

◇ 研究报告 ◇

# 结合改进DRSE-GCNN的电力调度语声识别模型\*

苕文涵<sup>†</sup> 张云翔 顾彬 相增辉 陈轩 李霁轩

(国网江苏省电力有限公司信息通信分公司 南京 210024)

**摘要:** 针对现有电力调度中语声识别方法存在的字识别错误率高和耗时长等问题,在分析语声识别技术的基础上,提出了一种改进的深度学习方法用于电力调度语声识别。将改进的深度残差收缩网络和改进的门控卷积神经网络相结合,通过改进的深度残差收缩网络提取有效特征,在通过堆叠改进的门控卷积神经网络来获取有效上下文信息。通过试验对所提方法的性能进行分析,验证其优越性。结果表明,所提方法与常规识别方法相比,在模型参数、字识别错误率和平均识别时间上均具有一定的优势,模型参数量为 6.48 M,字识别错误率为 2.87%,平均识别时间为 0.187 s。该研究为电力调度语言识别方法的发展提供了一定的参考。

**关键词:** 电力调度;语言识别;深度残差收缩网络;门控卷积神经网络;字识别错误率

中图分类号: TM933

文献标识码: A

文章编号: 1000-310X(2024)06-1243-07

DOI: 10.11684/j.issn.1000-310X.2024.06.007

## Combining improved DRSE-GCNN for power dispatching voice recognition model

CHANG Wenhan ZHANG Yunxiang GU Bin XIANG Zenghui  
CHEN Xuan LI Jixuan

(Information and Communication Branch of State Grid Jiangsu Electric Power Co., Ltd., Nanjing 210024, China)

**Abstract:** In order to solve the problems of high error rate and long time consuming in speech recognition in power dispatching, an improved deep learning method is proposed for power dispatching language recognition based on the analysis of speech recognition technology. Combining the improved deep residual shrinkage network with the improved gated convolutional neural network, the effective features are extracted by the improved deep residual shrinkage network, and the effective contextual information is obtained improved gated convolutional neural network. The performance of the proposed method is analyzed through tests to verify its superiority. The results indicate that, compared with conventional recognition methods, the proposed method has certain advantages in model parameters, character recognition error rate, and average recognition time. The model parameter quantity is 6.48 M, the character recognition error rate is 2.87%, and the average recognition time is 0.187 s. This study provides a certain reference for the development of language recognition methods for power dispatch.

**Keywords:** Power dispatch; Language recognition; Deep residual shrinkage network; Gated convolutional neural network; Word recognition error rate

2024-01-23 收稿; 2024-03-25 定稿

\*国网科技项目 (DGF5687485)

作者简介: 苕文涵 (1991-), 女, 安徽泗县人, 本科, 工程师, 研究方向: 电力调度通信, 图像语声信号处理。

<sup>†</sup>通信作者 E-mail: changwenhan2023@163.com

## 0 引言

我国双碳目标的全面展开,促使电网朝着能源互联网转型<sup>[1]</sup>。电力调度是电网稳定运行的基础。传统电力调度中语声交互在有多个电话时存在信息阻塞,无法满足调度自动化的需求<sup>[2]</sup>。调度语声识别是调度自动化实现的关键技术之一,对其进行研究具有一定的现实意义。

目前国内外对语声识别技术进行了大量的研究,但该技术在电力调度中应用较少,研究主要集中在声学模型和语言模型方面<sup>[3-5]</sup>。文献[6]提出一种用于电力调度语声识别的改进隐马尔可夫模型。所提方法在小规模词汇识别中效果较好,识别精度达到95.76%。文献[7]提出一种结合深度神经网络和隐马尔可夫模型的电力调度语声识别方法。所提方法具有较好的识别效果,识别精度达到94.63%。文献[8]提出一种结合改进动态时间调度算法和改进隐马尔可夫模型的电力调度语声识别方法。所提方法具有较好的降噪效果,且识别准确性较高。文献[9]提出一种深度神经网络用于电力调度系统语声识别模型。所提方法在多种方言特征中具有较好的识别效果,准确性较高,非常适合电网调度语声交互。为了提高电力调度语声的识别准确率,文献[10]提出一种基于深度卷积生成对抗性网络的电力调度语声交互模型。结果表明,相比于常规方法,所提方法在不同数据集识别中具有较高的识别准确率。为了提高传统声纹识别方法的准确性,文献[11]提出一种改进的深度残差收缩网络(Deep residual shrinkage networks, DRSN)用于电力调度语声识别。结果表明,所提方法可以有效降低噪声影响,具有较高的语声识别准确性。截至目前,人工智能相关方法已初步应用于电力系统语声识别领域,但上述语声识别的准确性和参数量有待进一步优化。

