

基于改进 ResNet50 的皮肤病变图像分类

王师玮, 陈俊, 易才键

(福州大学物理与信息工程学院, 福建 福州 350108)

✉ shiweiwangsw@163.com; 56851@qq.com; ycjfzu1998@163.com



摘要:为解决皮肤科医生人工识别皮肤癌过程中存在效率低、劳动强度大等问题,提出一种皮肤病变图像分类模型 DS-ResNet50。该模型在 ResNet50(深度残差网络)的基础上进行改进:设计了双尺度空洞卷积模块,通过级联不同空洞率的深度卷积核提取不同尺度的特征信息并进行融合;引入轻量型注意力模块 SimAM,使模型更好地聚焦主体目标提炼关键特征。选用 Focal Loss 函数,调节损失权重,使模型更关注难分类样本,提高对难分类样本的分类准确率。DS-ResNet50 模型在 ISIC2017 数据集上的分类准确率比 ResNet50 模型提升了 0.88%,验证了此模型的有效性。

关键词: 皮肤病变图像分类; 空洞卷积; SimAM; Focal Loss

中图分类号: TP391.4 **文献标识码:** A

Image Classification of Skin Lesions Based on Improved ResNet50

WANG Shiwei, CHEN Jun, YI Caijian

(College of Physics and Information Engineering, Fuzhou University, Fuzhou 350108, China)

✉ shiweiwangsw@163.com; 56851@qq.com; ycjfzu1998@163.com

Abstract: In order to solve the problems of low efficiency and high labor intensity in the process of dermatologist's manual identification of skin cancer, this paper proposes a skin lesion image classification model DS-ResNet50, which is an improved model on the basis of ResNet50(deep residual network). A dual-scale dilated convolutional module is designed, and the feature information of different scales is extracted and fused by cascading deep convolutional kernel of different dilation rates. The lightweight attention module SimAM is introduced to make the model better focus on the main object and extract the key features. Focal Loss function is used to adjust the weight of loss, so that the model could pay more attention to the hard-to-classify samples and improve their classification accuracy. The classification accuracy of DS-ResNet50 model on ISIC2017 dataset is 0.88% higher than that of ResNet50 model, which verifies the effectiveness of the proposed model.

Keywords: image classification of skin lesions; dilated convolution; SimAM; Focal Loss

1 引言(Introduction)

皮肤病目前已成为重要公共卫生问题。皮肤科医生人工识别皮肤病的过程中,一定程度地存在效率低、劳动强度大等问题,因此借助深度学习技术辅助医生进行诊疗,在临床

应用中具有重要的实用价值。

近年来,各种卷积神经网络(Convolutional Neural Network, CNN)层出不穷。2012年,由 KRIZHEVSKY 等^[1]提出 AlexNet 网络,在当年的 ImageNet(大规模视觉识别挑战赛)上夺冠,

掀起了 CNN 在图像分类领域的应用热潮。在此基础上, 皮肤病变识别也逐渐引入深度学习的研究方法。2014 年, 牛津大学计算机视觉几何组 (Visual Geometry Group, VGG) 提出了 VGGNet^[2], LI 等^[3]使用 VGG16(16 层结构的 VGGNet) 解决皮肤肿瘤分类的任务。ADEGUN 等^[4]研究了一种基于 DenseNet(密集连接卷积网络) 的方法检测皮肤癌, 该方法由密集块组成。张荣梅等^[5]提出用 Inception 结构取代残差网络中的卷积层、池化层, 并将其运用于皮肤黑色素瘤的分类识别中。SAELIM 等^[6]提出了一种基于 MobileNet(轻量级卷积网络) 的皮肤病变分类方法。SATAPATHY 等^[7]利用 CapsNet(胶囊网络) 对空间姿态、方向信息敏感的特点, 将其作为皮肤癌的分类器。由于皮肤病变表现出的类内相异性和类间相似性, 以及不同病变类型之间存在严重的类别不平衡情况, 因此在一定程度上增加了皮肤病变的识别难度。为了解决这些问题, 进一步提升皮肤病变图像的分类精度, 本文提出了皮肤病变图像分类模型 DS-ResNet50。

2 皮肤病变图像分类模型(Image classification model of skin lesions)

