

浅谈电力调度语音识别技术研究进展

李长东, 欧阳文华, 徐瑞泽, 舒信宇

(国网江西省电力有限公司电力科学研究院, 江西 南昌 330096)

摘要: 电力调度是电网安全稳定运行的核心环节。传统人机交互模式已难以满足日益增长的调度信息处理需求。自动语音识别(automatic speech recognition, ASR)技术在电力调度领域展现出巨大潜力, 但面临噪声、专业词汇识别、数据匮乏、实时性、安全性和方言口音等多重挑战。文中全面总结了 ASR 技术在电力调度中的研究进展, 包括基于深度学习(如 DRSN-Conformer、CNN-TDNN-FSA、CT-LSTM)的声学模型在噪声抑制和语音特征提取方面的创新, 以及领域定制化语言模型(如 NE-BERT、融合专业术语模块、大语言模型微调)在提升专业词汇识别准确率和语义理解上的关键作用; 文章还探讨了流式识别、本地化部署、安全监督和声纹识别等系统实现与应用; 最后, 展望了未来在数据扩充、多模态融合及大语言模型应用等方向的趋势, 旨在推动电力调度智能化进一步发展。

关键词: 电力调度; 语音识别; ASR; 声学模型; 语言模型

中图分类号: TM 391.42 **文献标志码:** A **文章编号:** 1006-348X(2026)03-0012-05

0 引言

电力调度作为电力系统运行的“大脑”, 承担着电力资源优化配置和电网安全调控的重任, 任何指令误判或误操作均可能引发重大事故。长期以来, 电力调度主要依赖调度员通过电话指令进行人机交互与人工记录, 这种传统人机交互模式在日益复杂的智能电网环境中暴露出诸多弊端。具体而言, 人工记录耗时耗力, 易造成信息遗漏和记录不准确; 大量通话信息涌入, 易导致信息阻塞和调度效率下降; 极端情况下, 甚至可能因误操作引发重大事故。

自动语音识别(automatic speech recognition, ASR)技术能将语音自动转为文本或指令, 是提升调度交互效率、降低误操作风险的有效手段, 引入自动语音识别技术成为提升电力调度效率与安全性的必然趋势。

文中旨在总结电力调度语音识别技术的研究进展、面临挑战及未来方向, 以期为电力调度智能化转型提供参考。

1 电力调度语音识别面临的挑战

1.1 复杂声学环境制约

电力调度中心和变电站等工作场所环境复杂多变, 设备运行产生的轰鸣声、电流声、风扇声以及环境混响等构成持续性背景噪声, 严重劣化语音信号质量。这些噪声使得语音信号的信噪比显著降低, 语音特征提取困难, 影响 ASR 模型的识别精度。研究表明, 通用 ASR 模型在相对安静的场景中表现良好, 但在真实电力调度场景下准确率显著下降。

1.2 领域语言特性限制

电力调度语言具有高度专业性和规范性, 与日常用语存在显著差异, 主要表现为:

1) 专业术语和命名实体密集: 调度指令中包含大量电力行业特有的专业词汇, 如“XX 变电站”“冷备用转运行”“断路器”“隔离开关”等。这些词汇在用于通用模型训练的语料库中出现频率极低, 导致通用语言模型难以准确识别。

2) 特殊发音规则: 数字、时间等在电力调度中有

收稿日期: 2026-04-10

作者简介: 李长东(1997), 男, 硕士, 工程师, 研究方向为电力人工智能、调度语音识别、主网故障研判。

特定的读法,例如“1780”常读作“幺拐八洞”,而非“一千七百八十”;“09:00”读作“洞拐点洞洞”。这种差异性极易造成 ASR 系统的误识别。

3) 口语化与非规范表达:调度员在实际通话中可能存在口语化、省略或非标准表达,增加了语音识别的难度。

1.3 数据稀缺与标注挑战

电力调度数据因其敏感性和保密性,难以对外公开,导致缺乏大规模的公开数据集进行模型训练。自建语料库需要耗费巨大的人力物力进行语音采集、整理和精确标注,且数据更新缓慢。这种数据匮乏严重制约了模型的训练效果,尤其是在专业术语识别方面。

1.4 硬件计算资源限制

电力调度指令的发布和接收需要极高的实时性,尤其在处理紧急故障时,系统响应延迟可能导致严重后果。然而,高性能的端到端 ASR 模型通常参数量庞大,计算复杂度高,难以在电力调度终端有限的硬件资源(低功耗、低算力)上实现高效、低延迟的部署和运行。

1.5 安全性制约

电力系统中的语音数据包含大量敏感信息(如电力运行状态、关键指令、位置信息等),一旦泄露可能引发严重安全后果。因此,电力调度系统强制要求 ASR 系统进行本地化、离线化部署,避免数据通过外部网络传输至第三方云服务平台,这与当前主流 ASR 服务普遍依赖云端计算的模式形成冲突。

1.6 地域口音差异

我国幅员辽阔,不同地区的调度人员与现场操作人员常带有一定的地域性口音或方言特征。以江西省为例,当地方言在语音、语调、词汇等方面与普通话存在显著差异,且省内不同区县的方言之间也有明显区别。这种语音特征的区域性差异,极大增加了自动语音识别模型对不同口音泛化能力的要求。如果模型无法有效适应多样化的口音,将直接影响到语音识别的准确率与系统在实际场景中的可用性。

