

基于改进 CapsNet 的色素性皮肤病识别的研究*

李励泽, 张晨洁, 杨晓慧, 孙文斌, 郭 滨

(长春理工大学 电子信息工程学院, 吉林 长春 130022)

摘要: 皮肤病是医学上的常见的、多发性疾病, 因此皮肤检测技术越来越受关注。卷积神经网络是常见的皮肤检测方法, 其模型结构会丢失很多信息。CapsNet(胶囊网络)是在卷积神经网络之后的一种新的神经网络。CapsNet 的矢量化特征能够较好地表达空间关联性, 每一个 capsule(胶囊)独立地服务各自的任务。分析了 CapsNet 的基本结构和主要算法, 改进了网络模型从而避免过拟合现象发生, 试图基于改进 CapsNet 针对预处理之后的皮肤图像进行识别, 并与传统卷积神经网络的模型作对比。实验结果表明, 使用改进 CapsNet 对色素性皮肤病进行识别可以有较好的识别效果, 并且准确率比传统方法高出 8%~10%。

关键词: 色素性皮肤病; 皮肤图像识别; 神经网络; 胶囊网络

中图分类号: TN98; TP391.4

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.200121

中文引用格式: 李励泽, 张晨洁, 杨晓慧, 等. 基于改进 CapsNet 的色素性皮肤病识别的研究[J]. 电子技术应用, 2020, 46(11): 60-64.

英文引用格式: Li Lize, Zhang Chenjie, Yang Xiaohui, et al. Pigmented skin lesion recognition based on improved CapsNet[J]. Application of Electronic Technique, 2020, 46(11): 60-64.

Pigmented skin lesion recognition based on improved CapsNet

Li Lize, Zhang Chenjie, Yang Xiaohui, Sun Wenbin, Guo Bin

(School of Electronics and Information Engineering, Changchun University of Science and Technology, Changchun 130022, China)

Abstract: Dermatitis is a common and multiple disease in medicine, so skin detection technology has attracted more and more attention. Convolutional neural network is a common skin detection method, and its model structure will lose a lot of information. CapsNet is a new kind of neural network after convolutional neural network. The vectorization of CapsNet can better express the spatial relevance, with each capsule serving its own mission independently. This paper analyzed the basic structure and main algorithm of CapsNet, the network model was improved to avoid over fitting, and tried to identify the pre-processed skin image based on improved CapsNet, and compared it with the model of traditional convolutional neural network. Experimental results show that improved CapsNet can be used to identify pigmented skin diseases with good effect, and the accuracy is about 8~10 percent higher than the traditional method.

Key words: pigmented skin lesion; skin image recognition; neural network; CapsNet

0 引言

皮肤在身体上属于最大的器官, 是对身体的第一层保护。据相关调查报告显示, 存在与皮肤相关的身体健康问题的人类在全世界占据了 30%~70%^[1]。色素性皮肤病的形成是因为黑素和黑素细胞出现异常, 皮肤的颜色变化与病灶的而改变息息相关, 因此识别难度较大。自从 1987 年对色素性皮肤病的分类研究^[2]出现之后, 陆续出现了大量的皮肤病图像诊断技术的研究成果^[3-4]。在文献[5]的算法研究中, 首先将图像进行预处理, 然后将利用卷积神经网络提取出的人脸特征进行训练、分类。文献[6]利用深度卷积神经网络和支持向量机对皮

肤图像进行分类和识别。文献[7]利用卷积网络的 VGG16 模型, 采取 3 种不同的方法来实现皮肤镜二分法, 其结果显著提升。在皮肤病诊断领域的算法研究较为复杂, 而且国外的研究要比国内相对较多, 如何用简单的算法结构完成精确的识别是皮肤病检测研究的主要问题。

在进行人脸图像检测时, 传统的卷积神经网络只要能够检测出耳朵、眼睛、鼻子、嘴巴便认作为人的面部, 此方式忽略了这些实例的空间位置信息, 可能会导致一定的识别错误。目前 CapsNet 技术已经比较成熟, 其应用也比较广泛。文献[8]中应用 CapsNet 在手语识别领域已经达到了令人满意的准确率。LI Y 等人使用 CapsNet 识别无人飞行器拍摄的水稻图像, 既可以预防害虫与疾病, 又能实时观察水稻^[9]。AFSHAR Y 等人采用 CapsNet