ResNet50^[8] 网络由很多残差块构建而成, 这些残差块不仅易优化, 计算负担也小, 使得 ResNet50 在网络层数加深的同时能有效解决梯度消失、梯度爆炸的现象。ResNet50 网络结构包含 49 个卷积层和 1 个全连接层。残差块中的 shortcut 连接将输入的信息进行跨层传递, 与经过卷积操作后的输出进行叠加, 使底层网络能够进行充分训练。随着残差块的堆叠, 网络深度不断增加, 网络的正确率也会显著提升。

DS-ResNet50 模型在 ResNet50 的基础上进行改进, 构建了新的瓶颈模块: 设计了双尺度空洞卷积模块, 通过级联空洞率为 1 和 3 的深度卷积核, 增加感受野, 保留更多的原始信息, 并对提取到的不同尺度的特征信息进行融合; 引入轻量型注意力模块 SimAM, 在不增加任何参数量的同时, 使模型更好地聚焦主体目标提炼关键特征, 提升易混淆样本的分类准确率。采用 Focal Loss 损失函数调整难分类、易分类样本的权重, 使模型更关注难分类样本的损失结果。模型整体结构如表 1 所示。

表 1 ResNet50 与 DS-ResNet50 结构对比

Tab.1 Structural comparison between ResNet50 and DS-ResNet50

层级	ResNet50	DS-ResNet50
conv1	7×7, 64, stride2 3×3maxpool, stride2	7×7, 64, stride2 3×3maxpool, stride2
conv2_x	$\begin{bmatrix} 1 \times 1 \times 64 \\ 3 \times 3 \times 64 \\ 1 \times 1 \times 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1 \times 64 \\ \text{SimAM} \\ 3 \times 3 \times 64, r=1 \\ 3 \times 3 \times 64, r=3 \\ \text{SimAM} \\ 1 \times 1 \times 256 \end{bmatrix} \times 3$

续表

层级	ResNet50	DS-ResNet50
conv3_x	$\begin{bmatrix} 1 \times 1 \times 128 \\ 3 \times 3 \times 128 \\ 1 \times 1 \times 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1 \times 128 \\ \text{SimAM} \\ 3 \times 3 \times 128, r=1 \\ 3 \times 3 \times 128, r=3 \\ \text{SimAM} \\ 1 \times 1 \times 512 \end{bmatrix} \times 4$
conv4_x	$\begin{bmatrix} 1 \times 1 \times 256 \\ 3 \times 3 \times 256 \\ 1 \times 1 \times 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \times 256 \\ \text{SimAM} \\ 3 \times 3 \times 256, r=1 \\ 3 \times 3 \times 256, r=3 \\ \text{SimAM} \\ 1 \times 1 \times 1024 \end{bmatrix} \times 6$
conv5_x	$\begin{bmatrix} 1 \times 1 \times 512 \\ 3 \times 3 \times 512 \\ 1 \times 1 \times 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1 \times 512 \\ \text{SimAM} \\ 3 \times 3 \times 512, r=1 \\ 3 \times 3 \times 512, r=3 \\ \text{SimAM} \\ 1 \times 1 \times 2048 \end{bmatrix} \times 3$
avgpool, fc, softmax		

2.1 双尺度空洞卷积模块

增加感受野可以使神经元接触到原始图像的更大范围, 从而捕捉到更全局的语义信息。传统的池化、增大卷积核都能扩大感受野, 但使用此方法扩宽感受野的同时, 特征图尺寸也会随之缩短, 因此采用上采样的方法复原特征图尺寸。特征图尺寸缩短再放大的过程通常会造精度上的损失, 遗失有用信息, 故本文通过空洞卷积扩宽感受野。空洞卷积是在标准的卷积中引入空洞, 在扩宽感受野的同时保持特征图的尺寸不变, 相比原来的标准卷积, 空洞卷积多了一个参数, 称为膨胀率。空洞卷积可以通过调整膨胀率控制感受野的大小, 膨胀率越大, 则感受野越大, 调整的同时不会改变特征图的分辨率。假设空洞卷积的卷积核大小为 k , 膨胀率为 r , 则其等效卷积核大小 k' 通过如公式(1)计算:

$$k' = k + (k - 1) \times (r - 1) \tag{1}$$

当 $r=1$ 时, 就是标准卷积。

本文模型在此基础上提出, 将如图 1 所示的双尺度空洞卷积模块置于原 ResNet50 的瓶颈结构中。将 2 个扩张率为 1 和 3 的空洞卷积进行级联, 扩宽感受野以获取多尺度的特征信息, 并融合这 2 个空洞卷积的输出特征图, 提高特征提取能力, 进而获得更精细的特征结构。