在此基础上,本文将改进的DRSN和改进的门控卷积神经网络(Gated convolutional neural network, GCNN)结合起来用于电力调度语言识别。改进的DRSN提取有效特征,然后通过堆叠改进的GCNN来获取有效上下文信息。通过试验验证了所提方法的优越性。

## 1 语声识别技术概述

语声识别主要由四部分组成:声学模型、语言模型、解码模块和特征提取。通过数据库对声学模

型和语言模型进行训练,结合字典构成网络空间。通过搜索算法找到识别的结果,语言识别系统结构如图1所示。

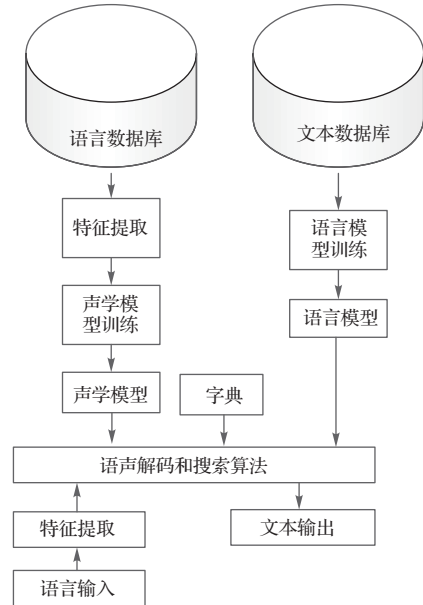


图1 语声识别系统结构

Fig. 1 Structure of speech recognition system

- (1) 特征提取:从输入的语言信号中提取声音特征,对声学模型进行训练<sup>[12]</sup>。
- (2) 声学模型:对特征提取的声学特征进行解析,本文主要对声学模型进行了研究<sup>[13]</sup>。
- (3) 语言模型:表征词汇间的概率统计,可以区分相同发音时的识别结果<sup>[14]</sup>。
- (4) 字典:将词或发声连接声学模型和语言模型<sup>[15]</sup>。
- (5) 搜索解码:根据声学模型,结合语言模型和字典找到最有可能的识别结果<sup>[16]</sup>。

## 2 声学模型

本文提出一种声学模型,将改进的DRSN和改进的GCNN相结合,通过改进的DRSN提取有效特征,再通过堆叠改进的GCNN来获取有效上下文信息。结合改进DRSN和改进GCNN用于电力调度语言识别,最后进行训练(CTC函数)<sup>[17-19]</sup>。图2所示本文声学模型整体构成。

### 2.1 改进DRSN

DRSN是一种人工智能方法,本质上DRSN由三部分组成——深度残差网络、注意力机制和软阈值函数<sup>[20-22]</sup>。本文将DRSN应用于语声识别,去

