

基于深度学习的鱼类识别与检测的算法研究*

王文成¹, 蒋 慧¹, 乔 倩¹, 祝捍皓², 郑 红²

(1. 浙江海洋大学 船舶与机电工程学院, 浙江 舟山 316022;

2. 浙江海洋大学 海洋科学与技术学院, 浙江 舟山 316022)

摘 要: 鱼类分类识别在渔业资源研究、鱼类知识的科学推广、水产养殖加工、稀有物种保护等领域具有广泛的应用前景。针对大菱鲆、黄鳍鲷、金钱鱼、鲻鱼这四种鱼类, 利用 PyTorch 框架为基础, 通过 ResNet50 网络模型, 用不同的算法对其进行分类识别, 不断对模型进行优化, 对四种鱼类训练学习, 通过测试其准确率达到 96% 以上。同时用 PyQt5 开发了 GUI 可视化界面, 通过界面图片的选择和预测功能按钮的操作, 测试结果实际类别与预测类别一致, 用 DSOD 框架做了水下目标实时跟踪检测, 提高了对小目标的检测率, 同时保持了模型的检测速度, 检测结果达到期望。

关键词: PyTorch 框架; ResNet50 网络; PyQt5 可视化界面; DSOD 目标检测器

中图分类号: **文献标识码:** A **DOI:** 10.19358/j.issn.2096-5133.2020.08.011

引用格式: 王文成, 蒋慧, 乔倩, 等. 基于深度学习的鱼类识别与检测的算法研究[J]. 信息技术与网络安全, 2020, 39(8): 57-61, 66.

Research on fish recognition and detection algorithm based on deep Learning

Wang Wencheng¹, Jiang Hui¹, Qiao Qian¹, Zhu Hanhao², Zheng Hong²

(1. College of Shipping and Electromechanical Engineering, Zhejiang Ocean University, Zhoushan 316022, China;

2. College of Marine Science and Technology, Zhejiang Ocean University, Zhoushan 316022, China)

Abstract: Fish classification and identification has broad application prospects in the fields of fisheries resource research, scientific promotion of fish knowledge, aquaculture processing, and protection of rare species. In this paper, for the four species of turbot, yellowfin snapper, golden fish, mullet, based on the PyTorch framework, the ResNet50 network model is used to classify and identify them with different algorithms, and the model is continuously optimized. The accuracy rate of fish training is 96%. At the same time, the GUI visual interface was developed with PyQt5. Through the selection of interface pictures and the operation of the prediction function button, the actual category of the test results is consistent with the predicted category. The real-time tracking and detection of underwater targets was done using the DSOD framework, which improved the detection rate of small targets. At the same time, the detection speed of the model is maintained, and the detection results meet expectations.

Key words: PyTorch framework; ResNet50 network; PyQt5 visual interface; DSOD target detector

0 引言

海洋作为地球上最大的生态系统包含极其丰富的生物和非生物资源。虽然我国海洋面积有 300 多万平方千米, 但是渔业资源作为海洋资源中重要的一项资源, 近年来呈现逐渐衰退趋势^[1]。对周围海域内海底环境进行实时监测, 分析研究海域中代表性鱼类的分布概况及生活习性, 对于渔业资源可

持续发展、探测未知生物资源有着重要意义。

目标识别是计算机视觉领域重要的研究方向之一。随着科学技术的进步、计算硬件设备水平的提高, 基于深度学习的目标识别检测方法受到了人们广泛关注。杜卫东等^[2]提出一种基于支持向量机 (SVM) 的多方位声散射数据决策融合的鱼类识别方法, 识别准确率达到 90% 以上, 姚润璐等^[3]通过分割鱼体, 获得鱼肚、鱼背图像, 通过提取纹理特征和

* 基金项目: 舟山市科技计划项目 (2017C41003)