* 基金项目: 吉林省科技厅项目(20200404216YY)

对在脑部的 NMR 中的脑肿瘤类型进行了分类研究,其预测精度比传统方法更高^[10]。文献[11]将 CapsNet 用于指静脉识别算法中,最终得到的精确度比 VGG 方法高出了 13.6%。本文将使用胶囊网络对皮肤图像进行处理,减少算法叠加的重复性,为皮肤图像处理领域打开新的大门,为后续的皮肤病诊断工作提供良好的基础。

1 CapsNet 模型

本文档胶囊神经网络由 HINTON G E 等人提出,研究者把具有相同特点的神经元捆绑成一组,称之为神经胶囊,胶囊网络的模型结构如图 1 所示^[12],其中包含两个卷积层(Conv)和一个全连接(FC)层。在胶囊网络中,采用胶囊的方法来代替传统的神经元,并提出了一种动态路由的算法,推进深度学习的预判能力。

(1) Conv1 层

该层是 Capsule 网络的第一个卷积层,其主要作用是对图像像素进行一个局部的特征检测,也就是图像输入到低级特征的一步常规卷积操作。经过卷积后,可以发现张量的维度尺寸已经发生了变化。维度变化的计算公式如下:

$$H_{out} = \left\lfloor \frac{H_{in} - H_{ck}}{H_{stride}} \right\rfloor + 1 \quad (1)$$

$$W_{out} = \left\lfloor \frac{W_{in} - W_{ck}}{W_{stride}} \right\rfloor + 1 \quad (2)$$

式中, H_{out} 表示图像的输出高度, H_{in} 表示图像的输入高度, H_{ck} 表示卷积核的高度, H_{stride} 表示垂直维度步长的大小; W_{out} 表示图像的输出宽度值, W_{in} 表示图像的输入宽度值, W_{ck} 表示卷积核的宽度值, W_{stride} 表示水平维度步长的大小。

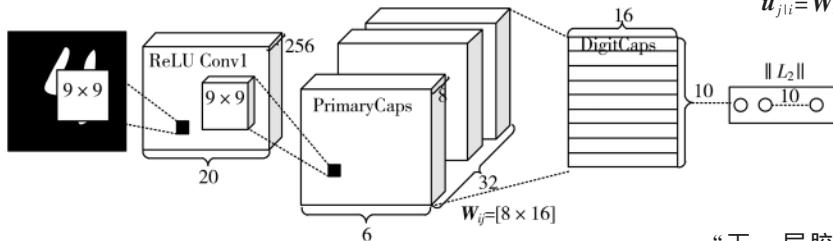


图 1 胶囊网络的模型结构

(2) Conv2 层

Conv2 层又称 PrimaryCaps 层,这一层为特有的卷积层,主要存储低级别特征的向量。与 CNN 比较有差异的是,在 Capsule 中,维度为 $6 \times 6 \times 8 \times 32$ 的层里有 $6 \times 6 \times 32$ 个元素,每个胶囊是一个 1×8 的向量。

(3) DigitCaps 层

DigitCaps 层储存高级别特征的向量,其包含两类参数。其中一类是与 PrimaryCaps 层之间的权值向量,另一类是耦合系数,该系数决定了该层各个胶囊与上一层各个胶囊之间的密切程度,这部分的参数取决于由动态路由(Dynamic Routing)算法。