除噪声和冗余信息,提取有效特征。但 ReLU 强制的稀疏处理使一些参数无法激活,出现神经元坏死,残差结构在训练过程中网络层数较多时训练误差较大。

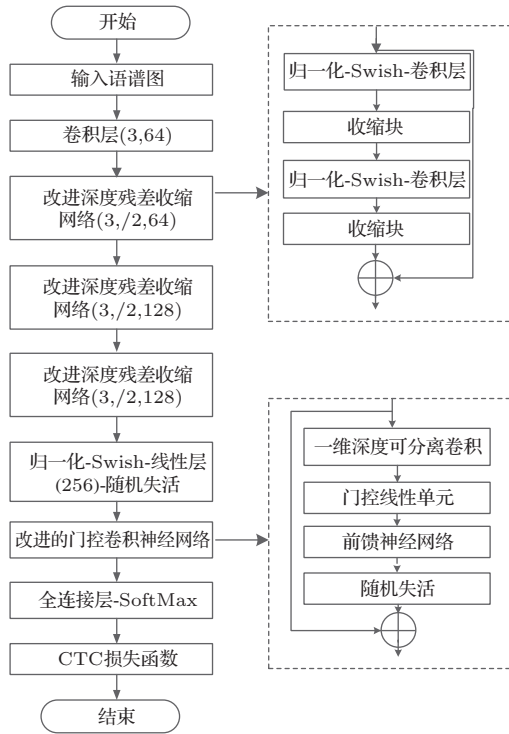


图 2 声学模型整体构成

Fig. 2 Overall composition of acoustic model

本文引入 Swish 激活函数对其进行优化,增强表达能力<sup>[23]</sup>。

Swish 激活函数如式(1)所示:

$$f(x) = x \cdot \text{sigmoid}(\beta x), \quad (1)$$

式(1)中, $\beta$ 为常量。Swish 具备无上界有下界、平滑、非单调的特性。

为了进一步提高 DRSN 的性能,将收缩模块与各层卷积神经网络(Convolutional neural network, CNN)结合<sup>[24]</sup>。DRSN 改进后充分发挥了收缩模块的优点,提高了模型的性能。改进的 DRSN 如图 3 所示。

### 2.2 改进的 GCNN

GCNN 主要由两部分组成——卷积层和门控线性单元 GLU, GLU 为门控机制,仅具有输出门,通过堆叠改进的 GCNN 来获取有效上下文信息<sup>[25-26]</sup>。但其层数和容量还有待提高,本文通过两个方面对其进行优化,即深度可分离卷积和前馈神经网络。

深度可分离卷积减少了卷积运算中的参数,提高了计算效率,已成功应用于图像分类任务<sup>[27]</sup>。其结构如图 4 所示。

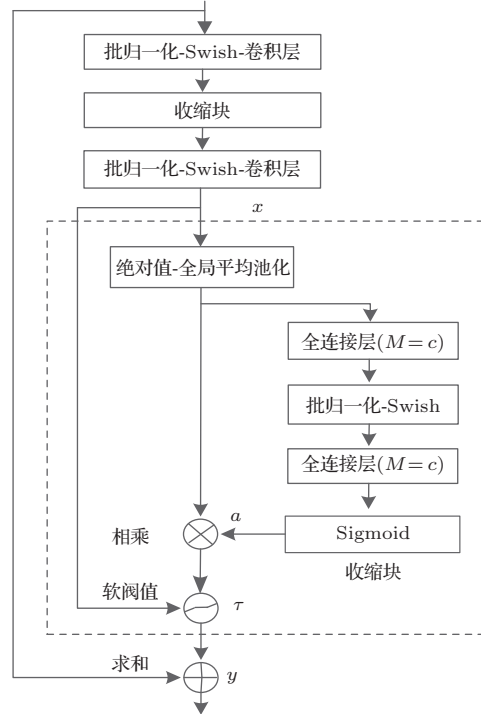


图 3 改进的 DRSN

Fig. 3 Improved DRSN

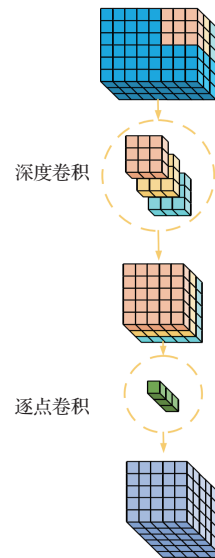


图 4 深度可分离卷积结构

Fig. 4 Depth wise separable convolutional structure

同时,引入前馈神经网络层变换 GCNN 输出空间,增加有效容量,表达能力增强<sup>[28-29]</sup>。为了防止过拟合,加入 dropout 层<sup>[30]</sup>。改进的 GCNN 如图 5 所示。

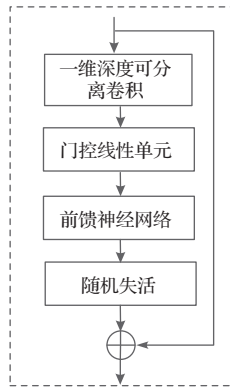


图5 改进 GCNN

Fig. 5 Improved GCNN

### 3 试验结果与分析

#### 3.1 试验参数

为了验证所提识别方法的优越性,在Python环境下基于Tensorflow深度学习框架进行了试验,表1为系统参数。

表1 系统参数  
Table 1 System parameters

配置	参数
CPU	I7-9750H
深度学习框架	Tensorflow2.3
操作系统	Ubuntu 17.10
GPU	RTX2080Ti
编程语言	Python3.8