形状特征,识别率达到 75%,相较于传统的目标检测算法,基于卷积神经网络(CNN)的目标检测算法通过学习自动提取颜色、轮廓等底层的特征和更高级、抽象的特征,然后利用分类模型进行分类检测,其检测结果具有更高的精度和鲁棒性。林明旺等^[4]采用 VGG16 模型进行鱼类分类识别,但是数据集背景干扰强度大,虽然进行微调,稳定性仍然不好。王文成等^[5]用 ResNet50 网络模型对十种鱼类做了分类识别,准确率达到 93.33%,但是数据集数量有限,识别精度不高;张俊龙等^[6]提出在预处理过程中使用权重化卷积操作,对海洋鱼类识别准确率提升 23%,在视频观测中精准识别鱼类;梁红等^[7]通过对图像进行水下降噪的方法,并且对预训练的网络微调参数,在海洋鱼类数据集上识别准确率达到 85.08%。

基于 CNN 的目标检测算法依据是否需要提取候选区域分为:基于有候选区域的目标检测算法和基于无候选区域的目标检测算法两种。区域卷积神经网络(R-CNN)^[8]、快速区域卷积神经网络(Fast R-CNN)^[9]、超快速区域卷积神经网络(Faster R-CNN)^[10],都属于代表性的基于候选区域的目标检测算法。该类算法从图像中事先提取一系列可能的对象候选区域,然后从这些候选区域中提取特征并进行判断分类。2019 年,袁红春等人^[11]利用 Faster R-CNN 目标检测方法应用到水下鱼类种类识别中,识别准确率达到 98.12%。该类算法的检测准确率较高,但是处理速度较慢。无候选区域的目标检测算法的代表性算法有单级式目标检测(YOLO)^[12]、单点多盒检测器(SSD)^[13]及其改进的算法 YOLO v2^[14]和深度监督对象检测器(DSOD)。此类算法去除候选区域提取过程,使用神经网络直接将一整张图片输入神经网络,然后预测出 bounding box 的坐标和物体的类别、置信度,检测物体速度较快。李庆忠等^[15]对 YOLO 实时目标检测算法进行了改进,采用迁移学习法训练网络,预处理中图像增强以及视频帧数检测速率上做了改变,该算法提升了对海底小目标检测的性能,但存在检测精度低、检测小目标会有遗漏等问题。SSD 算法同时融合了 proposal-free 的提取和 Faster R-CNNs 算法的优点,与其他检测算法相比,即使输入图像尺寸小,SSD 依然有更好的精度。DSOD 是一种改进的 SSD 网络框架算法。相比于 SSD 目标检测算法,DSOD 整个网络采用密集的短连接和通道拼接技

术,使得 DSOD 的模型参数更小,同时能直接在目标数据集上进行训练,而无需在 ImageNet、COCO、PASCAL VOC 等公开数据集上进行预训练。

基于上述的分析,由于分类识别和检测任务在统计分布和损失函数的差异,存在学习上的偏差,为了解决这个问题,本文利用深度监督对象检测器(DSOD)方法对选定鱼类图像进行检测分类识别研究,该模型具有高效的目标检测网络,保持了处理速度,提升了对小目标的检测性能。

1 技术路线

CNN 是当前深度学习算法实现的主要途径,每个模块主要由数据输入层、卷积层、激励层、池化层和全连接层堆叠在一起组成。一般来说网络层越深,训练得到的效果就越好,但是有时候有些网络会出现过拟合的现象,需要进行微调并不断测试。引入残差网络可以很好地抑制过拟合现象,ResNet 网络模型就用到了残差结构^[16],CNN 各卷积层共享权重,减少很大一部分参数,提高模型的计算效率。池化层通过对卷积的输出进行采样来降低下一层的数据速率,可以将图像用作直接输入,消除复杂的提取 FeatureMap 的过程。

1.1 网络模型

ResNet50 包含 49 个卷积层和 1 个全连接层。其中,第二至第五阶段的 ID BLOCK $\times 2$ 表示两个大小不变的残差,CONV BLOCK 表示增加比例的残差,每个残差该块包含三个卷积层,结构如图 1 所示。

