

文章编号: 2095-2163(2023)10-0009-06

中图分类号: R735.9

文献标志码: A

基于 CT 图像的胰腺癌智能辅助诊断模型研究

郭冰冰, 谷雪莲, 胡秀枋, 孙运文, 徐秀林

(上海理工大学 健康科学与工程学院, 上海 200093)

摘要: 为了更好的为临床疗效评估提供参考, 本研究使用胰腺癌患者的 APT、DPT、NPT、VPT 4 期 CT 图像数据集进行肿瘤良恶性分类。分别建立深度学习和机器学习模型, 评价对胰腺肿瘤良恶性分类的准确率, 比较了深度学习和机器学习模型在 CT 图像应用性能的优势与不足。深度学习模型采用 VGG16, 训练测试后评估识别率较高的是 APT 数据集和 VPT 数据集, 其测试集准确率均达到了 86% 以上; 机器学习模型采用支持向量机、逻辑回归、随机森林、K 近邻和 AdaBoost 5 种模型, 随机森林和 AdaBoost 两种模型效果良好, 训练测试后识别率较高的是 NPT 数据集和 VPT 数据集。在使用 AdaBoost 模型时, VPT 时期的准确率达到 77%。综合表现说明, 胰腺癌在 VPT 时期利用计算机辅助诊断技术具有良好的应用效果。

关键词: CNN; 机器学习; 胰腺癌; CT 图像; AdaBoost; 随机森林

Research on intelligent assistant model for pancreatic cancer diagnosis based on CT image

GUO Bingbing, GU Xuelian, HU Xiufang, SUN Yunwen, XU Xiulin

(School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

[Abstract] In order to provide a better reference for the evaluation of clinical efficacy, the four-phase APT, DPT, NPT and VPT CT image datasets of pancreatic cancer patients are used to classify benign and malignant tumors. In this paper, deep learning and machine learning models are established respectively to evaluate the accuracy of benign and malignant pancreatic tumors classification, and the advantages and disadvantages of deep learning and machine learning models in CT image application are analyzed. The deep learning model adopts VGG16, and the APT dataset and VPT dataset have higher recognition rates after training, where the accuracy of the test set is above 86%. Support vector machine, logistic regression, random Forest, K-nearest neighbor and AdaBoost are used in the machine learning model. In these machine learning methods, random forest and AdaBoost have good effects, and NPT dataset and VPT dataset have high recognition rates. When AdaBoost model is used, the accuracy rate in VPT period reaches 77%. The comprehensive experiments show that the application of computer-aided diagnosis in VPT stage of pancreatic cancer has good effect.

[Key words] CNN; machine learning; pancreatic cancer; CT images; AdaBoost; random forest

0 引言

胰腺癌是预后最差的恶性肿瘤之一, 其具有侵袭性强、隐匿度高、病程短的特点。据全球癌症统计数据显示, 胰腺癌在癌症的死亡原因里位居前列, 且有研究预测未来 10 年内胰腺癌相关死亡人数将上升至所有癌种的第 2 位, 仅次于肺癌^[1-2]。胰腺癌 IV 期患者确诊后生存期小于 1 年, 5 年平均生存率小于 1%^[3]。同时, 由于大部分胰腺癌患者在被确诊时就已错过最佳治疗时期, 因此寻找到一种有助于早期诊断的方法显得尤为重要, 可大幅提升患者的生存时间^[4]。目前, 病理切片依然被视为确诊的

“金标准”, 但总体上依赖的是医生的经验, 不同的医生可能生成不同的诊断结果。随着人工智能在医学领域的深入应用, 计算机辅助诊断 (Computer-Aided Diagnosis, CAD) 技术的优势不断显现出来, 其可以有效减少医生的阅片工作量, 也可为医生提供诊断决策支持, 有助于医生减少临床实践中的误诊率和误治率^[5], 甚至能预测病灶迁移位置, 以便为临床疗效评估提供参考。