为了验证模型在真实场景中的性能,采用了某供电公司真实场景中50000条调度语声数据作为试验数据集。

在安静环境下采用传声器进行录制,录制人员为10名普通话标准的大学生,采用频率为16 kHz。将数据集划分为训练集和测试集,训练集、验证集和测试集的比值为4:1:1<sup>[31]</sup>。电力调度语声识别中很多指令(如变电站名和线路名等)与常规语声识别词汇不同,出现的频率也较低,无法有效利用上下文信息,在电力调度语声识别的准确率较低。本文通过波束搜索进行解码,采用文献[32]的BERT语言模型,试验参数如表2所示。

不同的指标通常会产生不同的评价结果,因此需要选择有效的指标进行评价。本文选择字错误率(Character error rate, CER)对结果进行评价,如式(2)所示<sup>[33]</sup>:

$$\text{CER} = \frac{R + I + D}{N} \times 100\%, \quad (2)$$

式(2)中, $R$ 、 $I$ 、 $D$ 分别为替换、插入、删除的错误字数; $N$ 为总字数。

表2 试验参数  
Table 2 Experimental parameters

参数	数值
输入特征	语谱图
时间窗宽度	25 ms
窗移	10 ms
初始学习率	0.0001
批量大小	8
优化器	Adam

#### 3.2 试验分析

为了验证所提语声识别方法在优化前后的有效性,将本文模型与DRSN-GCNN、改进DRSN-GCNN和DRSN-改进GCNN三个优化前模型进行了比较分析。不同模型的CER如表3所示。

表3 不同模型的CER  
Table 3 CER of different models

模型	训练集 CER/%	测试集 CER/%
本文方法	2.58	2.87
DRSN-GCNN	10.21	10.90
改进DRSN-GCNN	6.81	7.70
DRSN-改进GCNN	4.51	5.21

由表3可以看出,无论是训练集还是测试集,本文模型的CER都是最低的,分别为2.58%和2.87%。通过比较DRSN-GCNN、改进DRSN-GCNN和DRSN-改进GCNN的试验。验证了增加前馈神经网络层的有效性,使模型具有较强的表达能力。表明本文方法可以更好地提高词谱图的特征学习能力和捕捉有效上下文信息能力。

为了验证模型受改进GCNN层数的影响程度,将改进GCNN层数分别设定为8、9、10、11、12、13、14。不同改进GCNN层数本文模型的语声识别结果如图6所示。

从图6可以看出,随着GCNN层数的增加,CER先降低在升高,层数为10时,无论是训练集还是测试集,CER都为最低,分别为2.58%和2.87%;层数少时,模型的表现能力弱导致泛化能力很弱。如果GCNN层数较多,则由于训练数据较少,模型可能欠拟合。这是因为GCNN存储信息容

量为定量,多层GCNN导致后面网络有效信息越来越少。

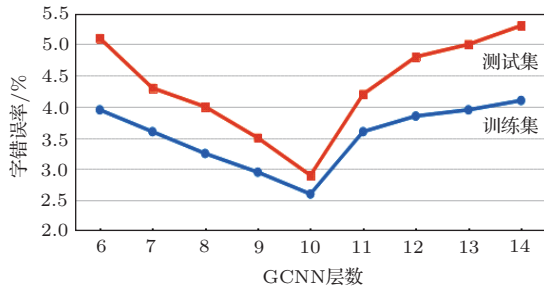


图6 不同GCNN层数的语音识别结果

Fig. 6 Speech recognition results with different GCNN layers

为了验证本文模型对不同输入特征的识别效果,分别输入语谱图、Fbank(40维的声学特征)和梅尔频率倒谱系数(Mel-frequency cepstrum coefficient, MFCC)(1阶、2阶差分的39维的特征)。不同输入特征的试验结果如表4所示。

