: lizhongxian@immu.edu.cn

准确诊断,且可提供云计算服务。

由于牛津大学“视觉几何实验室”(Visual Geometry Group)提出的VGG网络结构简单,可提供更加丰富的图像特征,而且VGG网络有多个分支,由简到繁,可适用于不同类型的图像训练,因此本研究使用VGG对新冠肺炎的CT图像进行训练,并选取其中的3个分支:VGG-11、VGG-13和VGG-16分别进行训练,比较其在诊断新冠肺炎方面的准确度、精确度等各项指标,选出最优的网络模型。在最优的网络模型的基础上,利用受试者工作特征曲线(Receiver Operating Characteristic Curve, ROC)对阈值进行优化,得出诊断新冠肺炎的最佳模型。

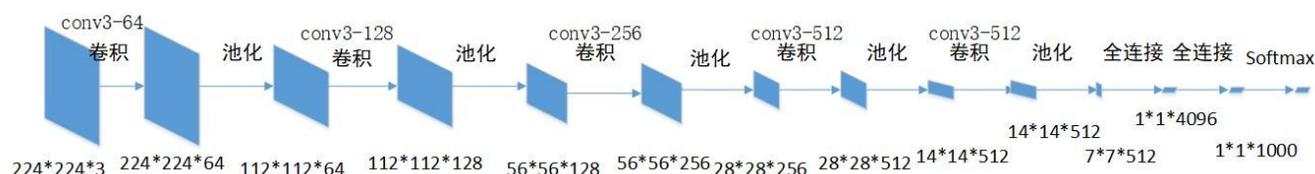


图1 VGG网络算法流程图

Figure 1 VGG network algorithm flowchart

1.1 卷积层

VGG网络使用的是3×3卷积核,使用小卷积层的问题是感受野会变小,VGG通过连续使用多个卷积层的方式来增加感受野,例如VGG-13每一层都包括2个3*3的卷积核,而VGG-16的第3~5层每一层都包括3个3*3的卷积核。VGG网络中的通道数相对较多,更多的通道代表图像更加丰富的特征,VGG网络中第1层的通道数量就达到64个,并且逐层翻倍,最后一层达到了512个,能够获取图像更多的特征^[18]。另外,VGG网络在图像的边缘填充1个像素,这样卷积后保持图像的分辨率不变。

1.2 池化层

池化层是对数据进行降维,用于缓解卷积层对位置的过度敏感性。VGG网络使用2*2的小池化窗口,池化窗口从输入数组的最左上方开始,按从左往右、从上往下且步长为2的顺序滑动,取2*2窗口上的最大值作为降维后图像的值^[19]。

1.3 全连接层

全连接层用于与最后的输出相连接,可以设计为一层或者多层。VGG网络包括3个全连接层,前两层是4096个通道,第3层是输出层,输出1000个分类,根据最后分类数目的不同需要对全连接层进行修改^[20]。

VGG网络所有的结构都按照如上的结构设计,而VGG-11、VGG-13和VGG-16的不同在于卷积层数的不同,也就是深度的不同。VGG-11有11个加权

1 VGG网络结构

VGG是牛津大学计算机视觉实验室参加2014年ImageNet Large Scale Visual Recognition Challenge (ILSVRC)比赛的网络结构,获了图像分类赛的第二名以及定位挑战赛的冠军^[15]。VGG的特点是结构简洁,使用的是3*3卷积核(conv3),并且使用连续多个卷积层来增大感受野,通道数更多,能够得到更加丰富的图像特征^[16-17]。卷积神经网络的深度增加和小卷积核的使用对网络的最终分类识别效果有很大的作用,VGG网络算法流程图见图1。

层,其中包括8个卷积层和3个全连接层;VGG-13有13个加权层,其中包括10个卷积层和3个全连接层;VGG-16包括了13个卷积层和3个全连接层,网络的深度逐渐加深,如表1所示。VGG网络的优点很明显,但是也有不足,一个是网络架构权重数量比较大,很消耗磁盘空间,另一个是训练比较慢。