图 1 中的 CONV 是卷积操作的卷积层,Batch-

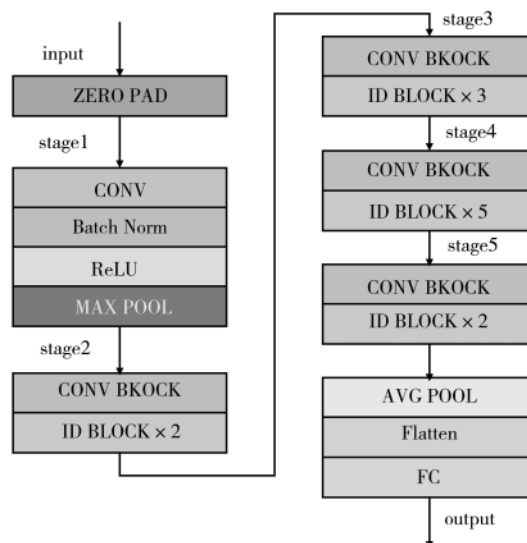


图 1 网络模型

Norm 是批量正则化处理,ReLU 是激活函数,MAX POOL 表示最大池化操作,AVG POOL 表示全局平均池化层操作,stage1 到 stage5 表示残差块。输入层把不同图片的装换为 224×224 的大小进行输入,激活函数使用 ReLU,即 $f(x)=\max(0,x)$,最后一层是全连接层,需要将每个类别对应的计算得分转换为与总和为 1 的每个类别对应的概率值,使用 Softmax 函数进行分类,即 $y_i=\exp(a_i)/\sum_{i=1}^m \exp(a_i)$ 。

1.2 DSOD 框架

DSOD 方法是一种类似于 SSD 的多尺度 proposal-free 检测框架,不需要提取候选框,也不需要预训练过程,能够从头开始训练模型,并达到很好的检测效果。DSOD 的网络结构可以分为两部分:用于特征提取的主干网络和多尺度预测网络。主干网络是深度监督 DenseNets 结构的一种变体,由一个主干块、四个密集块、两个过渡层和两个 w/o 组成池过渡层。如图 2 所示,DSOD 的架构由一个主干(Stem)模块、四个密集(Dense Block)模块、两个过渡层(Transition Layer)模块和两个无池化层的过渡层(Transition w/o Pooling Layer)模块组成。其中主干模块由 3 个 3×3 的卷积层和一个 2×2 的最大池化层组成,第一个卷积层的步长为 2,另外两个卷积层步长为 1,池化层步长为 2;密集模块是由 1×1 的卷积层和 3×3 的卷积层组成的,卷积层数量密集排列;过渡层由 1×1 卷积层和 2×2 的最大池化层组成,步长为 2;无池化层的过渡层只有一个 1×1 卷积层。最先进的对象检测器过分依赖在大规模数据集上预训练成的网络,分类和检测任务在统计分析损失函数上的差别,存在学习偏差,DSOD 模型是为了解决这两个方面的问题而产生的模型。

2 实验及结果分析

2.1 实验数据集

本文实验所需的数据均来源于手动搜集的网络图片,四种鱼类,分别为大菱鲂、黄鳍鲷、金钱鱼、鲮鱼,共 1 123 张图像,每种鱼类设置验证集 42

张图像,并采取相同的预处理操作,以确保最终验证的准确性和测试结果将数据分成相同大小的批次,方便后续模型训练。

2.2 实验平台

基于 PyTorch 框架实现的检测器。所有的模型都是在 NVidia GTX2080Ti GPU 上用 SGD 求解器从零开始训练的。由于 DSOD 特征映射的每个尺度都由多个分辨率串联而成,因此采用 L2 归一化技术^[17]。大多数训练设置都遵循 SSD,包括默认框的数据扩充、规模和高宽比,有自己的学习速率调度和小批量大小设置。细节将在实验部分给出。

2.3 结果分析

课题组在前期研究中针对 10 种鱼类进行分类识别,十种鱼类分别为鲮鱼、巴鲮鱼、金钱鱼、银鲳鱼、大菱鲂、鲈鱼、颌针鱼、黄鳍鲷、绿鳍鱼、黑鲷鱼,共 908 张图像,设置训练集 807,验证集 101 进行分层采样。前期研究的实验结果如表 1 所示,分别从 precision、recall、f1-score 的这几个指标中反映出来,计算这些指标有四个概念:TP、FP、TN、FN,从表 1 可以看出,该模型的准确率和召回率都比较高。