DigitCaps 中胶囊输出向量的 L2 范数(向量大小)决定了该网络的识别结果。如果一个向量很大,那么这个

向量的范数也很大,并且这个向量对应的数出现的可能性也很大。

2 算法研究

2.1 胶囊网络工作原理

胶囊网络中最重要的部分就是胶囊,它能够表征检测类型的多维实体的实例化参数和存在性。胶囊的工作原理如图 2 所示。

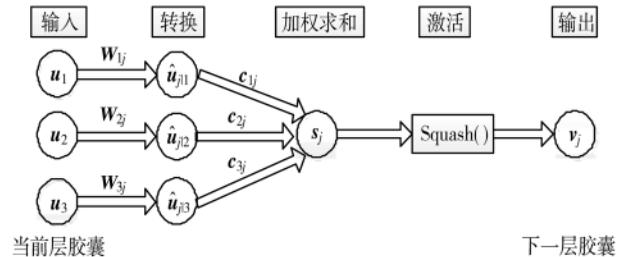


图 2 胶囊工作原理示意图

图 2 表达了胶囊网络中两层胶囊之间转换关系,其中 s_j 为胶囊的总体输入。对于除胶囊的第一层以外的所有层,向量 s_j 的总输入为来自“下一层胶囊”中所有“预测向量” \hat{u}_{jli} 的加权总和,每个胶囊在传入下一个胶囊之前,都要乘以一个权值共享矩阵 W_{ij} 。相邻胶囊层之间的特征信息由权值共享矩阵传递。将低层所有胶囊对高层胶囊的预测向量乘权重 c_{ij} 相加得到 s_j 。向量的输入分为以下两个阶段:

$$s_j = \sum_i c_{ij} \hat{u}_{jli} \quad (3)$$

$$\hat{u}_{jli} = W_{ij} u_i \quad (4)$$

其中,迭代动态路由过程确定的 c_{ij} 为耦合系数,且 \hat{u}_{jli} 是由低级别的输出与权重矩阵的乘积得出。“当前层胶囊”与上述层中的所有胶囊之间的耦合系数之和为 1,由路由最大值“routing softmax”决定。通过了解

“下一层胶囊”的输出与“当前层胶囊”做出的预测 \hat{u}_{jli} 的一致程度,迭代地优化起始的耦合系数。

$$c_{ij} = \frac{e^{b_{ij}}}{\sum_k e^{b_{ik}}} \quad (5)$$

其中, b_{ij} 为当前层的输出向量与上一层的预测向量的相似程度,通过这种方式实现一层一层迭代更新得到耦合系数 c_{ij} 的值。为了保证实体出现的概率在 0~1,采用一个 Squashing 函数(非线性激活函数)得到输出。本文使用 0.5 替代 HINTON G E 公式中的 1,这一改进之处可保证小于 0.5 的向量范数可以被压缩,比 0.5 大的向量范数可以被扩大。

$$v_j = \frac{\|s_j\|^2}{0.5 + \|s_j\|^2} \frac{s_j}{\|s_j\|} \quad (6)$$

式(6)的前一部分表示对所有函数进行压缩,后一部

分表示在向量的方向不发生改变的情况下进行单位化。改进之后的向量输出符合向量范数正则化的曲线拟合原理。模值越大,概率越大,这样可以提高个体识别准确率^[13],避免所有向量都被压缩而产生的过拟合现象,保证在规则化参数的同时最小化误差。

2.2 动态路由算法实现

算法 1: CapsNet 中的动态路由迭代算法

(1) 令: $\mathbf{b}_j \leftarrow 0 (j \in l+1)$

(2) for r iterations do

(3) $\mathbf{c}_i \leftarrow \text{softmax}(l+1)(\mathbf{b}_j) (i \in l)$

// 执对 l 层的 \mathbf{VN}_i , 将 \mathbf{b}_j 用 softmax 转化成概率 \mathbf{c}_j

(4) $\mathbf{s}_j \leftarrow \sum_i \mathbf{c}_i \hat{\mathbf{u}}_{j|i} (j \in l+1)$

(5) $\mathbf{v}_j \leftarrow \text{squash}(\mathbf{s}_j) (j \in l+1)$

// 对 $l+1$ 层的 \mathbf{VN}_j , 压缩 \mathbf{s}_j 得到 \mathbf{v}_j

(6) $\mathbf{b}_j \leftarrow \mathbf{b}_j + \hat{\mathbf{u}}_{j|i} \cdot \mathbf{v}_j (i \in l, j \in l+1)$