2 实验方案

本实验利用现有的VGG网络VGG-11、VGG-13和VGG-16对新冠肺炎患者的CT图像和正常肺部CT图像进行分类,并比较3种网络对图像分类的效果。

2.1 实验平台

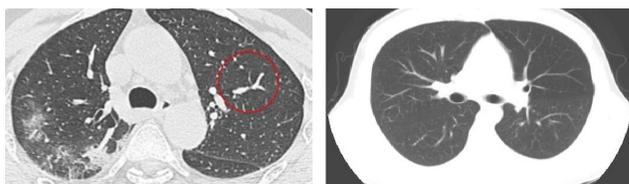
本实验是基于Python和Tensorflow框架实现的,采用显存为8GB的NVIDIA GeForce GTX 1050 Ti和主频2.50 GHz的Inter(R)Core(TM)i5-7300HQ CPU。

2.2 实验数据

本文所用数据源自Zhao等^[21]收集的数据集COVID-CT-Dataset: a CT scan dataset about COVID-19。数据集包括746张肺部CT图像,其中新冠肺炎CT图像349张,非新冠肺炎图像397张。本研究中使用分层抽样抽取了100张图像作为测试数据;剩余646张图像作为训练数据,并按照5:1的比例继续划分训练集和验证集。训练参数中学习率为0.00001,最大迭代次数为20,并对图像做均值预处理。数据集中样例图片见图2。

表1 VGG网络结构
Table 1 VGG network structure

VGG	VGG-11	VGG-13	VGG-16
卷积第一层	conv3-64	conv3-64	conv3-64
	-	conv3-64	conv3-64
池化第一层	maxpool	maxpool	maxpool
卷积第二层	conv3-128	conv3-128	conv3-128
	-	conv3-128	conv3-128
池化第二层	maxpool	maxpool	maxpool
卷积第三层	conv3-256	conv3-256	conv3-256
	conv3-256	conv3-256	conv3-256
	-	-	conv3-256
池化第三层	maxpool	maxpool	maxpool
卷积第四层	conv3-512	conv3-512	conv3-512
	conv3-512	conv3-512	conv3-512
	-	-	conv3-512
池化第四层	maxpool	maxpool	maxpool
卷积第五层	conv3-512	conv3-512	conv3-512
	conv3-512	conv3-512	conv3-512
	-	-	conv3-512
池化第五层	maxpool	maxpool	maxpool
全连接第一层	FC4096	FC4096	FC4096
全连接第二层	FC4096	FC4096	FC4096
输出层	FC1000	FC1000	FC1000



a: 新冠肺炎CT图像

b: 非新冠肺炎CT图像

图2 数据集中肺部CT图像

Figure 2 Lung CT image from data set

3 实验结果

本实验使用VGG的3个模型对数据进行训练,目的是对比这3个模型在新冠肺炎自动诊断中的优劣,找出最佳的模型,本研究分别从训练集和测试集两个方面来进行比对。

3.1 训练集结果

训练中使用计算softmax交叉熵来得出分类的概率误差^[22],该方法可以很好地避免梯度弥散,而且计算简单高效,主要包括以下两个步骤。

(1)对输出层做softmax,得出输出层属于其中某一类的概率,softmax公式如下:

$$\text{softmax}(x)_i = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \quad (1)$$

其中, x_i 是全连接网络的第*i*类输出, x_j 是第*j*类输出。

(2)对上一步的输出与样本的真实标签做交叉熵,公式如下:

$$H_y(y) = -\sum_i y_i \log(y_i) \quad (2)$$

其中, y_i 是第一步的输出结果 $\text{softmax}(x)_i$ 。

在训练集上的实验结果如图3~图5所示,分别从训练集准确率、训练集损失值、验证集准确率和验证集损失值这4个参数比较3个模型。结果发现,这3个模型的训练集准确率最终都可以达到100%,但相比较而言,VGG-13模型较早达到了较高的准确率且相对稳定,而VGG-16在训练后期才达到较高的准确率,且不稳定。在验证集上,VGG-11的表现较好,VGG-13次之,而VGG-16相对较差。在损失值的变化曲线中,3个模型在开始阶段迅速下降,后期趋于稳定,其中VGG-11的损失曲线相对比较理想,下降速度和稳定度都优于其他两个模型。

3.2 测试集结果

测试集数据为100张图像,采用分层抽取的方法在总的数据集中抽取得到。将3个训练好的模型分别在测试集上进行验证,并分别比对精确率、召回率和准确率,精确率、召回率和准确率分别通过混淆矩阵^[23](表2)得出。