表 1 分类识别实验结果

	precision	recall	f1-score	support
颌针鱼	1.00	1.00	1.00	61
鲮鱼	1.00	0.99	1.00	137
金钱鱼	1.00	1.00	1.00	89
大菱鲂	1.00	1.00	1.00	113
银鲳鱼	0.99	1.00	0.99	83
黑鲷鱼	0.99	0.98	0.99	105
鲈鱼	1.00	1.00	1.00	100
黄鳍鲷	0.99	1.00	0.99	96
绿鳍鱼	0.99	1.00	0.99	76
巴鲮鱼	1.00	0.97	0.98	33

$$\text{precision}=\frac{\text{TP}}{\text{TP}+\text{FP}} \tag{1}$$

$$\text{recall}=\frac{\text{TP}}{\text{TP}+\text{FN}} \tag{2}$$

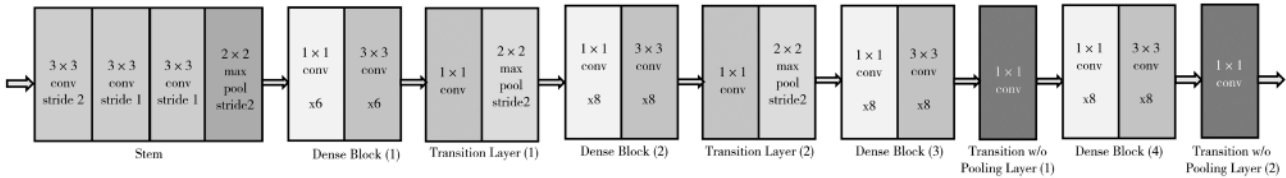
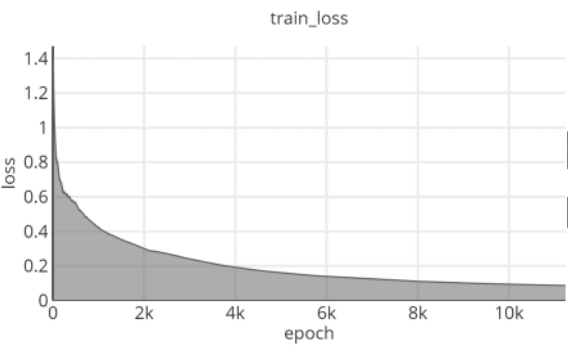


图 2 DSOD 架构

$$f1-score = \frac{2 * precision * recall}{precision + recall} \quad (3)$$

本文实验选择其中四种鱼类，分别为大菱鲆、黄鳍鲷、金钱鱼、鲷鱼，并且增加了每种鱼类的数量，共 1 123 张图像，每种鱼类设置验证集 42 张图像，剩余的图片用作训练学习。本文是在 PyTorch 框架下，使用 ResNet50 网络模型进行实验，分类识别用的优化器是 Adam，该模型的参数设置如下，学习率 $1e^{-4}$ ， $lr_decay = 1e^{-4}$ ， $weight_decay = 1e^{-4}$ 。如图 3(a) 所示，横坐标为迭代次数，纵坐标为损失。从图中可以看出随着迭代次数的增加损失值逐渐下降，迭代次数 2 000 次之前下降速率较快，往后曲线逐渐平稳。图 3(b) 所示为随迭代次数的增加识别准确率的变化，从图中可以看出随着迭代次数的增加准确率逐渐增加，当迭代次数为 2 000 次的时候，训练准确率达到 90% 以上，当迭代次数达到 8 000 次的时候，训练准确率达到 95% 以上，之后持续为平稳状态。