// 根据 $\mathbf{u}_{j|i}$ 和 \mathbf{v}_j 的关系来更新 \mathbf{b}_j

(7) Return \mathbf{v}_j

r 表示迭代次数,通常为 3 次。一般来说,输出为输入的一个聚类的结果,而这个聚类的过程需要经过迭代算法的实现,这个迭代算法被称为“动态路由”。经过 softmax 函数回归得到的皆为非负数且总和为 1,所以 \mathbf{c}_j 属于一组概率变量, squash 可以保证 \mathbf{s}_j 的方向不发生变化,其长度永远比 1 小。因为这代表 \mathbf{VN} (向量神经元)具有给定特征的概率。第 6 行是此算法的重点所在,用 $\mathbf{u}_{j|i}$ 和 \mathbf{v}_j 的点积(dot product)更新 \mathbf{b}_j ,其中前者是 l 层 \mathbf{VN}_i 对 $l+1$ 层 \mathbf{VN}_j 的“个人”预测,而后者是所有 l 层 \mathbf{VN} 对 $l+1$ 层 \mathbf{VN}_j 的“共识”预测;当两者相似, \mathbf{b}_j 随点积的变大而变大,低级别 \mathbf{VN}_j 与高级别 \mathbf{VN}_j 的相连的可能性就变大;当两者相异,则三者都变小。

softmax 回归本身可以定位为对一个具体问题的预测。定义一个神经网络的输出节点为预测值,这里使用的损失函数(loss 函数)为均方误差(MSE):

$$\text{MSE}(y, y') = \frac{\sum_{i=1}^n (y_i - y'_i)^2}{n} \quad (7)$$

其中, y_i 为一个 batch 中的第 i 个数据的正确结果, y'_i 是神经网络提供的一个预测的值。损失函数对预期的结果考量,损失函数的值越小,神经网络的训练效果越好。

3 实验及结果分析

3.1 皮肤病数据集

临床上,黑色素细胞和黑色素出现不正常的构成,是导致色素性皮肤病的主要原因,色素的变化会导致外部皮肤明显的颜色加深,这使得色素性皮肤病的识别处理更加不容易。本文研究使用了国际皮肤图像合作组织(International Skin Imaging Collaboration)组织的 2017 年黑色素瘤检测挑战进行的皮肤镜图像的数据集,此数据集

包括 3 个独立的皮肤镜图像数据集:2 000 张带有标注的训练集、150 张带有标注的验证集(包含 30 张恶性黑色素瘤样本)和 600 张带有标注的测试集(其中恶性黑色素瘤样本为 117)。如图 3 所示,每份数据集有黑色素瘤、色素痣和脂溢性角化病。图像分布如表 1 所示。



图 3 3 类色素性皮肤病

表 1 3 种类别图像分布

数据集	黑色素瘤	色素痣	脂溢性角化病	合计
训练集	374	1 372	254	2 000
验证集	30	78	42	150
测试集	117	393	90	600

根据在医学诊断中和 ISIC2017 中的建议,除了采用识别准确率(acc)以外,本文也将损失函数值作为主要的评价指标。准确率指在分类中,使用测试集对模型进行分类,分类正确的记录个数占总记录个数的比例,其公式如下:

$$\text{acc} = \frac{N_{tp} + N_{tn}}{N_{tp} + N_{fp} + N_{tn} + N_{fn}} \quad (8)$$

其中, N_{tp} 表示真阳性像素点的个数, N_{tn} 表示真阴性像素点的个数, N_{fp} 表示假阳性像素点的个数, N_{fn} 表示假阴性像素点的个数。当准确率取值较大时,说明模型的识别性较好。

3.2 实验与结果

在 TensorFlow 框架下,用 GPU 对训练过程加速,在 CapsNet 训练过程中,通过验证集评估模型的预测水平,并在整个过程中进行监控,不断改变参数数值使模型达到最优,分别输出边缘损失值 margin-loss(如图 4 所示)、重构损失值 reconstruction-loss(如图 5 所示)和总损失值 total-loss(如图 6 所示)3 组损失值。总损失值为边缘损失值与一定比例的重构损失值之和。随着训练次数的增加,振动趋于稳定。