精确率如式(3)所示:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

召回率如式(4)所示:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

准确率如式(5)所示:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (5)$$

通过表3可以发现,VGG-13模型在测试集中的表现是最好的,精确率、召回率和准确率分别都是最高,而VGG-11模型次之,而VGG-16模型最差。

4 模型改进

从上面的结论可以看出VGG-13模型在本实验中的效果是最好的,不论是训练集还是测试集,但是在测试集验证的时候发现,许多分类错误的结果的分类概率值在0.5附近,也就是用默认的0.5阈值进行分类时很多图片都被分类错误,0.5并不是最佳阈值。因此,对3个模型分别画出其ROC曲线^[24],通过曲线可以看到,VGG-13模型的ROC曲线下的面积(Area Under Curve, AUC)最大为0.9464,如图6所示。从而进一步证明在新冠肺炎的自动诊断中,VGG-13模型的效果最好,而且通过分析ROC曲线,

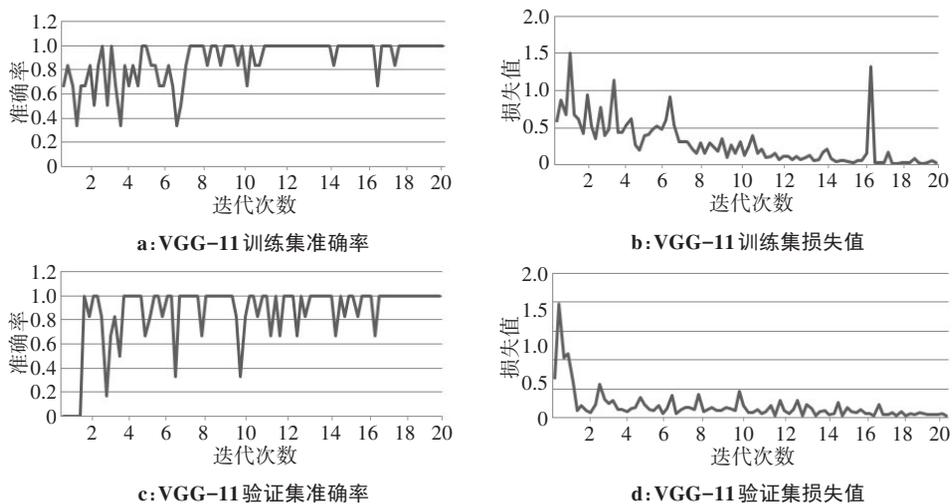


图3 VGG-11 训练集和验证集的准确率和损失值

Figure 3 Accuracy and loss of VGG-11 in training set and validation set

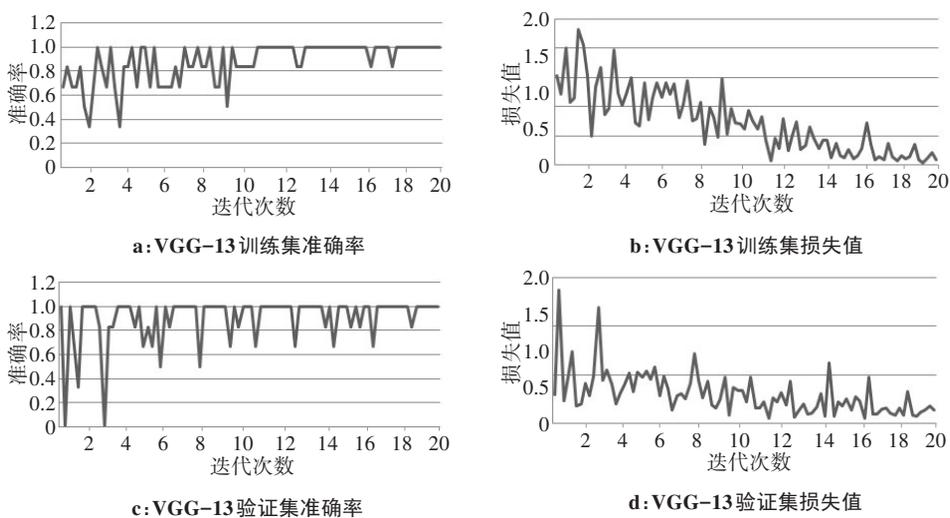