表4 不同输入特征 CER  
Table 4 CER with different input features

输入特征	训练集 CER/%	测试集 CER/%
语谱图	2.58	2.87
Fbank	4.62	5.49
MFCC	10.51	11.53

由表4可以看出,不管是在验证集还是测试集,语谱图输入CER都是最低的,与Fbank CER相比分别下降了44.16%和47.72%,与MFCC CER相比分别下降了75.45%和75.11%。这主要是因为Fbank和MFCC在多次提取后丢失了部分原始信息,MFCC原始信息丢失最多,所以CER最高。而语谱图原始信息保留最完整,虽然含有一定的噪声和冗余信息,但本文模型可以很好地清除这些不利因素,达到最好的识别效果。

本文进一步验证了该文模型在噪声环境下的鲁棒性,将咖啡馆噪声、汽车噪声和白噪声加入,语音随机包含一种噪声,将5组信噪比分别设为-10 dB、-5 dB、0 dB、5 dB、10 dB加入测试集,不同信噪比测试集的识别结果如图7所示。

由图7可以看出,本文模型抗噪性能优于DRSN-GCNN模型、改进DRSN-GCNN模型和DRSN-改进GCNN模型,优化前后模型均下降缓慢。这是因为模型中的收缩模块可以自适应地学习

语谱图相关噪声和冗余信息的阈值,通过阈值消除噪声和冗余信息,可以得到有效的特征,且特征更具有代表性。另外,前馈神经网络层的引入增加了GCNN的有效容量,使得模型提取的上下文信息更加有效,识别效果较好。

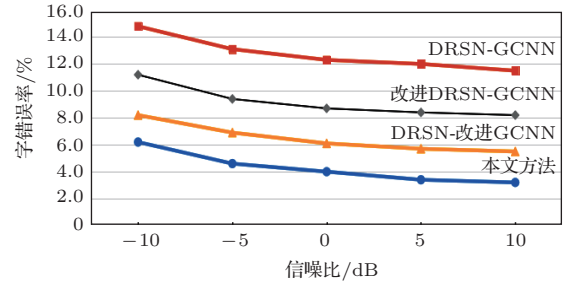


图7 不同信噪比的语音识别结果

Fig. 7 Speech recognition results with different signal-to-noise ratios

为了验证本文模型的优越性,将本文方法与文献[33]CNN结合BLATM模型、文献[34]的DFSMN-T模型、文献[35]的DL-T模型进行了比较分析,均采用本文数据集进行训练和测试,分别对比了模型参数、CER和平均识别时间。不同模型在测试集上的试验结果如表5所示。

表5 不同模型测试集结果对比  
Table 5 Results comparison of different model test sets

模型	模型参数/M	CER/%	平均耗时/s
本文方法	6.48	2.87	0.187
文献[33]	26.32	4.52	0.452
文献[34]	262.00	7.78	0.654
文献[35]	28.00	4.77	0.338

由表5可以看出,本文模型在模型参数、CER和平均耗时方面均为最佳。与文献[33]、文献[34]和文献[35]模型相比,本文模型参数分别降低了75.38%、97.40%、76.86%; CER分别降低了36.50%、63.11%、39.83%; 平均耗时分别降低了58.63%、71.41%、44.67%。

## 4 结论

本文提出将改进的DRSN和改进的GCNN结合起来用于电力调度语言识别。通过改进DRSN来提取特征,通过堆叠改进GCNN来获取有效的上下文信息。结果表明,所提方法能够提高词谱

图的特征学习,捕捉有效的上下文信息。相比于文献[33]模型、文献[34]模型和文献[35]模型,本文模型参数量分别降低75.38%、97.40%、76.86%, CER分别降低36.50%、63.11%、39.83%,平均耗时分别降低58.63%、71.41%、44.67%。本文可为电力调度自动化提供一定参考。但仍存在一些不足,如仅对声学模型进行研究和识别模型性能还有待提高,后期将结合一些模型提高识别能力。