机器学习是人工智能的重要研究内容之一, 机器学习算法是一类从数据中自动分析获得规律, 并利用规律对未知数据进行预测的算法。针对不同的问题需要不同的学习方式, 主要的学习方式有有监

作者简介: 郭冰冰 (1998-), 女, 硕士研究生, 主要研究方向: 人工智能及大数据; 徐秀林 (1957-), 女, 硕士, 教授、硕士生导师, 主要研究方向: 康复设备及其关键技术、人工智能及大数据。

通讯作者: 徐秀林 Email: xxlin100@163.com

收稿日期: 2022-11-05

哈尔滨工业大学主办 ◆ 学术研究与应用

督学习和无监督学习。有监督学习常用于分类和回归的问题,其特点是所有数据都有标签;无监督学习常用于聚类的问题,其特点为所有数据都没有标签。常用的机器学习算法有支持向量机、逻辑回归、随机森林、Lasso 回归、Ada Boost (adaptive boosting, AdaBoost)等。近年来,由于医疗数据不断增长,因此机器学习的优势得到充分的体现。

2006年,Hinton教授及其团队提出了一种新的网络结构——深度置信网络(Deep Belief Net, DBN)^[6],至此深度学习开始出现在研究者视野里。深度学习常被应用于图像识别,但随着技术的不断推进,在医疗诊断、医用机器人控制、搜索引擎等领域均有不俗的表现。目前,最经典的深度学习网络还是卷积神经网络(Convolutional Neural Network, CNN),Lecun^[7-8]等在神经网络中创造性的代入了传播算法,并以此为基础构建了卷积神经网络。此外,其它常用的学习模型有 Inception^[9]、ResNet^[10]、AlexNet、VGG等。早期诊断是改善胰腺癌患者预后的关键因素之一,若能提升早期诊断率,就能极大改善胰腺癌的预后并延长生存时间。本文主要研究CT分层图像对胰腺癌的早期良恶性诊断的有效性,通过构建 AdaBoost等5种机器学习模型和经典深度学习网络模型 VGG16,分别获得胰腺癌良恶性分类的准确率,同时对不同模型进行对比分析,进一步验证CT图像在VGG16的应用性能。

1 资料与方法

1.1 数据集预处理

本研究使用的数据来自医院的胰腺癌病人CT图像,将其分成APT、DPT、NPT、VPT4期的扫描图像,并对图像进行去噪处理,去除相应体数据的绝对位置、数据密度等信息,4个数据集存在交集。其中,APT(静脉期)包含786张图像;DPT(延迟期)包含757张图像;VPT(动脉期)包含1382张图像;NPT(平扫期)包含1368张图像。

在对数据集进行训练时,将数据集分成训练集、验证集、测试集3部分。训练集用于训练模型,验证集用来验证模型的效能,测试集用来对模型进行评估。实践中,测试集和验证集采用同一批数据,在分配时将所有数据的80%分配给训练集,20%分配给测试集。

1.2 模型训练方法

VGGNet^[11]是2014年ImageNet Challenge图像识别比赛的亚军,其不仅在图像识别应用非常广泛,

在目标分割、人脸识别等方面的应用也会使用VGGNet作为基础模型,其证明了增加网络的深度能够在一定程度上影响网络最终的性能。VGG16有16层网络,能不断提取学习从低级到高级的特征,卷积核大小为3×3,逐层增加卷积核的数量以增强非线性表达能力^[12]。本文深度学习模型选用VGG16对已有数据集进行训练学习。

AdaBoost^[13]是机器学习工具箱中最强有力的工具之一,其基分类器一般是单层决策树,决策树在分裂的时候会选择最优属性进行分裂,最优属性为误差最小的属性,AdaBoost在每次迭代时会选择误差最小的那棵树。其优势在于Adaboost提供了一种能使用其他方法构建子分类器的框架,不仅能显著提高学习精度,且可以根据子分类器的反馈,对假定的错误率进行自适应调整。