图4 VGG-13 训练集和验证集的准确率和损失值

Figure 4 Accuracy and loss of VGG-13 in training set and validation set

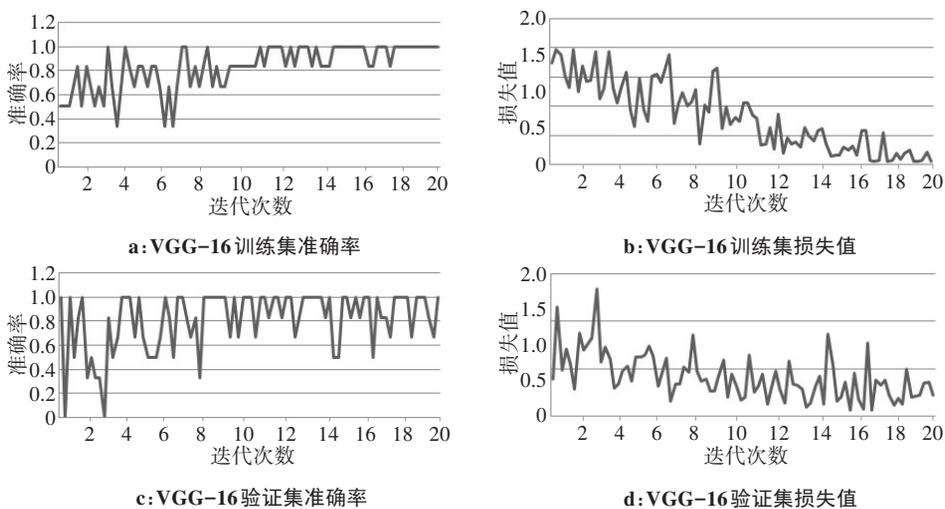


图5 VGG-16 训练集和验证集的准确率和损失值

Figure 5 Accuracy and loss of VGG-16 in training set and validation set

表2 混淆矩阵

Table 2 Confusion matrix

	预测值(阳性)	预测值(阴性)
真实值(阳性)	TP(真阳性)	FN(假阴性)
真实值(阴性)	FP(假阳性)	TN(真阴性)

表3 测试集结果(%)

Table 3 Performances of 3 models in test set (%)

模型	精确率	召回率	准确率
VGG-11	82.00	86.00	84.00
VGG-13	83.33	90.00	86.00
VGG-16	79.63	86.00	82.00

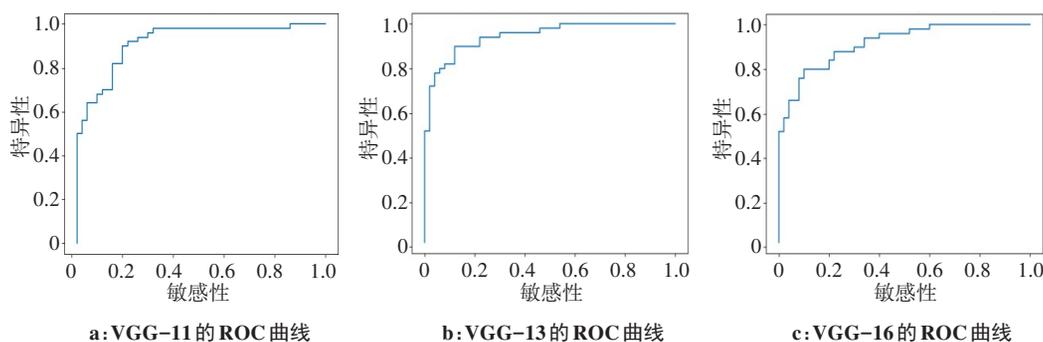


图6 VGG-11、VGG-13、VGG-16的ROC曲线

Figure 6 ROC curve of VGG-11, VGG-13, VGG-16

如表4所示, VGG-11、VGG-13、VGG-16的准确率分别提高到了86%、89%和85%。

表4 阈值修改后的测试集结果(%)

Table 4 Test set results after threshold modification (%)

