

结合批规范化层的深度学习模型 在水中目标识别中的应用^{*}

孙悦, 彭圆, 贾连徽, 曹琳, 郭欣雨, 徐剑秋

(水下测控技术重点实验室, 辽宁 大连 116013)

摘要: 针对深度学习模型在训练水声样本数据时会出现稳定性差进而导致分类识别效果不佳的问题, 从网络的局部连接、空间位置排列以及模型设计的角度出发, 研究在原有的一维序列卷积核与一维序列池化的基础上, 引入批规范化层, 构建了深度学习网络模型。通过归一化处理, 达到加速网络模型的收敛过程以及提高训练过程中的稳定性的目的。最终为验证该模型的有效性, 对3类水声目标样本数据进行网络训练和模型验证, 证明该模型对水声目标数据分类识别有一定程度的性能提升。

关键词: 水声目标; 深度学习; 分类; 网络模型

中图分类号: TP183; TP391

文献标识码: A

DOI: 10.19358/j.issn.2097-1788.2024.04.007

引用格式: 孙悦, 彭圆, 贾连徽, 等. 结合批规范化层的深度学习模型在水中目标识别中的应用 [J]. 网络安全与数据治理, 2024, 43(4): 41-45, 52.

Application of deep learning model combined with batch normalization layer in underwater target recognition

Sun Yue, Peng Yuan, Jia Lianhui, Cao Lin, Guo Xinyu, Xu Jianqiu

(Science and Technology on Underwater Test and Control Laboratory, Dalian 116013, China)

Abstract: In view of the poor stability of deep learning in training underwater acoustic targets, resulting in poor classification and recognition performance, from the perspectives of local connectivity, spatial arrangement, and model design of the network, based on the original one-dimensional sequence convolution kernel and one-dimensional sequence pooling, this paper introduces batch normalization layer to build a deep learning network model. By normalizing, the goals of accelerating the convergence process of the network model and improving the stability during the training process are achieved. To verify the effectiveness of the model, network training and model validation are carried out on sample data of three types of underwater acoustic targets, which proves that the model also has a certain degree of performance improvement in improving the classification and recognition performance of underwater acoustic target data.

Key words: underwater acoustic target; deep learning; classification; network model

0 引言

传统识别方式是针对确定性的数据, 目标样本确定, 目标类型确定, 应用已有的水声特征提取方法, 对确定性的数据进行特征提取, 最终实现分类识别^[1]。然而, 由于数据值有误或缺失以及各个传感器的影响造成水声数据的不确定性, 影响了对水下目标的分类识别准确率。

再加上当前具备可分辨性的物理特征数量稀少, 受当前信号处理机制性能瓶颈限制, 多目标强干扰、低信噪比等条件下的目标特征提取与识别能力尚无法满足需求^[2]。

智能识别的出现革新了特征提取的模式, 小样本学习、迁移学习、深度学习等技术的迅猛发展, 为智能水声目标识别提供了新的机遇与挑战^[3]。深度学习是机器学习领域中的一种研究方法, 它通过构建具有多个层次结构的神经网络来模拟人类大脑的功能和思维过程。近年来, 随着计算机硬件性能的提升和大数据的普及, 深

^{*} 基金项目: 国防科技重点实验室基金项目 (2020 - JCJQ - LB - 027)

深度学习已经取得了众多重要的研究和应用成果,成为人工智能领域的热点之一。深度学习作为人工智能领域的一个子集被广泛地应用于水声目标识别领域。

本文从网络的局部连接、空间位置排列以及模型的设计角度出发,采用序列卷积与批规范化层的方式构造网络模型,通过对比实验得出添加了批规范化层的网络模型收敛速度更快、分类识别效果也更好。

1 深度学习网络模型

如图1所示,深度学习网络模型的设计主要从减少参数数量的必要性、生物学模拟的可行性以及提取深层特征的重要性三个方面考虑。

首先,采用局部连接的方式来达到减少参数量、缩短网络训练时间的目的^[4]。其次,局部连接的另一个作用是实现对网络神经元之间的局部感知,从生物学模拟的角度出发摒弃全局感知的影响,通过局部感知策略,提取某一部分的特征信息,然后将不同的特征进行组合,这样做既保留了网络的局部敏感性又达到了全局处理的效果^[5]。最后,对输入数据进行预处理和特征提取,通过这种端到端的学习方式,能更好地提取有效特征,从而提升网络性能^[6]。