1.3 模型评价指标

在对模型进行学习训练时,本文采用的评价指标为模型的训练集和测试集的准确率(Accuracy)及其损失函数(Loss Function)。

准确率的计算公式为

$$c = \frac{\text{分类正确的样本量}}{\text{总输入样本量}} \quad (1)$$

本文选择的损失函数为交叉熵损失函数(categorical_crossentropy),可评估当前训练得到的概率分布与真实分布的差异情况。交叉熵的值越小,概率分布与真实分布就越接近。其具体公式为:

$$C = -\frac{1}{n} \sum_x [y \ln a + (1 - y) \ln(1 - a)] \quad (2)$$

$$a = \sigma(z) \quad (3)$$

$$z = \sum W_j * X_j + b \quad (4)$$

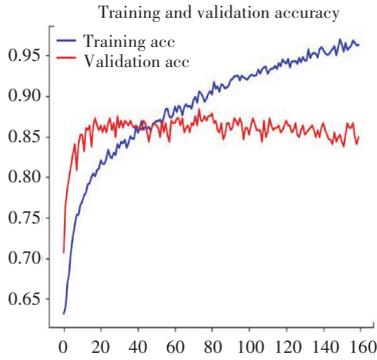
其中, y 为期望的输出, a 为神经元的实际输出。

2 基于深度学习良恶性分类诊断模型

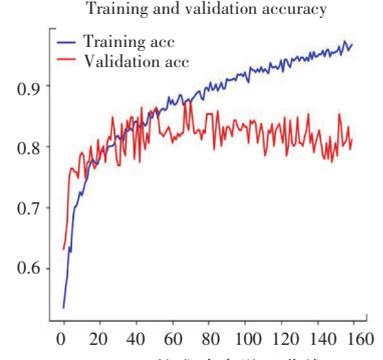
首先基于Keras搭建VGG16模型,其中最后一个全连接层的激活函数为softmax,其优点是计算简单且计算量较小,能把数据的值映射到0~1区间,保证概率为非负值;其次将输入的数据集图片调整尺寸大小为(224,224)的单通道灰度图像,以加快模型训练速度。本实验目标是识别胰腺癌肿瘤的良恶性,属于二分类问题,设置训练次数为160次,优化函数为Adam^[14],是一种对随机梯度下降法的扩展,其优点在于效率高、内存小、且超参数具有直观的解释,通常只需要较少的调谐。

本文分别从 APT、DPT、VPT、NPT 的数据集中随机提取 20% 的图像作为验证集, 将 4 组图像数据分别导入搭建好的 VGG16 模型进行训练, 经过不断

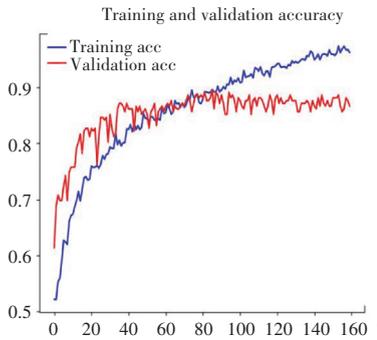
对参数调优, 最终训练集和验证集的训练结果如图 1 所示, 其对应训练集和验证集的损失函数图像如图 2 所示:



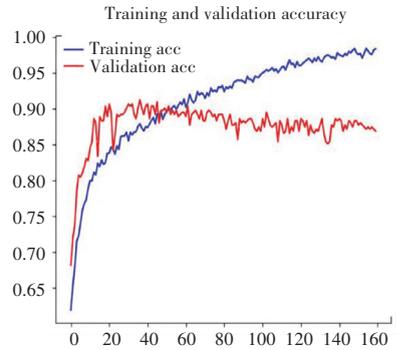
(a) APT 的准确率学习曲线



(b) DPT 的准确率学习曲线



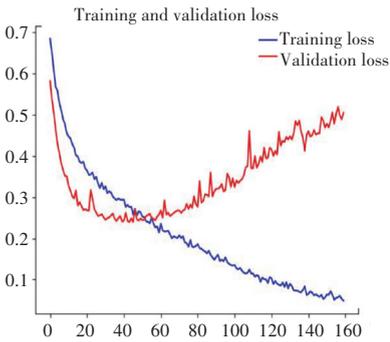
(c) NPT 的准确率学习曲线



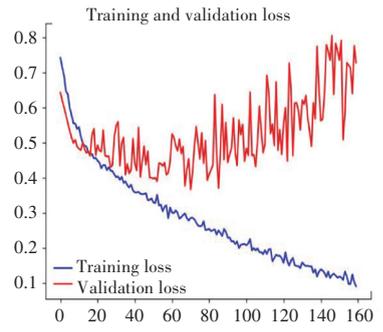
(d) VPT 的准确率学习曲线

图 1 4 组数据集的准确率学习曲线

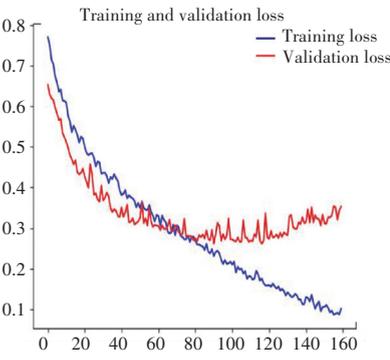
Fig. 1 Learning curve of accuracy of four datasets



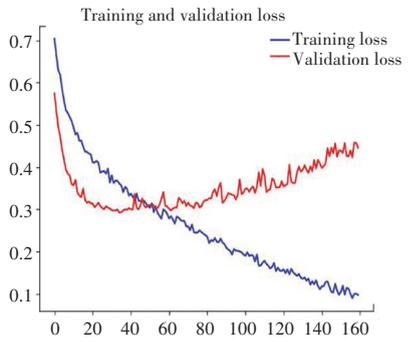
(a) APT 的损失函数曲线



(b) DPT 的损失函数曲线



(c) NPT 的损失函数曲线



(d) VPT 的损失函数曲线

图 2 4 组数据集的损失函数曲线

Fig. 2 Loss curves of four datasets

根据以上训练和验证结果,得到如表 1 所示的混淆矩阵。

由表 1 不难看出,VGG16 模型对于 APT、DPT

数据集的灵敏度最好,APT、VPT 数据集的 *AUC* 和精确度较高。

表 1 4 组数据集训练后的混淆矩阵

Tab. 1 Confusion matrix after training on four datasets

Models	<i>AUC</i>	<i>Accuracy</i>	<i>Fpr</i>	<i>Tpr</i>	<i>Sensitivity(Recall)</i>	<i>Loss</i>	<i>Pecison</i>	<i>AP</i>	<i>F1</i>
APT	0.952	0.8663	0.222	0.925	0.925	4.616	0.861	0.969	0.892
DPT	0.920	0.810	0.395	0.963	0.963	6.544	0.766	0.938	0.853
NPT	0.928	0.849	0.093	0.741	0.741	5.205	0.809	0.877	0.773
VPT	0.940	0.868	0.099	0.808	0.808	4.544	0.815	0.898	0.811

为了进一步验证模型的准确率,本研究分别将上述 4 期的测试集导入已训练好的 VGG16 模型,对患者进行肿瘤良恶性分类诊断。其中包含 202 张 APT 的胰腺癌图像,190 张 DPT 的胰腺癌图像,345 张 NPT 的胰腺癌图像,342 张 VPT 的胰腺癌图像,预测与真实值的对比结果见表 2:

表 2 4 组数据集预测值与真实值结果对比

Tab. 2 Comparison of prediction results and ground truth on four datasets

Confusion matrix	Actual	Prediction		<i>Accuracy</i>
		Benign	Malignant	
APT	Benign	63	18	0.778
	Malignant	9	112	0.926
DPT	Benign	49	32	0.605
	Malignant	4	105	0.963
NPT	Benign	204	21	0.907
	Malignant	31	89	0.742
VPT	Benign	200	22	0.901
	Malignant	23	97	0.808