将实验数据按比例分为两部分:一部分用于网络

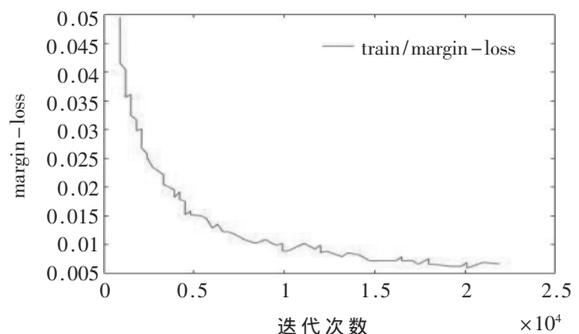


图 4 边缘损失值

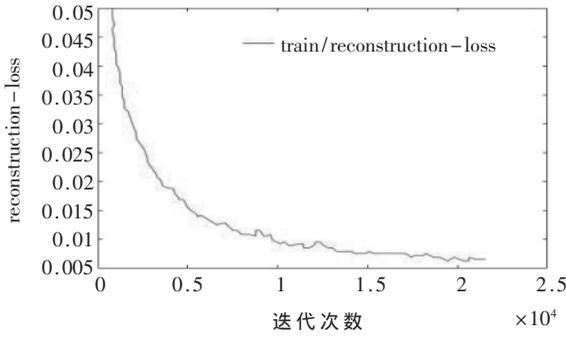


图5 重构损失值

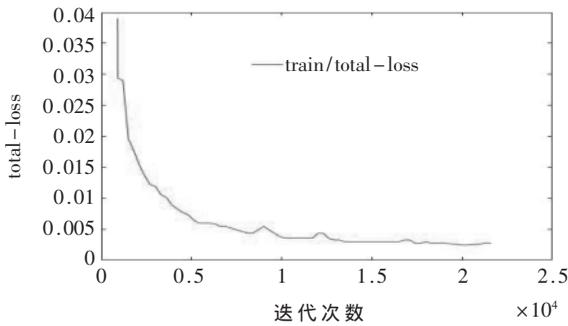


图6 总损失值

训练,另一部分用于网络测试,其实验最终训练结果如图7所示。该实验条件下,正确地识别皮肤图像为黑素细胞痣,识别结果如图8所示,图中横纵坐标均表示图像的像素大小。在此实验条件下,CapsNet模型的识别准确率达到82.24%。

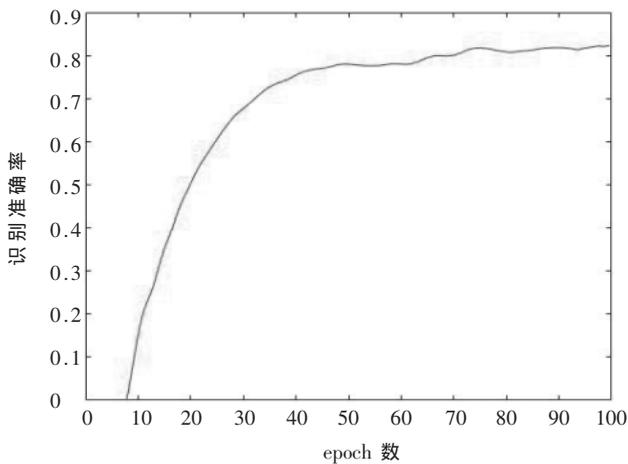


图7 训练结果

为了验证本文模型的准确性,基于ISIC2017的数据集,选择卷积神经网络的ResNet、GoogLeNet+、VGG+网络进行对比实验^[14],表示训练中的权重值。本文方法与其他3种深度学习模型实验对比如图9所示。本文模型的识别结果优于其他模型,证明此方法的准确性和可靠性。另外,几种方法的识别率均不高,表明了几种色素性