模型	特异度	敏感度	准确率
VGG-11	92	80	86
VGG-13	90	88	89
VGG-16	80	90	85

5 讨论

本研究发现在VGG的3个模型中,VGG-11的模型复杂度最低,VGG-16的模型复杂度最高,而VGG-13的模型复杂度介于两者之间,但从新冠肺炎的诊断准确率上看,VGG-13的效果最好,VGG-11次之,而复杂度最高的VGG-16的效果最差。因此,在训练的过程中,要根据实际情况选取合适的网络进行训练,并根据数据集的情况进行参数的调整,才能够得到较好的训练效果。

同时,本研究在VGG-13模型的基础上,发现阈

值越靠近左上角,实验的准确性就越高。最靠近左上角的ROC曲线的点是最佳阈值,其假阳性和假阴性的总数最少。

最佳阈值通过约登指数(Youden Index, YI)^[25]进行选择,YI也称正确指数,是指召回率和特异度之和减去1:

$$YI = Recall + Specificity - 1 \tag{6}$$

其中,召回率见公式(4),特异度计算公式如下:

$$Specificity = \frac{TN}{TN + FP} \tag{7}$$

得出在VGG-13中的最佳阈值为0.571,VGG-11的最佳阈值为0.447,VGG-16的最佳阈值为0.700,那么通过该阈值的设定可以将该疾病的准确率提高,

值的选取对结果的影响较大,对于默认的0.5的阈值并不能进行很好的分类,因此通过ROC曲线和YI得出了最佳阈值为0.571,进一步提高了准确率。虽然本研究在准确度上还需要进一步的提高,但此模型结构简单,运算速度快,可快速对CT图像进行分类,而且为后续的改进工作奠定了一定的基础。

【参考文献】

- [1] Naserghandi A, Allameh SF, Saffarpour R. All about COVID-19 in brief[J]. New Microbes New Infect, 2020, 35: 100678.
- [2] Pascarella G, Strumia A, Piliengo C, et al. COVID-19 diagnosis and management: a comprehensive review[J]. J Intern Med, 2020, 288(2): 192-206.
- [3] Tenda ED, Yulianti M, Asaf MM, et al. The importance of chest CT scan in COVID-19[J]. Acta Med Indones, 2020, 52(1): 68-73.
- [4] Li MZ, Lei PG, Zeng BL, et al. Coronavirus disease (COVID-19): spectrum of CT findings and temporal progression of the disease[J]. Acad Radiol, 2020, 27(5): 603-608.
- [5] Lawrence DR, González CP, Harris J. Artificial intelligence[J]. Camb Q Healthc Ethic, 2016, 25(2): 250-261.
- [6] Hamet P, Tremblay J. Artificial intelligence in medicine [J]. Metabolism, 2017, 69S: S36-S40.
- [7] Mintz Y, Brodie R. Introduction to artificial intelligence in medicine [J]. Minim Invasive Ther Allied Technol, 2019, 28(2): 73-81.
- [8] Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare[J]. Nat Biomed Eng, 2018, 2(10): 719-731.
- [9] 李欢,程尼涛,孙文博,等.人工智能辅助定量分析新型冠状病毒肺炎的CT进展类型[J].武汉大学学报(医学版),2021,42(1): 1-5.
- Li H, Cheng NT, Sun BW, et al. AI-assisted quantitative analysis of