由表 2 可知,APT、DPT 对胰腺癌恶性肿瘤的识别率高达 0.92 以上,而 NPT、VPT 对良性肿瘤识别率也超过了 0.90。

3 基于机器学习模型的良恶性分类诊断

为了建立识别胰腺癌 CT 图像良恶性分类最佳的机器学习模型,本文将 4 组数据集分别导入支持向量机、逻辑回归、随机森林、K 近邻、AdaBoost 5 个模型进行训练学习,同时从原数据集提取 20% 作为测试集以验证模型训练效果。各模型对比结果见表 3。

由表 3 可知,随机森林和 AdaBoost 训练集和测试集准确率较好,对于 APT、DPT、NPT、VPT 4 个数据集的训练效果均达到 90% 以上,测试集准确率达到 70% 以上。随机森林的学习曲线如图 3 所示,AdaBoost 的学习曲线如图 4 所示。

表 3 5 种机器学习模型对于 4 组数据集的训练结果对比

Tab. 3 Comparison of prediction results of five machine learning models on four datasets

		支持向量机	逻辑回归	随机森林	K 近邻	AdaBoost
APT	训练集准确率	0.949	0.925	0.996	0.700	0.998
	测试集准确率	0.670	0.696	0.741	0.671	0.728
DPT	训练集准确率	0.959	0.953	0.993	0.699	0.995
	测试集准确率	0.730	0.671	0.757	0.638	0.737
NPT	训练集准确率	0.935	0.935	0.953	0.782	0.952
	测试集准确率	0.679	0.704	0.763	0.715	0.777
VPT	训练集准确率	0.940	0.962	0.926	0.795	0.948
	测试集准确率	0.736	0.758	0.751	0.758	0.773