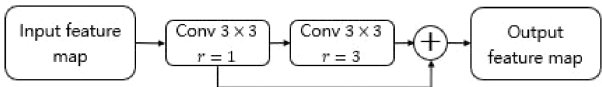


图 1 双尺度空洞卷积

Fig. 1 Dual-scale dilated convolution

2.2 轻量型注意力 SimAM 模块

注意机制是深度学习中一种常用的技术,基本思想是基于输入序列的不同部分与当前预测或输出的相关性加权它们的贡献,一般通过为输入序列的每个元素计算一组注意力得分,然后使用这些得分计算序列元素的加权和,使得模型在进行预测或生成输出时专注于输入序列的特定部分。通过注意力机制,模型可以关注更值得关注且更应该关注的地方。

SimAM 模块^[9]是一种无参数的轻量型注意力模块,可以无缝地集成到模型架构中一起进行训练。在 DS-ResNet50 模型中引入 SimAM 模块,在不额外增加参数的情况下,可以捕获更多的细节特征,使模型更加关注易混淆样本。

SimAM 根据神经科学理论设计一种能量函数用于发现重要神经元,并据此实现注意力权值的计算。在神经科学中,与周围神经元相比,信息较丰富的神经元通常会展现出不同的放电模式,并且活跃的神元往往会抑制周围神经元的活动,即所谓的空域抑制。所以,应当给予具有明显空域抑制效应的神元更高的优先级,定义如公式(2)所示的能量函数找到这些重要的神经元:

$$e_t^* = \frac{4(\hat{\sigma}^2 + \lambda)}{(t - \hat{\mu})^2 + 2\hat{\sigma}^2 + 2\lambda} \quad (2)$$

其中, $\hat{\mu} = \frac{1}{M} \sum_{i=1}^M x_i$, $\hat{\sigma}^2 = \frac{1}{M} \sum_{i=1}^M (x_i - \hat{\mu})^2$ 。

由公式(2)可以看出,神经元 t 与周围神经元之间的差异会随着能量的变化而变化,即能量越低,神经元 t 与周围神经元的区别越大,重要性也越高,故神经元的重要性可以通过 $1/e_t^*$ 得到。通过能量函数判断神经元的重要性,然后使用缩放运算对这些特征进行细化处理。

2.3 Focal Loss 损失函数

多分类实验中,数据集经常会出现类别不平衡的问题,导致模型偏向数据量大的类,对数据量小的类的学习就不够充分,最终使少样本的类别分类结果较差;数据集中也会存在难分类样本,难分类样本与其他类别样本差异细微,很容易造成误分,影响模型性能。

针对上述问题,本文所提方法首先在数据集的预处理过程中,采用数据增强方法进行样本补充,缓解数据不平衡的问题,平衡模型在训练过程中的偏好。其次选择 Focal Loss^[10]作为损失函数,通过调整难、易分类样本对总损失结果的权重,增加难分类样本的损失值,减小易分类样本的损失值,使模型更关注难分类样本,避免难分类样本淹没在大量易分类样本中,让模型对难分类样本进行更充分的学习,降低难分类样本的误分率,从而提升总体的分类准确率。Focal Loss 函数是基于交叉熵损失函数的改进,具体表达式如下:

$$F_L(p_t) = -a_t(1-p_t)^r \log_2 p_t \quad (3)$$

其中, p_t 为预测概率;抑制参数 $r \geq 0$,抑制参数与权重成反比,即 r 越大,权重越低,一般根据经验取 $r=2$;各类别权重 a_t 的取值范围为(0,1),根据样本数据量的相对大小取值,样本多的类别权重取值较小,样本少的类别权重取值较大。

3 实验与结果分析(Experiment and result analysis)

3.1 实验环境

本文实验在 Windows 10 操作系统下完成,基于 Pytorch 深度学习框架。硬件环境为 16 GB 内存的 IntelCore i7-9700K CPU 处理器,8 GB 显存的 NVIDIA GeForce RTX 2070 显卡。