- chest CT progression patterns of COVID-19[J]. Medical Journal of Wuhan University, 2021, 42(1): 1-5.
- [10] Liu FJ, Zhang Q, Huang C, et al. CT quantification of pneumonia lesions in early days predicts progression to severe illness in a cohort of COVID-19 patients[J]. Theranostics, 2020, 10(12): 5613-5622.
- [11] 冯毅博, 仇大伟, 曹慧, 等. 基于深度可分离稠密网络的新型冠状病毒肺炎X线图像检测方法研究[J]. 生物医学工程学杂志, 2020, 37(4): 557-565.
Feng YB, Qiu DW, Cao H, et al. Research on coronavirus disease 2019 (COVID-19) detection method based on depthwise separable DenseNet in chest X-ray images[J]. Journal of Biomedical Engineering, 2020, 37(4): 557-565.
- [12] 李文哲, 李凯, 张楠, 等. 高原肺水肿与新型冠状病毒肺炎计算机断层扫描特征及影像鉴别[J]. 生物医学工程学杂志, 2020, 37(6): 1031-1036.
Li WZ, Li K, Zhang N, et al. Differential diagnosis of high altitude pulmonary edema and COVID-19 with computed tomography feature [J]. Journal of Biomedical Engineering, 2020, 37(6): 1031-1036.
- [13] Belfiore MP, Urraro F, Grassi R, et al. Artificial intelligence to codify lung CT in Covid-19 patients[J]. Radiol Med, 2020, 125(5): 500-504.
- [14] Ozturk T, Talo M, Yildirim EA, et al. Automated detection of COVID-19 cases using deep neural networks with X-ray images[J]. Comput Biol Med, 2020, 121: 103792.
- [15] Sengupta A, Ye YT, Wang R, et al. Going deeper in spiking neural networks: VGG and residual architectures[J]. Front Neurosci, 2019, 13: 95.
- [16] Lei G, Zhang SQ, Jun T, et al. Lung segmentation method with dilated convolution based on VGG-16 network [J]. Comput Assist Surg (Abingdon), 2019, 24 (Suppl 2): 27-33.
- [17] 王羽微, 程远, 毕海, 等. 基于改进式VGG网络的海洋单细胞藻类识别算法研究[J]. 大连海洋大学学报, 2020, 161. DOI: 10.16535/j.cnki.dlhyxb.2020-161. DOI: 10.16535/j.cnki.dlhyxb.2020-161.
Wang YZ, Cheng Y, Bi H, et al. Research on marine single cell algae recognition algorithm based on improved VGG network[J]. Journal of Dalian Ocean University, 2020, 161. DOI: 10.16535/j.cnki.dlhyxb.2020-161. DOI: 10.16535/j.cnki.dlhyxb.2020-161.
- [18] 王浩雨. 基于改进VGG16神经网络的异常检测模型研究[J]. 现代计算机, 2020(30): 3-7.
Wang HY. Research on anomaly detection model based on improved VGG16 neural network[J]. Modern Computer, 2020(30): 3-7.
- [19] 曹宇, 邢素霞, 逢键梁, 等. 基于改进的VGG-16卷积神经网络的肺结节检测[J]. 中国医学物理学杂志, 2020, 37(7): 940-944.
Cao Y, Xin SX, Pang JL, et al. Detection of pulmonary nodules based on improved VGG-16 convolution neural network[J]. Chinese Journal of Medical Physics, 2020, 37(7): 940-944.
- [20] 邓竹琴, 俞永伟. 改进型深度学习模型在乳腺肿瘤恶性鉴别中的应用[J]. 中国医学物理学杂志, 2020, 37(11): 1469-1473.
Deng ZQ, Yu YW. Application of improved deep learning model in differential diagnosis of benign and malignant breast tumors [J]. Chinese Journal of Medical Physics, 2020, 37(11): 1469-1473.
- [21] Zhao JY, Zhang YC, He XH, et al. COVID-CT-dataset: a CT scan dataset about COVID-19[J]. arXiv preprint, arXiv: 2003.13865, 2020.
- [22] 杨海龙, 田莹, 王澧冰. 基于优化损失函数的YOLOv2目标检测器[J]. 辽宁科技大学学报, 2020, 43(1): 52-57.
Yang HL, Tian Y, Wang LB. Object detector YOLOv2 based on optimized loss function [J]. Journal of University of Science and Technology Liaoning, 2020, 43(1): 52-57.
- [23] 徐健锋, 苗奇谦, 张远健. 基于混淆矩阵的多目标优化三支决策模型[J]. 模式识别与人工智能, 2017, 30(9): 859-864.
Xu JF, Miao DQ, Zhang YJ. Three-way decisions model for multi-object optimization based on confusion matrix [J]. Pattern Recognition and Artificial Intelligence, 2017, 30(9): 859-864.
- [24] Kamarudin AN, Cox T, Kolumunnage-Dona R. Time-dependent ROC curve analysis in medical research: current methods and applications [J]. BMC Med Res Methodol, 2017, 17(1): 53.
- [25] Li CX, Chen JY, Qin GS. Partial Youden index and its inferences [J]. J Biopharm Stat, 2019, 29(2): 385-399.

(编辑:薛泽玲)