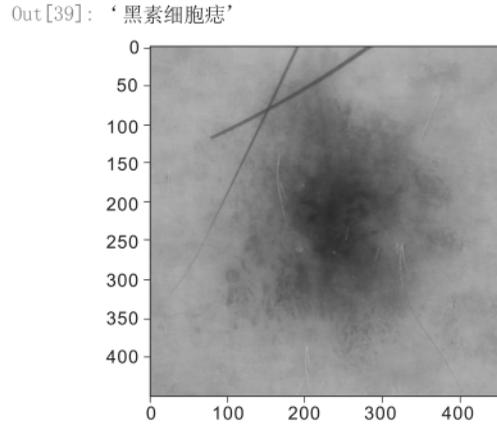


图8 皮肤病识别结果

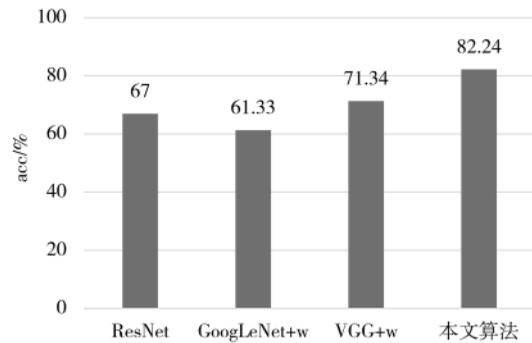


图9 不同方法下的识别率对比

皮肤病较难区分,识别难度较大。

4 结论

本文主要研究了对色素性皮肤病的分类识别方法,训练了一个经过优化的胶囊神经网络结构,开拓性地将CapsNet应用于皮肤检测领域,实现了对色素性皮肤病的识别处理。本文采用的深度学习模型在保证达到较高的准确率下,也避开了传统方法的局限性。实验表明,目前在深度学习方法中CapsNet在图像处理方面具有很强的潜在能力,并且拥有广泛的应用前景,在一定程度上满足了较高要求的领域需求。

参考文献

- [1] HAY R J,JOHNS N E,WILLIAMS H C,et al.The global burden of skin disease in 2010:an analysis of the prevalence and impact of skin conditions[J].Journal of Investigative Dermatology,2014,134(6):1527-1534.
- [2] CASCINELLI N,FERRARIO M,TONELLI T,et al.A possible new tool for clinical diagnosis of melanoma:the computer[J].Journal of the American Academy of Dermatology,1987,16(2):361-367.
- [3] STANGANELLI I,BRUCALE A,CALORI L,et al.Computer-aided diagnosis of melanocytic lesions[J].Anticancer Research,2005,25(6C):4577-4582.
- [4] RAZMJOOY N,SHEYKHAHMAD F R,GHADIMI N.A hybrid neural network-world cup optimization algorithm for

- melanoma detection[J].Open Medicine, 2018, 13: 9-16.
- [5] 黄海新, 张东. 基于深度学习的人脸活体检测算法[J]. 电子技术应用, 2019, 45(8): 44-47.
- [6] HAMEED N, SHABUT A M, HOSSAIN M A. Multi-class skin diseases classification using deep convolutional neural network and support vector machine[C]. 2018 12th International Conference on Software, Knowledge, Information Management and Applications, 2018: 1-7.
- [7] LOPEZ A R, GIRO-NIETO X, BURDICK J, et al. Skin lesion classification from dermoscopic images using deep learning techniques[C]. Proceeding of the 2017 13th IASTED International Conference on Biomedical Engineer. Piscataway, NJ: IEEE, 2017: 49-54.
- [8] 郝子煜, 阿里甫·库尔班, 李晓红, 等. 基于 CapsNet 的中国手指语识别[J]. 计算机应用研究, 2019, 36(10): 3157-3159.
- [9] LI Y, QIAN M, LIU P, et al. The recognition of rice images by UAV based on capsule network[J]. Cluster Computing, 2019, 22: 9515-9524.
- [10] AFSHAR P, MOHAMMADI A, PLATANIOTIS K N. Brain tumor type classification via capsule networks[J]. arXiv preprint, arXiv: 1802.10200, 2018.
- [11] 余成波, 熊递恩. 基于胶囊网络的指静脉识别研究[J]. 电子技术应用, 2018, 44(10): 15-18.
- [12] SABOUR S, FROSST N, HINTON G E. Dynamic routing between capsules[C]. Advances in Neural Information Processing Systems, 2017: 3856-3866.
- [13] 林少丹, 洪朝群, 陈雨雪. 结合胶囊网络和卷积神经网络的目标识别模型[J]. 电讯技术, 2019, 59(9): 987-994.
- [14] 何雪英, 韩忠义, 魏本征. 基于深度卷积神经网络的色素性皮肤病识别分类[J]. 计算机应用, 2018, 38(11): 3236-3240.