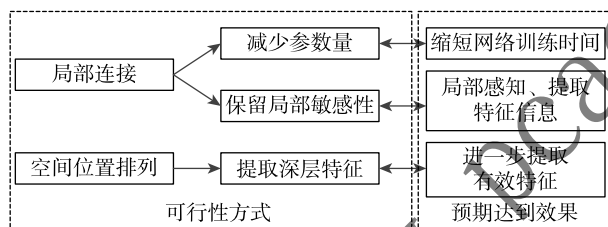


图1 深度学习网络模型设计

1.1 局部连接

当处理音频信号这种高维度数据时,采用全连接^[7]的方式会导致网络中的参数量非常多,训练的时间和空间代价都非常高昂^[8]。另外,从生物学模拟的可行性角度出发,深度学习网络中的每个神经元之间是没有必要对全局进行感知的,只需要对局部敏感^[9]。具体来说,构造卷积层时,每一个神经元只需要与上一层的部分神经元相连即可^[10]。

1.2 空间位置排列

对于水声信号来说,每一个卷积核能够提取某一部分的特征^[11]。在绝大多数情况下,单卷积核提取的特征是不充分的,这时就需要添加多个卷积核来提取多重特征^[12]。因此,通常来说,多个卷积核的效果要比单个卷积核的效果好^[13]。除此之外,还要指定卷积核在输入数据上的移动步幅,设步幅大小为 s ,表示卷积核每一步会

移动的步长为 s 。补零通常被用来对边界进行处理,由于卷积核的大小不一定能够被输入数据的维度大小所整除,因此,可能会出现卷积核不能完全覆盖边界元素的情况,故采取补零的策略以防止信息的丢失^[14]。

1.3 模型设计

深度学习网络模型在设计完成之后投入到训练过程中往往会出现难以训练的问题,具体表现为在模型训练期间,网络的每一次输入会随着前一层网络参数的变化而变化,因此在训练开始阶段必须使用很小的学习率以及对参数进行合理的初始化,这就导致训练过程变得很缓慢,而参数的随机性变化也导致了模型最终训练效果的不可控性。

引入批规范化层可达到加速网络模型的收敛过程以及提高训练过程中稳定性的目的,从而能在一定程度上提高模型的分类识别效果。

1.3.1 改进前网络模型结构设计

样本数据为一维数据,区别于二维图形数据处理,为适应数据形式,网络结构中的二维运算操作全部改成一维。具体为二维的 $n \times n$ 图形卷积核改成一维的序列卷积核;二维的 $n \times n$ 图形池化改成一维序列池化。网络模型结构1如图2所示。

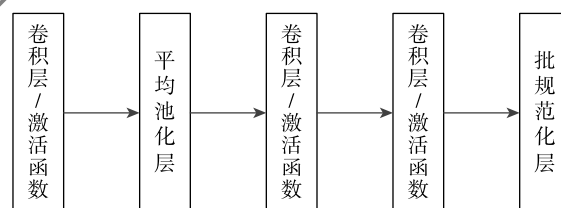


图2 模型结构1

使用初步改进的模型结构1结合不同的超参数组合进行训练。

(1) 学习率为0.003,批次大小为256,迭代次数为100次。网络训练过程中的损失函数变化曲线如图3所示。可以看出,随着迭代次数的增加,损失函数值逐渐降低,最后趋于稳定。准确率变化曲线如图4所示,随着迭代次数的增加,训练和测试集的识别率能达到98%左右,验证集的识别率为87.9%。

(2) 学习率为0.003,批次大小为128,迭代次数为100次。网络训练过程中的损失函数变化曲线如图5所示,可以看出,随着迭代次数的增加,损失函数值逐渐降低,最后趋于稳定。准确率变化曲线如图6所示,从图中可以看出,训练和测试集的识别率能达到98%左右,验证集的识别率为90.3%。