## 参 考 文 献

- [1] 顾晓东,唐丹宏,黄晓华. 基于深度学习的电网巡检图像缺陷检测与识别[J]. 电力系统保护与控制, 2021, 49(5): 91-97.  
Gu Xiaodong, Tang Danhong, Huang Xiaohua. Defect detection and recognition of power grid inspection images based on deep learning[J]. Power System Protection and Control, 2021, 49(5): 91-97.
- [2] 夏玉果,董天天,丁晟. 基于轻量化深度迁移神经网络的电子元器件识别[J]. 电子器件, 2023, 46(6): 1673-1679.  
Xia Yuguo, Dong Tiantian, Ding Sheng. Electronic component recognition based on lightweight deep transfer neural network[J]. Chinese Journal of Electron Devices, 2023, 46(6): 1673-1679.
- [3] 周艳真,查显煜,兰健,等. 基于数据增强和深度残差网络的电力系统暂态稳定预测[J]. 中国电力, 2020, 53(1): 22-31.  
Zhou Yanzhen, Cha Xianyu, Lan Jian, et al. Transient stability prediction of power systems based on data augmentation and deep residual networks[J]. Electric Power, 2020, 53(1): 22-31.
- [4] 赵涛,张羿,王永和,等. 基于深度学习的人机语音交互平台[J]. 信息系统工程, 2019, 12(1): 102-104.  
Zhao Tao, Zhang Yi, Wang Yonghe, et al. A human computer speech interaction platform based on deep learning[J]. China CIO News, 2019, 12(1): 102-104.
- [5] 赵晴,李庭瑞,罗睿,等. 基于双字典类标签语言模型的电力调度语音识别[J]. 电子测量技术, 2021, 44(13): 121-126.  
Zhao Qing, Li Tingrui, Luo Rui, et al. Power dispatch speech recognition based on dual dictionary class label language model[J]. Electronic Measurement Technology, 2021, 44(13): 121-126.
- [6] 鄢发齐,王春明,窦建中,等. 基于隐马尔可夫模型的电力调度语音识别研究[J]. 武汉大学学报(工学版), 2018, 51(10): 920-923.  
Yan Faqi, Wang Chunming, Dou Jianzhong, et al. Research on power dispatching speech recognition based on hidden Markov model[J]. Journal of Wuhan University(Engineering Edition), 2018, 51(10): 920-923.
- [7] 窦建中,罗深增,金勇,等. 基于深度神经网络的电力调度语音识别研究及应用[J]. 湖北电力, 2019, 43(3): 16-22.  
Dou Jianzhong, Luo Shenzeng, Jin Yong, et al. Research and application of speech recognition for power dispatching based on deep neural networks[J]. Hubei Electric Power, 2019, 43(3): 16-22.
- [8] 王素宁,朱俊杰,李志勇,等. 基于DTW算法的电力调度语音识别研究和应用[J]. 电力与能源, 2021, 42(1): 35-38, 64.  
Wang Suning, Zhu Junjie, Li Zhiyong, et al. Research and application of power dispatching speech recognition based on DTW algorithm[J]. Electricity & Energy, 2021, 42(1): 35-38, 64.
- [9] 胡翔,杨洋,蒋长江,等. 一种基于深度神经网络的电力系统调度控制语音识别模型[J]. 电子器件, 2023, 46(1): 90-95.  
Hu Xiang, Yang Yang, Jiang Changjiang, et al. A speech recognition model for power system scheduling control based on deep neural networks[J]. Chinese Journal of Electron Devices, 2023, 46(1): 90-95.
- [10] Wang Z H, Gao F. Research on voice interaction model of intelligent power dispatching based on DCGAN[J]. Nanotechnology for Environmental Engineering, 2021, 6(3): 53.
- [11] Zhang Q R, Zhai H T, Ma Y Y, et al. Enhanced-deep-residual-shrinkage-network-based voiceprint recognition in the electric industry[J]. Electronics, 2023, 12(14): 3017-3031.
- [12] 王泽霞,陈革,陈振中. 基于改进卷积神经网络的化纤丝饼表面缺陷识别[J]. 纺织学报, 2020, 41(4): 39-44.  
Wang Zexia, Chen Ge, Chen Zhenzhong. Surface defect recognition of synthetic fiber cake based on improved convolutional neural network[J]. Journal of Textiles, 2020, 41(4): 39-44.
- [13] 许洪强,蔡宇,万雄,等. 电网调控大数据平台体系架构及关键技术[J]. 电网技术, 2021, 45(12): 4798-4807.  
Xu Hongqiang, Cai Yu, Wan Xiong, et al. Architecture and key technologies of power grid regulation big data platform[J]. Power System Technology, 2021, 45(12): 4798-4807.
- [14] 邱志斌,石大寨,况燕军,等. 基于深度迁移学习的输电线路涉鸟故障危害鸟种图像识别[J]. 高电压技术, 2021, 47(11): 3785-3794.  
Qiu Zhibin, Shi Dazhai, Kuang Yanjun, et al. Bird species image recognition based on deep transfer learning for bird related faults in transmission lines[J]. High Voltage Technology, 2021, 47(11): 3785-3794.
- [15] 徐冬冬. 基于Transformer的普通话语音识别模型位置编码选择[J]. 应用声学, 2021, 40(2): 194-199.  
Xu Dongdong. Selection of position encoding for mandarin phonetic recognition model based on transformer[J]. Journal of Applied Acoustics, 2021, 40(2): 194-199.
- [16] 颜宏文,陈金鑫. 基于改进YOLOv3的绝缘子串定位与状态识别方法[J]. 高电压技术, 2020, 46(2): 423-432.  
Yan Hongwen, Chen Jinxin. A method for insulator string positioning and state recognition based on improved YOLOv3[J]. High Voltage Technology, 2020, 46(2): 423-432.
- [17] 杨德举,马良荔,谭琳珊,等. 基于门控卷积网络与CTC的端到端语音识别[J]. 计算机工程与设计, 2020, 41(9): 2650-2654.  
Yang Deju, Ma Liangli, Tan Linshan, et al. End to end speech recognition based on gated convolutional networks