3.2 数据来源与数据增强

本文实验数据来自 ISIC2017 皮肤镜图像分类数据集,数据集由 2 000 张训练集、150 张验证集、600 张测试集组成。数据集共有 3 种类别的皮肤病变图像,分别为黑色素瘤(Melanoma, MEL)、脂溢性角化病(Seborrheic Keratosis, SK)、黑色素细胞痣(Melanocytic Nevus, MN)。原训练集中各类别图像的具体分布情况如图 2 所示。

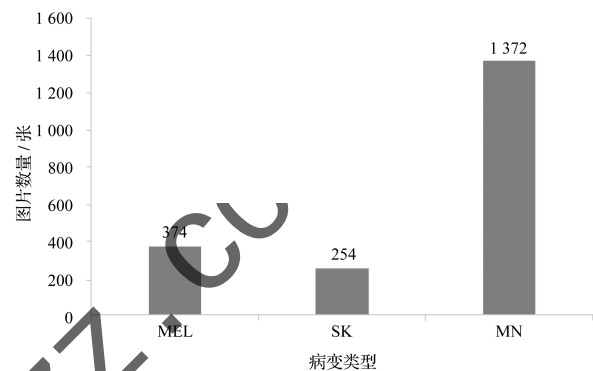


图 2 原训练集的图片分布

Fig. 2 Image distribution of original training set

由图 2 可以看出,训练集中各类别图像的数据量少且各类别间数量差异很大,所以本文通过数据增强的方法产生新的样本图像,在扩充样本数据量解决样本分布不平衡问题的同时,也增强了模型的泛化能力。数据增强的具体操作为逆时针旋转 90°、180°、270°、水平或垂直翻转、随机缩放等。经调整后的训练集样本分布如图 3 所示。

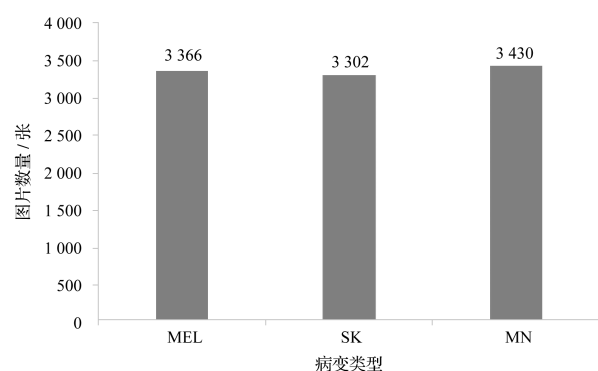


图 3 数据增强后的训练集图片分布

Fig. 3 Image distribution of training sets after data enhancement

3.3 实验参数设置

本文实验使用 Adam 作为优化器,损失函数选择 Focal Loss 函数。设置每批次训练 8 张图像,模型的最大训练周期为 100 个,采用稳定减小学习率、早停法防止数据过拟合。初始学习率设置为 0.000 5,稳定减小学习率方法的 patience

参数设置为 7，当 7 个训练周期后验证性能没有提高，学习率会降低为原来的一半。早停法的 patience 参数设置为 15，当连续 15 个训练周期的验证性能没有提高，就会终止训练。

3.4 评价标准

为了全面评估模型的分类性能，使用准确率(ACC)、精确率(P)、召回率(R)及 F1-score 值(F1)作为评价标准。

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$P = \frac{TP}{TP + FP} \quad (5)$$

$$R = \frac{TP}{TP + FN} \quad (6)$$

$$F1 = \frac{2TP}{2TP + FN + FP} = \frac{2 \times P \times R}{P + R} \quad (7)$$

其中，TP(真阳性)表示将正例预测为正例的数量，FP(假阳性)表示将负例预测为正例的数量，FN(假阴性)表示将负例预测为负例的数量，TN(真阴性)表示将正例预测为负例的数量。一般而言，准确率、精确率、召回率和 F1-score 的值越高，模型的分类性能越好。

3.5 结果与分析

为了验证本文模型的有效性，在数据增强后的皮肤病变图像分类数据集上进行实验，表 2 展示了不同类别皮肤病变图像的分类结果。从表 2 中可知，本模型在 3 种皮肤病变类型的分类上的精确率、召回率和 F1-score 均表现良好。