(收稿日期: 2020-02-22)

作者简介:

李励泽(1995-), 女, 硕士研究生, 主要研究方向: 电子电路与系统设计、智能信息处理。

张晨洁(1983-), 女, 博士研究生, 主要研究方向: 智能信息处理。

杨晓慧(1963-), 通信作者, 女, 教授, 主要研究方向: 电路与系统, E-mail: yangxiaohui1963@163.com。



(上接第 59 页)

- on Audio, Speech, and Language Processing, 2006, 14(4): 1218-1234.
- [3] LIM L, OPPENHEIM A V. Enhancement and bandwidth compression of noisy speech[J]. Proceedings of the IEEE, 1979, 67(12): 1586-1604.
- [4] 王晶, 傅丰林, 张运伟. 语音增强算法综述[J]. 声学与电子工程, 2005(1): 22-26.
- [5] 何玉文, 鲍长春, 夏丙寅, 等. 基于 AR-HMM 在线能量调整的语音增强方法[J]. 电子学报, 2014, 42(10): 1991-1997.
- [6] 徐勇. 基于深度神经网络的语音增强方法研究[D]. 合肥: 中国科学技术大学, 2015.
- [7] WANG Y, CHEN J, WANG D L. Deep neural network based supervised speech segregation generalizes to novel noises through large-scale training[R]. Ohio State University Columbus, 2015.
- [8] CHEN J, WANG Y, YOHO S E, et al. Largescale training to increase speech intelligibility for hearingimpaired listeners in novel noises[J]. The Journal of the Acoustical Society of America, 2016, 139(5): 2604-2612.
- [9] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative adversarial nets[C]. Proceedings of Advances in Neural Information Processing Systems. US: NIPS, 2014: 2672-2680.
- [10] MEHDI M, SIMON O. Conditional generative adversarial nets[J]. arXiv: 1411.1784, 2014.
- [11] MARTIN A, SOUMITH C, LÉON B, et al. Wasserstein GAN[J]. arXiv: 1701.07875, 2017.
- [12] SANTIAGO P, ANTONIO B, JOAN S, et al. SEGAN: speech enhancement generative adversarial network[J]. arXiv: 1703.09452v1, 2017.
- [13] ANDERS B L L, OLE W. Autoencoding beyond pixels using a learned similarity metric[J]. arXiv: 1512.09300v2, 2016.
- [14] PIERRE B. Autoencoders, unsupervised learning, and deep architectures[C]. JMLR: Workshop and Conference Proceedings, 2012, 27: 37-50.
- [15] Hu Guoning. PNL100 nonspeech sounds[OL]. [2020-04-22]. http://web.cse.ohio-state.edu/pnl/corpus/HuNonspeech/HuCorpus.html.
- [16] GAROFOLO J S, LAMEL L F, FISHER W M, et al. TIMIT acoustic-phonetic continuous speech corpus[C]. Philadelphia: Linguistic Data Consortium, 1993.
- [17] VARGA A, STEENEKEN H J M. Assessment for automatic speech recognition: II. NOISEX-92: a database and an experiment to study the effect of additive noise on speech recognition systems[J]. Speech Communication, 1993, 12(3): 247-251.

(收稿日期: 2020-04-22)

作者简介:

夏鼎(1999-), 男, 本科, 主要研究方向: 机器学习。

徐文涛(1989-), 男, 博士, 讲师, 主要研究方向: 信号处理、机器学习。

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《电子技术应用》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）、DOAJ、美国《乌利希期刊指南》、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！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所