(3) 学习率为0.02,批次大小为256,迭代次数为

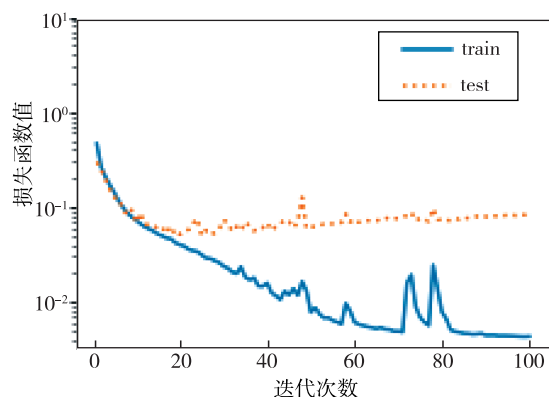


图3 损失函数曲线

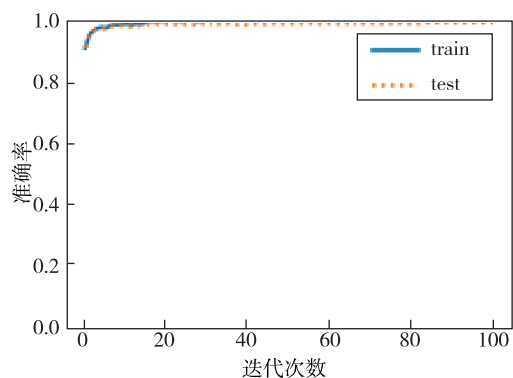


图6 准确率变化曲线

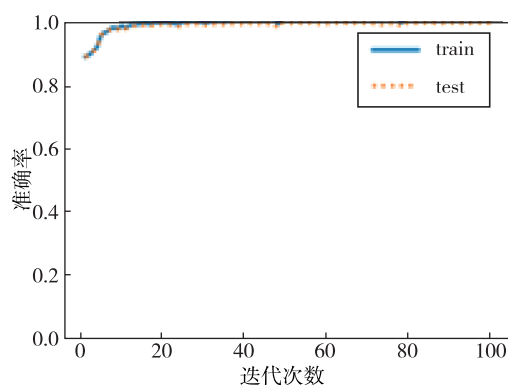


图4 准确率变化曲线

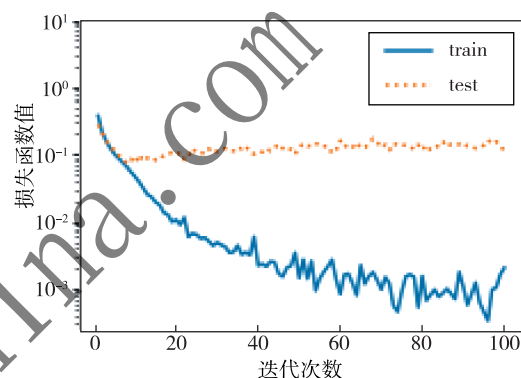


图7 损失函数曲线

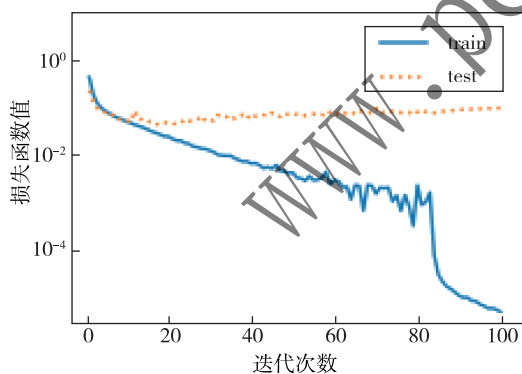


图5 损失函数曲线

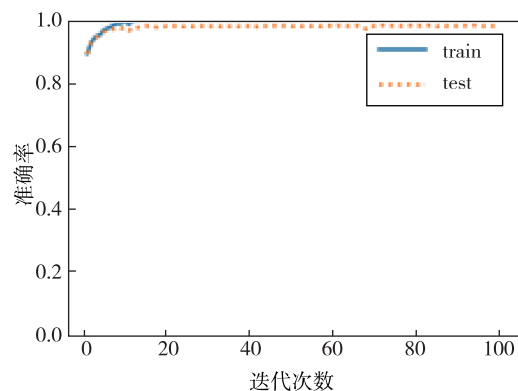


图8 准确率变化曲线

100次。网络训练过程中的损失函数变化曲线如图7所示,可以看出,随着迭代次数的增加,损失函数值逐渐降低,最后趋于稳定。准确率变化曲线如图8所示,从图中可以看出,训练和测试集的识别率能达到97%左右,验证集的识别率为90.3%。