表2 3种皮肤病变类型的分类结果

Tab.2 Classification results of 3 types of skin lesions

类别	P	R	F1
MEL	0.88	0.89	0.89
MN	0.90	0.95	0.92
SK	0.96	0.88	0.92

混淆矩阵是准确性评估的标准格式，本文的皮肤病变分类模型输出分为 3 类，因此用 3×3 矩阵表示，分类结果的混淆矩阵如图 4 所示。

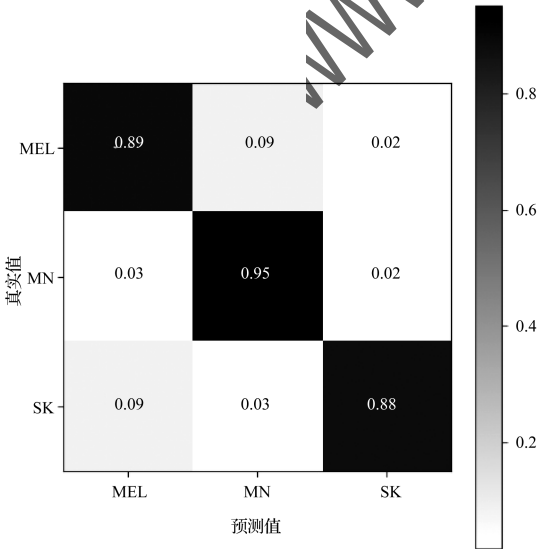


图4 分类结果的混淆矩阵

Fig. 4 Confusion matrix of classification results

将本文算法与经典模型 VGG16、GoogleNet、ResNet50、MobileNetV2、ShuffleNetV2、InceptionV3 进行分类准确率对比，实验结果如表 3 所示。

表3 不同模型分类准确率对比

Tab.3 Comparative experiment of different models' performance

模型	ACC/%
VGG16	86.26
GoogleNet	88.19
ResNet50	90.07
MobileNetV2	87.62
ShuffleNetV2	85.43
InceptionV3	88.03
本文模型	90.95

从表 3 中可以看出，本文算法在准确率上的实验结果为 90.95%，分别高于 VGG16、GoogleNet、ResNet50、MobileNetV2、ShuffleNetV2、InceptionV3 模型 4.69%、2.76%、0.88%、3.38%、5.52%、2.92%。由此可见，本文模型具有更好的分类效果，是一个高精度、高性能的皮肤病变分类网络模型。

本文模型在训练过程中的损失如图 5 所示，可以看出，在训练过程中采用早停法，训练损失和验证损失都不断减小，最终保存了训练过程中最低的损失值，模型得到了有效的训练。

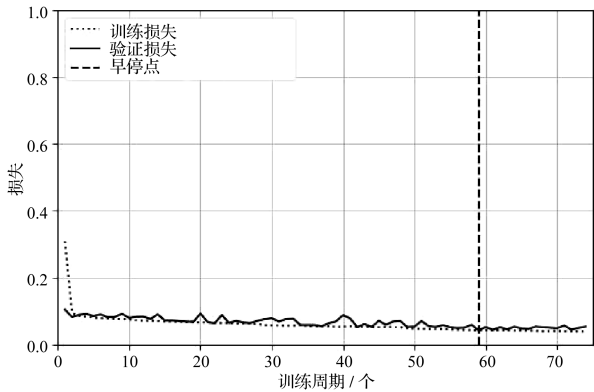


图5 训练过程中的损失

Fig. 5 Loss during training process

4 结论(Conclusion)

本文模型在 ResNet50 的基础上进行改进，提出了一种皮肤病变分类模型 DS-ResNet50：设计了双尺度空洞卷积模块，通过级联不同空洞率的深度卷积核，在特征提取阶段提取不同尺度的特征信息并进行融合；引入轻量型注意力模块 SimAM，使模型更好地聚焦主体目标提炼关键特征。训练过

程中选用 Focal Loss 损失函数,通过调节权重参数使模型更关注难分类的样本。在公开数据集 ISIC2017 上进行实验,并与几款经典模型进行对比,结果表明本文提出的 DS-ResNet50 模型在皮肤病变分类中准确率在几种模型中最高,达到了 90.95%,具有良好的分类效果,可为临床医学诊断提供辅助。