(a) 损失

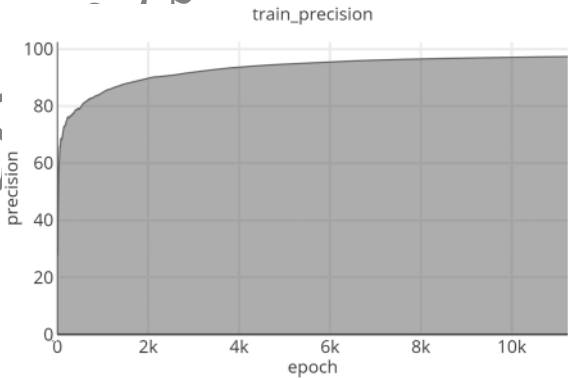
同时，数据结果显示该模型收敛速度快，识别精度高，验证集误差小。

2.4 PyQt 界面

图 4 所示为利用 PyQt5 做的 GUI 可视化界面，在界面中有鱼类识别和鱼类视频检测两个功能检测窗口，在鱼类识别窗口中有选择图片、预测、实际类别、预测类别以及退出这几个功能。当从数据库中随意选择一个图片进行预测，如图 4(a) 所示，系统给出的实际类别是金钱鱼。通过点击预测功能按钮，图 4(b) 给出预测类别为金钱鱼，与图(a)给出的实际类别相同。本文经过多次实验测试，图(a)中的实际类别和图(b)中的预测类别总是一致的，预测准确率是很高的，PyQt5 可视化界面中都能准确呈现出来识别种类，GUI 可视化界面能够直观地反映识别的准确度。

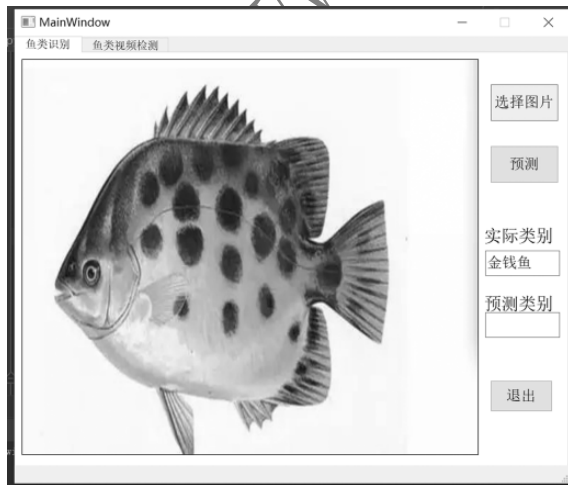
2.5 DSOD 算法目标检测分类结果

本文用于实验检测的视频来源于 Labeled fishes



(b) 准确率

图 3 准确率与损失变化



(a) 实际类别



(b) 预测类别

图 4 PyQt5 可视化界面

in the wild^[18]。目标检测的训练配置参数如下,初始学习率 0.01, momentum = 0.9, weight_decay = $5e^{-4}$, 学习率采用自适应的学习率下降策略,当学习率停滞不降一定次数之后,降低学习率。目标检测结果如图 5 所示,通过图(a)可以发现,该检测框架对于比较小的物体也能够检测出来,图中有一些小的目标,由于水下的恶劣环境的原因,并没有完全检测出来,在水下基础设施较好的情况下,可以达到较好的检测效果,视频中出现的鱼类都能准确地标注出来,检测效果很明显,而且 DSOD 继承了 SSD 的检测速度快的优点,对于遗漏小目标的缺点得到提升。

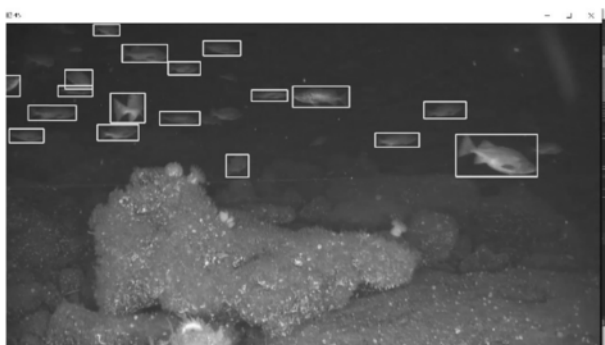


图 5 水下实时跟踪检测

DSOD 方法是一种类似于 SSD 的多尺度 proposal-free 检测框架,引入了密集网络(Densets)的思想,即将某些层的输入进行变换并与密集网络结合,DSOD 的参数数量大大减少。DSOD 不需要预训练模型的检测算法,训练检测网络从零开始训练,能达到最先进的精度,快速的处理速度,与从大数据训练而来的复杂模型相比,更好的模型结构可以实现相似或更好的性能。