从损失函数曲线和准确率变化曲线可以看出模型泛化能力不强,存在一定的过拟合现象。

1.3.2 引入批规范化层的网络模型结构设计

深度学习网络模型在进行训练时,由于每批训练数

据的分布各不相同,这就导致网络需要在每次迭代中去学习适应不同的分布,这会大大降低网络的训练速度。深度网络的训练是一个非常复杂的过程,如果网络的前面几层发生微小的改变,那么这些微小的改变在后面的层就会被累积放大下去。一旦网络某一层的输入数据的分布发生改变,那么该层网络就需要去适应学习新的数据分布,所以如果训练过程中训练数据的分布一直在发生变化,隐层的输入分布总是处于变化状态,会带来内部协变量转移问题,将会影响网络的训练速度。

批规范化层的设计主要分为以下 4 个步骤:

- (1) 求每一个训练批次数据的均值;
- (2) 求每一个训练批次数据的方差;
- (3) 使用求得的均值和方差对该批次的训练数据做归一化, 获得 0-1 分布。
- (4) 尺度变换和偏移。

因此通过引入批规范化层, 使得每个神经元的输入数据“独立同分布”。当数据被同时输入网络中一起进行前向计算时, 该批次中所有数据的误差也将累计起来

一起回传, 添加批规范化层后将对一个批次中的数据做归一化处理, 上层参数不再需要不断适应新的输入数据分布, 从而提高学习速度; 下层输入的变化不再趋向于变大或者变小, 避免上层落入饱和区, 使得学习过程能够正常进行, 达到加速网络模型的收敛过程以及提高训练过程中的稳定性的目的。改进后的模型结构如图 9 所示。结合批规范化层的深度学习模型识别流程如图 10 所示。结合批规范化层的深度学习模型结构如表 1 所示。

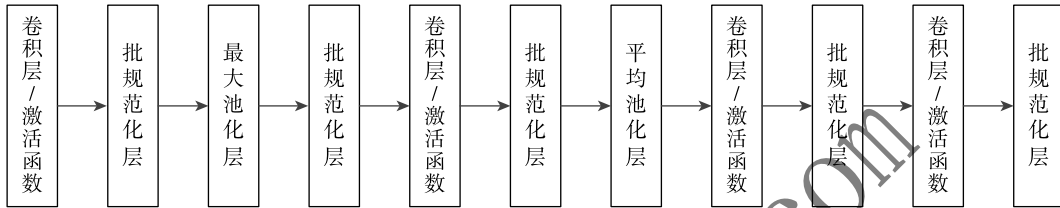


图 9 改进后的模型结构

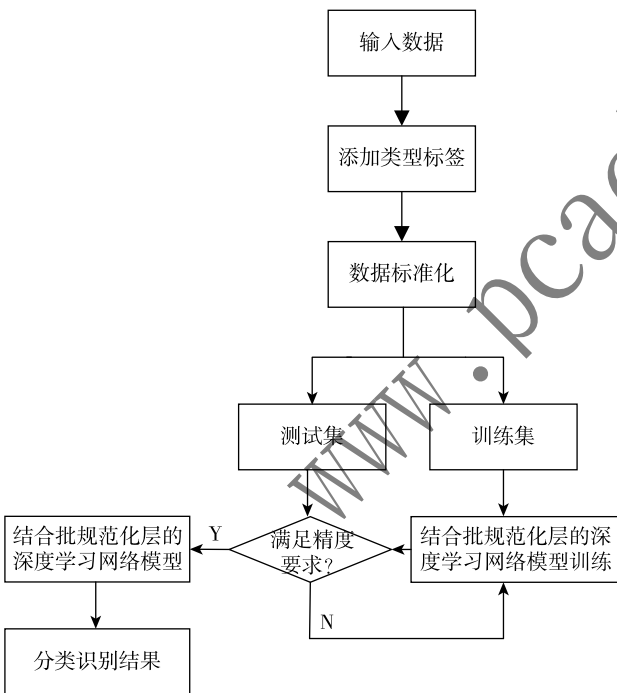


图 10 结合批规范化层的深度学习模型识别流程图

表 1 结合批规范化层的深度学习网络模型结构

网络层	通道	卷积核大小	优化参数
Conv1D0	16	9	0.9
Conv1D1	32	9	
Conv1D2	64	9	
BN			0.9
MaxPool1D		4	
BN			
Conv1D3	128	3	0.9
Conv1D4	256	3	
Conv1D5	256	3	
BN			0.9
GlobalAvgPool1D		4	
Conv1D6	512	3	
Conv1D7	512	3	0.9
Conv1D8	512	3	
BN			
Conv1D9	4 096	3	0.9
Conv1D10	4 096	3	
BN			
Dropout			0.5
Flatten	3		