参考文献(References)

- [1] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6):84-90.
- [2] KAREN S, ANDREW Z. Very deep convolutional networks for large-scale image recognition[J]. Computer Science, 2014, 6(1):1-14.
- [3] LI Q T, RUAN H G. Automatic skin lesion analysis towards melanoma detection[C]// IES Organizing Committee. 2017 21st Asia Pacific Symposium on Intelligent and Evolutionary Systems. New York: IEEE, 2017:106-111.
- [4] ADEGUN A A, VIRIRI S. FCN-based DenseNet framework for automated detection and classification of skin lesions in dermoscopy images[J]. IEEE Access, 2020, 8:150377-150396.
- [5] 张荣梅,张琦,刘院英.基于 Inception 深度残差网络的皮肤黑色素瘤图像分类算法[J].计算机系统应用,2021,30(7):142-149.
- [6] SAEIM W, WETTAYAPRASIT W, AIYARAK P. Convolutional neural networks using MobileNet for skin lesion classification[C]// JCSSE Organizing Committee. International Joint Conference on Computer Science and Software Engineering. New York: IEEE, 2019:242-247.
- [7] SATAPATHY S C, CRUZ M, NAMBURU A, et al. Skin cancer classification using convolutional capsule network[J]. Journal of Scientific and Industrial Research, 2020, 79(11):994-1001.
- [8] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]// CVPR Organizing Committee. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE Computer Society, 2016:770-778.
- [9] YANG L, ZHANG R Y, LI L, et al. Simam: A simple, parameter-free attention module for convolutional neural networks[C]// ICML Organizing Committee. Proceedings of Machine Learning Research. California: JMLR, 2021:11863-11874.
- [10] LIN T Y, GOYAL P, GIRSHICK R, et al. Focal loss for dense object detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 42(2):318-327.

作者简介:

王师玮(1998-),女,硕士生.研究领域:计算机视觉.

陈俊(1978-),男,硕士,副教授.研究领域:物联网通信.

易才键(1998-),男,硕士生.研究领域:深度学习,计算机视觉.

(上接第 58 页)

参考文献(References)

- [1] 崔金栋,李晨雨,李菲菲.大数据背景下主流社交媒体热点发现机制研究[J].情报科学,2021,39(12):72-79.
- [2] 雷刚,赖宇,余梦雨.基于公有云的微博网络爬虫及舆情可视化分析实验平台设计[J].实验室研究与探索,2021,40(10):273-276,294.
- [3] 毕硕本,黄铜,赵爱,等.基于社会网络分析法对信息流空间传播特征的研究[J].计算机科学,2018,45(12):279-287.
- [4] 舒万畅.爬虫技术在大数据领域中的应用分析[J].科学技术创新,2018(36):91-92.
- [5] 罗咪.基于 Python 的新浪微博用户数据获取技术[J].电子世界,2018(5):138-139.
- [6] 董洋溢,李伟华,陈世亮.中文领域知识半自动化 OWL 本体构建方法研究[J].计算机应用与软件,2016,33(5):17-19,65.
- [7] ALFRED V.编译原理[M].北京:机械工业出版社,2009:20-85.
- [8] SnowNLP: Simplified Chinese Text Processing[EB/OL]. (2020-1-19)[2021-2-18]. <https://github.com/isnowfy/snownlp>.
- [9] 徐光美,刘宏哲,张敬尊.基于特征加权的多关系朴素贝叶斯分类模型[J].计算机科学,2014,41(10):283-285.
- [10] 钱爱兵,江岚.基于改进 TF-IDF 的中文网页关键词抽取——以新闻网页为例[J].情报理论与实践,2008,31(6):945-950.
- [11] PURDA L, SKILLICORN D. Accounting variables, deception, and a bag of words: Assessing the tools of fraud detection[J]. Contemporary Accounting Research, 2015, 32(3):1193-1223.

作者简介:

霍英(1975-),女,博士,教授.研究领域:大数据与物联网,社会网络.

丘志敏(1973-),男,硕士,副教授.研究领域:社会网络.

李小帆(1996-),男,本科,工程师.研究领域:舆情处理.

李彦廷(1975-),男,博士,副教授.研究领域:舆情处理.