3 结论

本文从深度学习的角度出发,利用卷积神经网络在计算机视觉中的优势,提出了一种基于 PyTorch 深度学习框架的鱼类图像识别算法,以 ResNet50 为网络模型,采用 Adam 优化算法对模型参数进行更新。在 PyQt5 可视化界面中,实验结果预测值和实际值正确率很高。该模型能够抑制过拟合,收敛速度较快,训练时间较短,残差网络解决了深度网络训练艰难的问题,提高了性能。提出了一种基于深度监督的 DSOD 框架,该框架可以对目标检测器进行从无到有的学习。该模型检测速率快,对于目标检测遗漏较少。该算法在水产养殖、保护渔业资源方面都有广阔的应用前景。

参考文献

- [1] 李继龙,曹坤,丁放,等.基于渔获物统计的中国近海鱼类营养级结构变换及其与捕捞作业的关系[J].中国水产科学,2017,024(1):109-119.
- [2] 杜伟东,李海森,魏玉阔,等.基于 SVM 的决策融合鱼类识别方法[J].哈尔滨工程大学学报,2015,36(5):623-627.
- [3] 姚润璐,桂咏雯,黄秋桂.基于机器视觉的淡水鱼品种识别[J].微型机与应用,2017,36(24):37-39.
- [4] 林明旺.深度学习在鱼类图像识别与分类中的应用[J].数字技术与应用,2017(4):96-97.
- [5] 王文成,蒋慧,乔倩,等.基于 ResNet50 网络的十种鱼类图像分类识别研究[J].农村经济与科技,2019,30(19):60-62.
- [6] 张俊龙,曾国荪,覃如符.基于深度学习的海底观测视频中鱼类的识别方法[J].计算机应用,2019,39(2):376-381.
- [7] 梁红,金磊磊,杨长生.小样本情况基于深度学习的水下目标识别研究[J].武汉理工大学学报(交通科学与工程版),2019,43(1):10-14.
- [8] NING Z, DONAHUE J, GIRSHICK R, et al. Part-Based R-CNNs for fine-grained category detection[C]. Proceedings of the European Conference on Computer Vision, F, 2014.
- [9] GIRSHICK R. 2015 IEEE International conference on computer vision (ICCV)[C]. Proceedings of the IEEE International Conference on Computer Vision, F, 2015.
- [10] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2017, 39(6): 1137-1149.
- [11] 袁红春,张硕.基于 Faster R-CNN 和图像增强的水下鱼类目标检测方法[J/OL].大连海洋大学学报: 1-11[2020-05-25]. <https://doi.org/10.16535/j.cnki.dlhxb.2019-146>.
- [12] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: unified, real-time object detection[J]. arXiv: 1506.02640, 2015.
- [13] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]. Proceedings of the European Conference on Computer Vision, F, 2016.
- [14] REDMON J, FARHADI A. YOLO9000: better, faster,

(下转第 66 页)

型、LSSVM 模型相比在趋势、精度、平均绝对误差及均方根误差方面均有非常大的提升,验证了 LSTM 神经网络在时间序列处理上的优越性,也验证了 LSTM 模型在浆液 pH 值预测的有效性 & 通用性。

4 结论

针对燃煤电厂 WFGD 过程中石灰石浆液 pH 值的变化受多个变量的影响,且变量之间具有相关性和现场数据具有时序特性,本文提出了一种基于长短期记忆网络的 pH 值预测模型。首先,通过机理分析初步筛选辅助变量后进一步使用互信息确定辅助变量;然后,建立 LSTM 神经网络模型;最后,使用燃煤电厂数据对模型进行测试。模型测试结果表明本文所提出的 LSTM 模型相对 BP 神经网络模型、LSSVM 模型预测精度高、泛化能力强。