表 2 数据样本划分

水声目标分类	训练集/条	测试集/条
A 类	1 655	1 664
B 类	1 684	1 684
C 类	1 858	1 855

2 实验验证

为了验证基于深度学习的水声目标分类识别方法的可行性, 本文收集了 3 类水声目标样本集。训练和测试数据样本划分如表 2 所示。

2.1 改进前模型实验验证

改进前模型在不同参数条件下的训练结果如表 3 所示, 最终在测试集上的分类识别效果如表 4 所示。

表3 不同参数条件下的模型训练结果

学习率	批量大小	动量	训练集准确率/%
0.05	256	0.9	98.9
0.03	256	0.9	99.9
0.01	256	0.9	99.9
0.009	256	0.9	99.1
0.005	256	0.9	99.9
0.003	256	0.9	98.3
0.001	256	0.9	98.2
0.000 3	256	0.9	97.1
0.000 1	256	0.9	90.1
0.000 03	256	0.9	89.6
0.000 01	256	0.9	81.4

表4 分类识别测试结果

水声目标分类	测试样本个数	分类识别正确个数	识别准确率/%
A类	1 664	1 236	74.3
B类	1 684	1 411	83.8
C类	1 855	1 526	82.3

改进前的网络模型泛化能力差,在训练集上表现较好,而在测试集上表现较差。

2.2 改进后模型实验验证

采用前期学习率为0.05,使损失函数曲线快速收敛到局部最优附近,然后降低学习率为0.000 3,使损失函数值缓慢下降。改进后的模型训练完毕,准确率变化曲线如图11所示,最终在测试集上的分类识别效果如表5所示。

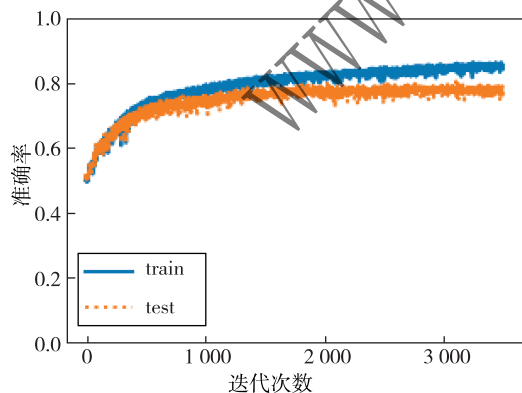


图11 准确率变化曲线

表5 分类识别测试结果

水声目标分类	测试样本个数	分类识别正确个数	识别准确率/%
A类	1 664	1 539	92.5
B类	1 684	1 520	90.3
C类	1 855	1 820	98.1

3 结论

本文提出了一种基于深度学习的水声目标分类识别方法,针对3类水声目标进行了实验验证,并给出了模型在改进之前的识别准确率,以及在原有网络模型的基础上添加批规范化层后的分类识别效果,可以看出分类识别效果有一定的提升。但是从深度学习网络模型的实际应用角度出发,目前的水声数据的数据量还远远不够支撑模型的大规模训练,因此需要进行数据扩增处理,除了从数据的角度出发以外,考虑如何将多源水声信号特征数据与深度神经网络模型进行融合来提升模型的分类识别性能,也是今后需要研究的重点方向。

参考文献

- [1] 张巧力,刘福臣.基于前馈神经网络的水下目标被动定位方法研究[C]//全国声学大会论文集,2019:211-212.
- [2] 刘峰,罗再磊,沈国圣,等.时频谱图和数据增强的水声信号深度学习目标识别方法[J].应用声学,2021,40(4):518-524.
- [3] 王升贵,胡桥,陈迎亮,等.基于深度学习的水下目标识别方法研究[J].舰船科学技术,2020,42(23):141-145.
- [4] ZHANG S, TONG H, XU J, et al. Graph convolutional networks: a comprehensive review [J]. Computational Social Networks, 2019, 6 (1): 1-23.
- [5] LUO Z, HOU T, ZHOU X, et al. Binary code similarity detection through LSTM and siamese neural network [J]. EAI Endorsed Transactions on Security and Safety, 2021, 8 (29): 11-20.
- [6] YANG S, CHENG L, ZENG Y, et al. Asteria: deep learning based AST-encoding for cross-platform binary code similarity detection [C]// 2021 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN). IEEE, 2021: 224-236.
- [7] SRIVASTAVA N, HINTON G, KRIZHEVSKY A. Dropout: a simple way to prevent neural networks from overfitting [J]. Journal of Machine Learning Research, 2014, 15 (1): 1929-1958.
- [8] GIDARIS S, KOMODAKIS N. Generating classification weights with GNN denoising autoencoders for few-shot learning [C]// Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2019: 21-30.
- [9] SHELL J, SWERSK K, ZEMEL R. Prototypical networks for few-shot learning [J]. Advances in Neural Information Processing Systems, 2017: 4078-4088.
- [10] FINN C, ABBEEL P, LEVINE S. Model-agnostic meta-learning for fast adaptation of deep networks [C]// International Conference on Machine Learning, 2017: 1126-1135.
- [11] RADFORD A, METZ L, CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks [J]. Computer Science, 2015: 127-132.