- and CTC[J]. *Computer Engineering and Design*, 2020, 41(9): 2650–2654.
- [18] 张晓旭, 马志强, 刘志强, 等. Transformer在语音识别任务中的研究现状与展望[J]. *计算机科学与探索*, 2021, 15(9): 1578–1594.
- Zhang Xiaoxu, Ma Zhiqiang, Liu Zhiqiang, et al. The current research status and prospects of Transformer in speech recognition tasks[J]. *Journal of Frontiers of Computer Science and Technology*, 2021, 15(9): 1578–1594.
- [19] 吕坤儒, 吴春国, 梁艳春, 等. 融合语言模型的端到端中文语音识别算法[J]. *电子学报*, 2021, 49(11): 2177–2185.
- Lyu Kunru, Wu Chunguo, Liang Yanchun, et al. End to end Chinese speech recognition algorithm integrating language models[J]. *Acta Electronica Sinica*, 2021, 49(11): 2177–2185.
- [20] 于蒙, 李雄, 杨海潮, 等. 基于图像识别的苹果的等级分级研究[J]. *自动化与仪表*, 2019, 34(7): 39–43, 47.
- Yu Meng, Li Xiong, Yang Haichao, et al. Research on apple grading based on image recognition[J]. *Automation & Instrumentation*, 2019, 34(7): 39–43, 47.
- [21] 李云红, 梁思程, 贾凯莉, 等. 一种改进的DNN-HMM的语音识别方法[J]. *应用声学*, 2019, 38(3): 371–377.
- Li Yunhong, Liang Sicheng, Jia Kaili, et al. An improved DNN-HMM speech recognition method[J]. *Journal of Applied Acoustics*, 2019, 38(3): 371–377.
- [22] 王晶晶, 姜明, 张旻. 基于上下文感知的方面类别情感分类[J]. *计算机应用研究*, 2021, 38(6): 1770–1774.
- Wang Jingjing, Jiang Ming, Zhang Min. Context aware aspect category sentiment classification[J]. *Computer Application Research*, 2021, 38(6): 1770–1774.
- [23] 王智超, 徐及, 张鹏远, 等. 卷积神经网络声学模型的结构优化和加速计算[J]. *重庆邮电大学学报(自然科学版)*, 2018, 30(3): 416–422.
- Wang Zhichao, Xu Ji, Zhang Pengyuan, et al. Structural optimization and accelerated computation of convolutional neural network acoustic model[J]. *Journal of Chongqing University of Posts and Telecommunications(Natural Science Edition)*, 2018, 30(3): 416–422.
- [24] Singh A, Kabra R, Kumar R, et al. On-device system for device directed speech detection for improving human computer interaction[J]. *IEEE Access*, 2021, 9: 131758–131766.
- [25] 杜永萍, 刘杨, 贺萌. 基于对抗式分布对齐的跨域方面级情感分析[J]. *模式识别与人工智能*, 2021, 34(1): 87–94.
- Du Yongping, Liu Yang, He Meng. Cross domain aspect level sentiment analysis based on adversarial distribution alignment[J]. *Pattern Recognition and Artificial Intelligence*, 2021, 34(1): 87–94.
- [26] 景丽, 李曼曼, 何婷婷. 结合扩充词典与自监督学习的网络评论情感分类[J]. *计算机科学*, 2020, 47(S2): 78–82, 91.
- Jing Li, Li Manman, He Tingting. Combining extended dictionaries with self supervised learning for sentiment classification in online comments[J]. *Computer Science*, 2020, 47(S2): 78–82, 91.