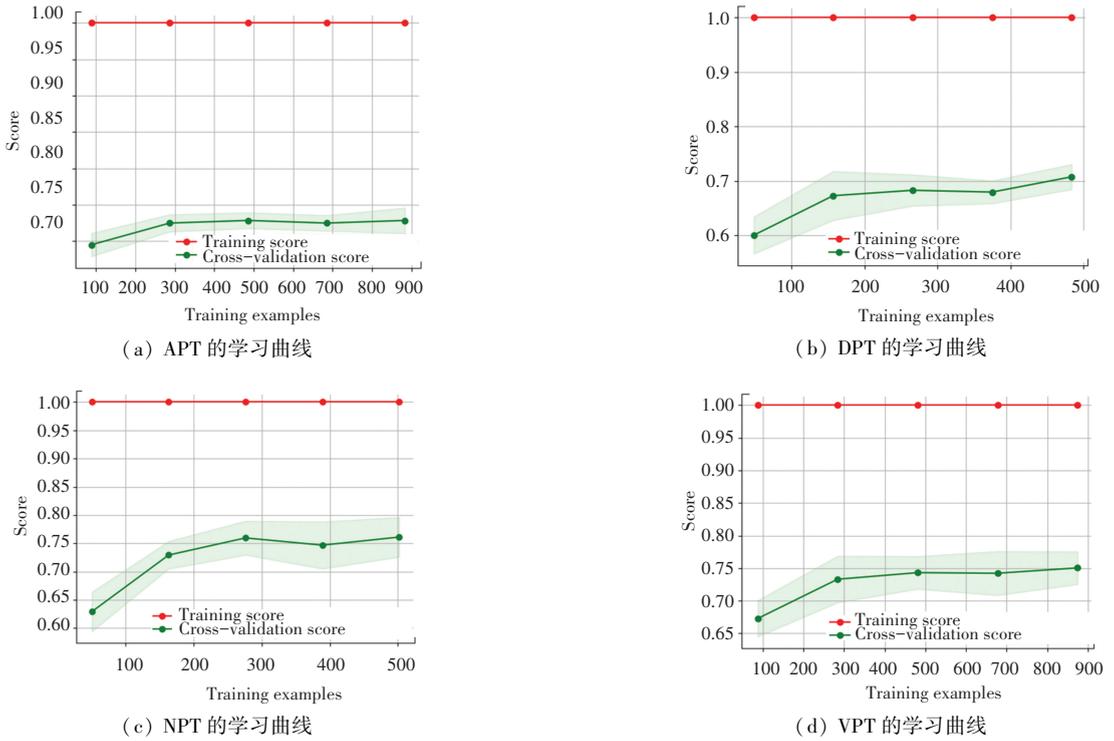


图 3 4 组数据集在随机森林模型下的学习曲线

Fig. 3 Learning curves of random forest model on four datasets

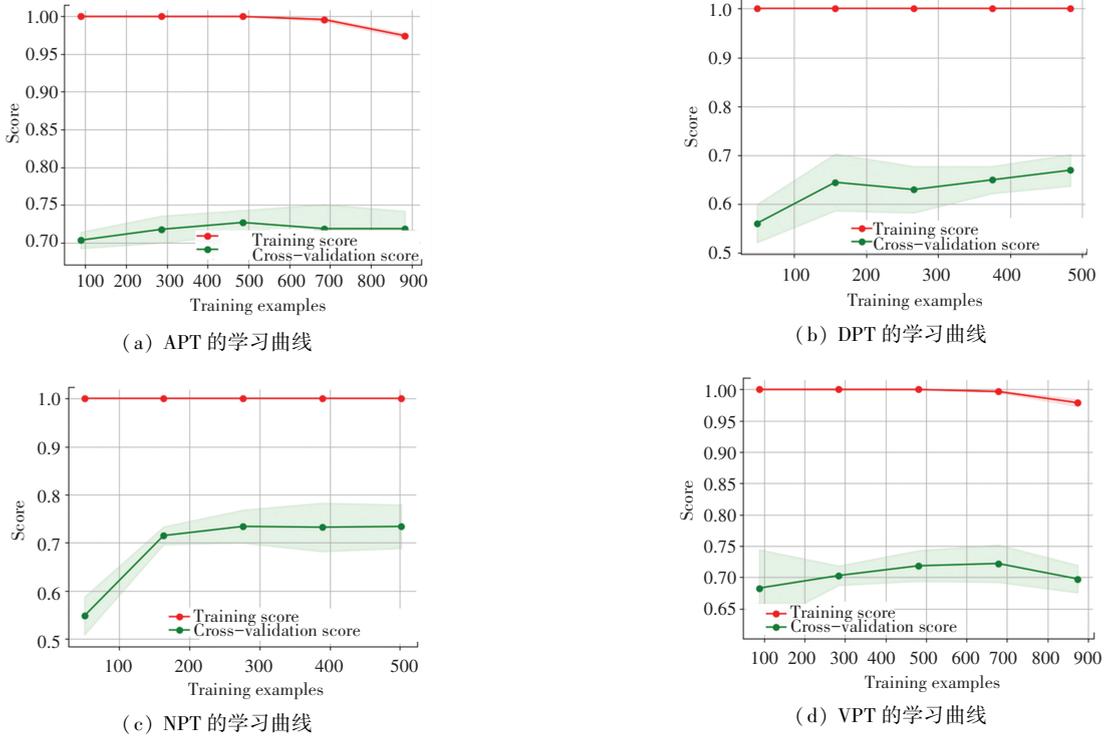


图 4 4 组数据集在 AdaBoost 模型下的学习曲线

Fig. 4 Learning curves of AdaBoost model on four datasets

4 结束语

随着人工智能技术的蓬勃发展,机器学习和深

度学习在医疗辅助诊断领域的应用也在不断深入,基于 CT 图像的胰腺癌良恶性分类的问题已经证明了其具有良好的可行性,获得了很好的分类效

果^[15]。

使用机器学习进行训练学习的优势在于网络较简单,速度快,但对于诊断CT图像良恶性准确率不高;训练集与测试集的差距超过20%,因此可能存在过拟合现象。

深度学习对于图像的识别率很高,CNN对于4期的CT图像的AUC均可以达到90%以上,但其不足在于网络模型较为复杂,训练时间长,对电脑配置要求较高。

本文分别将胰腺癌患者的APT、DPT、NPT、VPT 4期CT图像导入深度学习和机器学习的网络模型训练,虽然不同网络在准确率上有较为明显的差距,但基于深度学习模型的良恶性分类诊断中,综合得分和识别率较高的是APT和VPT时期;基于深度学习模型的良恶性分类诊断中,综合得分和识别率较高的是VPT和NPT时期。