需要说明的是,由于数据量有限,本研究只能归纳出高负荷状态下的浆液 pH 值预测模型。然而随着燃煤电厂智能化水平的提高,电厂数据海量化和高维化已经成为必然趋势。合理使用这些数据可以进一步建立全工况的浆液 pH 值预测模型,这也是后续的研究重点。本文结果对 LSTM 算法在燃煤电厂的实际应用具有一定借鉴意义。

参考文献

- [1] 康英伟,孙智滨,常俊.基于改进差分进化算法的湿法烟气脱硫系统模型辨识[J].动力工程学报,2019,39(10):834-839.
- [2] 李存杰,张军,张涌新,等.基于 pH 值分区控制的湿法烟气脱硫增效研究[J].环境科学学报,2015,35(12):4081-4087.
- [3] 张嘉,黄景涛,茅建波.基于 LSTM 的电站锅炉 NO_x 排放浓度建模研究[J].计算机与应用化学,2018,35(11):903-909.
- [4] 王志甄,邹志云.基于神经网络的 pH 中和过程非线性

性预测控制[J].化工学报,2019,70(2):678-686.

- [5] MOTA A S, MENEZES M R, SCHMITZ J E, et al. Identification and online validation of a pH neutralization process using an adaptive network-based fuzzy inference system[J]. Chemical Engineering Communications, 2016, 203(4): 516-526.
- [6] CHEN X, XUE A K, PENG D L, et al. Modeling of pH neutralization process using fuzzy recurrent neural network and DNA based NSGA-II[J]. Journal of the Franklin Institute, 2014, 351(7): 3847-3864.
- [7] 李泽龙,杨春节,刘文辉,等.基于 LSTM-RNN 模型的铁水硅含量预测[J].化工学报,2018,69(3):992-997.
- [8] 焦李成,杨淑媛,刘芳,等.神经网络七十年:回顾与展望[J].计算机学报,2016,39(8):1697-1716.
- [9] 王康成,尚超,柯文思,等.化工过程深度神经网络软测量的结构与参数自动调整方法[J].化工学报,2018,69(3):900-906,1253.
- [10] 王文广,赵文杰.基于 GRU 神经网络的燃煤电站 NO_x 排放预测模型[J].华北电力大学学报(自然科学版),2020,47(1):96-103.
- [11] 彭燕,刘宇红,张荣芬.基于 LSTM 的股票价格预测建模与分析[J].计算机工程与应用,2019,55(11):209-212.
- [12] 陈卫华,徐国祥.基于深度学习和股票论坛数据的股市波动率预测精度研究[J].管理世界,2018,34(1):180-181.

(收稿日期:2020-05-09)

作者简介:

金秀章(1969-),男,博士,副教授,主要研究方向:先进控制策略在大型电力机组的控制研究、信息融合技术。

景昊(1994-),男,硕士,主要研究方向:机器学习算法研究、火电厂控制策略优化。

(上接第 61 页)

stronger[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), 2016.

- [15] 李庆忠,李宜兵,牛炯.基于改进 YOLO 和迁移学习的水下鱼类目标实时检测[J].模式识别与人工智能,2019,32(3):193-203.
- [16] 孙琪翔,何宁,张则宇.基于深度学习方法的人体行为识别综述[C].中国计算机用户协会网络应用分会 2019 年第二十三届网络新技术与应用年会论文集,

F, 2019.

- [17] LIU W, RABINOVICH A, BERG A C. ParseNet: looking wider to see better[J]. Computer Science, 2015.

(收稿日期:2020-05-26)

作者简介:

王文成(1995-),男,硕士,主要研究方向:农业工程与信息技术。

蒋慧(1996-),女,硕士,主要研究方向:声学仪器设备。

版权声明

经作者授权，本论文版权和信息网络传播权归属于《信息技术与网络安全》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《信息技术与网络安全》编辑部
中国电子信息产业集团有限公司第六研究所