- [27] 卢麒仰, 徐超林, 杨育, 等. 基于改进YOLOv5网络的印刷电路板缺陷检测[J]. *电子器件*, 2023, 46(6): 1504–1508.
- Lu Qiyang, Xu Chaolin, Yang Yu, et al. Printed circuit board defect detection based on improved YOLOv5 network[J]. *Chinese Journal of Electron Devices*, 2023, 46(6): 1504–1508.
- [28] 曾伟, 武华, 冯秀平, 等. 一种低反向恢复电流的无电压回跳RC-IGBT设计[J]. *电子器件*, 2023, 46(6): 1480–1483.
- Zeng Wei, Wu Hua, Feng Xiuping, et al. A voltage free rebound RC-IGBT design with low reverse recovery current[J]. *Chinese Journal of Electron Devices*, 2023, 46(6): 1480–1483.
- [29] 邵良杉, 周玉. 基于语义规则与RNN模型的在线评论情感分类研究[J]. *中文信息学报*, 2019, 33(6): 124–131.
- Shao Liangshan, Zhou Yu. Research on sentiment classification of online comments based on semantic rules and RNN model[J]. *Journal of Chinese Information Processing*, 2019, 33(6): 124–131.
- [30] 陈佳伟, 韩芳, 王直杰. 基于自注意力门控图卷积网络的特定目标情感分析[J]. *计算机应用*, 2020, 40(8): 2202–2206.
- Chen Jiawei, Han Fang, Wang Zhijie. Specific target sentiment analysis based on self attention gated graph convolutional network[J]. *Journal of Computer Applications*, 2020, 40(8): 2202–2206.
- [31] 王阁藩, 刘博, 李恺, 等. 一种低回踢噪声低失调高频CMOS比较器[J]. *电子器件*, 2023, 46(6): 1474–1479.
- Wang Gefan, Liu Bo, Li Kai, et al. A low bounce noise, low offset high-frequency CMOS comparator[J]. *Chinese Journal of Electron Devices*, 2023, 46(6): 1474–1479.
- [32] 陈蕾, 郑伟彦, 余慧华, 等. 基于BERT的电网调度语音识别语言模型研究[J]. *电网技术*, 2021, 45(8): 2955–2961.
- Chen Lei, Zheng Weiyang, Yu Huihua, et al. Research on speech recognition language model for power grid dispatching based on BERT[J]. *Power System Technology*, 2021, 45(8): 2955–2961.
- [33] Sultana S, Iqbal M Z, Selim M R, et al. Bangla speech emotion recognition and cross-lingual study using deep CNN and BLSTM networks[J]. *IEEE Access*, 2022, 10: 564–578.
- [34] 胡章芳, 蹇芳, 唐珊珊, 等. DFSMN-T: 结合强语言模型Transformer的中文语音识别[J]. *计算机工程与应用*, 2022, 58(9): 187–194.
- Hu Zhangfang, Jian Fang, Tang Shanshan, et al. DFSMN-T: Chinese speech recognition combined with strong language model Transformer[J]. *Computer Engineering and Applications*, 2022, 58(9): 187–194.
- [35] 张威, 刘晨, 费鸿博, 等. 基于DL-T及迁移学习的语音识别研究[J]. *工程科学学报*, 2021, 43(3): 433–441.
- Zhang Wei, Liu Chen, Fei Hongbo, et al. Research on speech recognition based on DL-T and transfer learning[J]. *Chinese Journal of Engineering*, 2021, 43(3): 433–441.