综上,不论使用深度学习还是机器学习的方法,良恶性特征较为明显的时期是VPT,即动脉期。

虽然本文的研究取得了一定成果,但依然存在一些不足。如:4期的CT图像数据虽然存在交集,但并不是完全重合的数据;由于深度学习网络模型耗时较长,只选取了经典的VGG16导入数据训练等。后续可尝试ResNet、Inception等其他网络模型进行进一步研究。

参考文献

- [1] RAHIB L, SMITH B D, AIZENBERG R, et al. Projecting Cancer Incidence and Deaths to 2030: The Unexpected burden of thyroid, liver, and pancreas cancers in the united states [J]. Cancer Research, 2014, 74(14): 2913-2921.
- [2] RAHIB L, WEHNER M, MATRISIAN M C, et al. Projection of cancer incidence and death to 2040 in the US: Impact of cancer screening and a changing demographic [J]. Journal of Clinical

Oncology, 2020, 38(15):1556-1566.

- [3] CHAN-SENG-YUE M, KIM J C, WILSON G W, et al. Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution [J]. Nature Genetics, 2020, 5(2):231-240.
- [4] 何鹏飞. 基于深度学习的多模态影像融合及应用研究[D]. 哈尔滨:哈尔滨工业大学, 2021.
- [5] SS A, FFRA B, MN A, et al. Artificial intelligence in healthcare: opportunities and risk for future [J]. Gaceta Sanitaria, 2021, 35(s1):s67-s70.
- [6] HINTON G E, SALAKHUTDINOV R R. Reducing the Dimensionality of Data with Neural Networks [J]. Science, 2006, 313(5786): 504-507.
- [7] LECUN Y, BENGIO Y, HINTON G. Deep learning [J]. Nature, 2015, 521(7553): 436.
- [8] ZHANG X, ZHAO J, LECUN Y. Character-level convolutional networks for text classification [J]. Advances in Neural Information Processing Systems, 2015(1):649-657.
- [9] 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.
- [10] 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.
- [11] SIMONYAN K, ZISSERMAN A. Very Deep Convolutional Networks for Large-Scale Image Recognition [J]. Computer Science, 2014, 7(1):1409-1418
- [12] 侯向宁,徐草草,杨井荣. 基于Spark的花卉图像分类研究[J]. 计算机技术与发展, 2022, 32(7):5.
- [13] FRIEDMAN J, TIBSHIRANI R, HASTIE T. Additive logistic regression: A statistical view of boosting [J]. The Annals of Statistics: An Official Journal of the Institute of Mathematical Statistics, 2000,28(2):337-374.
- [14] KINGMA D, BA J. Adam: A Method for Stochastic Optimization [J]. Computer Science, 2014,abs/1412.6980.
- [15] LI S, JIANG H, WANG Z, et al. An effective computer aided diagnosis model for pancreas cancer on PET/CT images [J]. Computer methods and programs in biomedicine, 2018, 165: 205-214.

(上接第8页)

- [17] YANG M, YU K, ZHANG C, et al. Denseaspp for semantic segmentation in street scenes [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 3684-3692.
- [18] ZHANG Z, SABUNCU M. Generalized cross entropy loss for training deep neural networks with noisy labels[C]//Proceedings

of the 32nd International Conference on Neural Information Processing System, 2018:8792-8802.

- [19] BERTASIUS G, TORRESANI L, YU S X, et al. Convolutional random walk networks for semantic image segmentation [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 858-866.