(下转第52页)

3 结论

为了在复杂背景下对小尺寸多角度人脸进行更好的检测,同时能够更容易地实现在边缘设备上部署模型,本文提出了一种新的神经网络模型(GhostNet-MTCNN),该模型以MTCNN为基础,利用Ghost Bottleneck模组重新构建特征提取网络。通过与原始MTCNN和MobileNetV3-MTCNN的对比实验以及模型消融实验证明了所提模型可以在复杂场景下较好地实现小尺寸人脸的检测,同时也能够有效平衡参数量和精度使其成为边缘设备部署的更优选择。经过以上改进可以在人流量较多的场景下取得不错的检测效果。如何将模型部署到边缘设备是接下来可以研究的方向。

参考文献

[1] ROWLEY H A, BALUJA S, KANADE T. Neural network-based face detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20 (1): 23 – 38.

[2] VIOLA P, JONES M J. Robust real-time face detection [J]. International Journal of Computer Vision, 2004, 57: 137 – 154.

[3] ZHAO Z. Application of improved CNN-based face detection technology in public administration [J]. Journal of Computational Methods in Sciences and Engineering, 2023, 23 (4): 1985 – 1997.

[4] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580 – 587.

[5] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39 (6): 1137 – 1149.

[6] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look

once: unified, real-time object detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779 – 788.

[7] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7263 – 7271.

[8] REDMON J, FARHADI A. Yolov3: an incremental improvement [J]. arXiv preprint arXiv: 1804. 02767, 2018.

[9] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector [C]//European Conference on Computer Vision. Springer, Cham, 2016: 21 – 37.

[10] ZHANG K P, ZHANG Z P, LI Z F, et al. Joint face detection and alignment using multitask cascaded convolutional networks [J]. IEEE Signal Processing Letters, 2016, 23 (10): 1499 – 1503.

[11] KU H, DONG W. Face recognition based on MTCNN and convolutional neural network [J]. Frontiers in Signal Processing, 2020, 4 (1): 37 – 42.

[12] 陈政生. 基于深度学习的口罩佩戴检测方法研究 [D]. 武汉: 华中师范大学, 2021.

[13] HAN K, WANG Y H, TIAN Q, et al. Ghostnet: more features from cheap operations [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580 – 1589.

(收稿日期: 2024 – 01 – 08)

作者简介:

黄杰 (1999 –), 男, 硕士研究生, 主要研究方向: 目标检测、深度学习、边缘部署。

刘芬 (1979 –), 通信作者, 女, 硕士, 副教授, 主要研究方向: 微波电路设计、数字信号处理、人工智能。E-mail: liufute@126.com。

(上接第45页)

[12] ADAM S, MATTHEW B. One-shot learning with memory-augmented neural networks [J]. IEEE Transactions on Software Engineering, 2022, 49 (4): 1661 – 1682.

[13] MUNKHDALAI T, YU H. Meta networks [C]//International Conference on Machine Learning, 2017: 2554 – 2563.

[14] XU L, CHOY C S, LI Y W. Deep sparse rectifier neural networks for speech denoising [C]//2016 IEEE International Workshop on Acoustic Signal Enhancement (IWAENG), 2016: 26 – 33.

(收稿日期: 2023 – 09 – 21)

作者简介:

孙悦 (1994 –), 通信作者, 女, 硕士, 工程师, 主要研究方向: 模式识别。E-mail: 1286857718@qq.com。

彭圆 (1972 –), 女, 博士, 研究员, 主要研究方向: 模式识别。

贾连徽 (1985 –), 男, 博士, 高级工程师, 主要研究方向: 目标特性。

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部

www.pcachina.com