2 电力调度语音识别技术研究进展

语音识别系统的基本框架通常包括信号预处理、特征提取、声学模型、语言模型等模块。为应对上述

挑战,研究人员主要在声学模型、语言模型和系统集成等方面进行了大量创新。

2.1 声学模型研究进展

声学模型主要负责将语音信号转换为音素或字符序列的概率分布,其发展大致经历了模板匹配阶段、统计模型阶段和深度学习阶段。其中,模板匹配、统计模型归为传统声学模型。

2.1.1 传统声学模型

模板匹配阶段:语音识别技术起源于 20 世纪 50 年代,贝尔实验室的 Audrey 系统实现了对 10 个孤立数字的识别^[1]。该阶段主要采用模板匹配算法,通过动态时间规整技术解决了不同时长音频的匹配问题。虽然,模板匹配方法对孤立词识别效果较好,但对陌生发音识别准确率较低^[2]。

统计模型阶段:20 世纪 80 年代,基于概率统计的隐马尔可夫模型(hidden markov model, HMM)及其与高斯混合模型(gaussian mixture model, GMM)结合的 GMM-HMM 框架成为语音识别的主流方法。研究人员运用统计方法,采用梅尔倒谱系数作为声学特征,在小规模词汇量下,基于 GMM-HMM 的声学建模方法在电网调度语音数据集上达到了 95.76% 的识别精度^[3]。GMM-HMM 模型能够表征语音信号的时变性和波动性,具有很好的模型参数估计与连续解码能力,并在之后的二十年间得到广泛应用。然而, GMM-HMM 在识别率达到 80% 左右后,难以再取得突破,面临模型表征能力不强、鲁棒性差等问题。

2.1.2 基于深度学习的声学模型

进入 21 世纪,深度学习技术的飞速发展极大地促进了 ASR 性能的提升,尤其是在处理复杂语音和大规模数据方面。

2009 年, Hinton 等人首次将深度神经网络(deep neural network, DNN)应用于声学建模并取得成功^[3];随后,大量学者用 DNN 替代 GMM,对 HMM 状态的后验概率进行建模,形成了 DNN-HMM 混合模型,在连续语音识别任务中显著提升;窦建中等人将 DNN-HMM 技术引入电力调度,识别准确率达到 94.63%,显著优于传统的 GMM-HMM^[4]。

循环神经网络(recurrent neural network, RNN)特别是长短期记忆网络(long short-term memory network, LSTM),因其在处理序列数据方面的优势,被

广泛应用于语音识别; Alex Graves 等人研究了深度循环神经网络, 并结合 LSTM 架构, 在 TIMIT 音素识别基准测试中实现了 17.7% 的测试集错误率, 表明其在语音识别领域具有先进水平^[5]。LSTM 基于门控机制解决了 RNN 的梯度消失和梯度爆炸问题, 有效提升了语音识别系统的整体性能。

2017 年, Ashish Vaswani 等人舍弃了循环和卷积, 提出 Transformer 模型, 在捕获长距离信息和并行训练方面表现出色, 成为序列建模任务的特殊 AED 结构^[6]; 另一方面, 随着卷积神经网络 (convolutional neural network, CNN) 的广泛应用, 研究人员发现其在处理语音信号受环境干扰的问题上表现出色, CNN 逐渐被大量研究者用于语音识别任务。Anmol Gulati 等人将 CNN 与 Transformer 相结合, 提出了 Conformer 网络, 有效建模音频序列的局部和全局依赖性, 在 LibriSpeech 基准测试中取得了 2.1% 的词错误率 (word error rate, WER)^[7]; 针对电力调度噪声场景, 有研究提出 DRSN-Conformer (DRformer) 模型^[8], 该模型在声学编码器 Conformer 前加入深度残差收缩网络 (DRSN), 通过动态软阈值模块对含噪音频进行降噪预处理, 有效减少噪声干扰; 在公开数据集 Aishell-1、Thchs30 及某电网公司内部电力调度数据集上的实验表明, DRformer 相较于 Transformer 网络模型, 字错误率 (character error rate, CER) 降低了 8.5%, 专业术语识别字错误率降低了 5.2%; 卢蓬媛提出 CTDNN-FSA 模型, 通过 CNN 与深层 TDNN-F 级联增强局部特征提取, 并引入受时间约束的多头自注意力层捕捉上下文关联, 使字符错误率降低 2.31%^[1]; 金保含提出的 CT-LSTM 声学模型结合了 CNN、TDNN-F 和 LSTM 的优势, 通过关注语音特征的局部和全局信息, 相比单一 LSTM 网络, 字错误率降低 5%^[9]。

2.2 语言模型研究进展

语言模型 (language model, LM) 旨在用于计算上下文中某一句子或单词出现的概率, 是语音识别系统中另一个重要模块, 其可分为基于统计的语言模型和基于神经网络的语言模型两类。

2.2.1 基于统计的语言模型

传统的语言模型多基于 N-gram 统计方法。N-gram 模型通过统计文本数据中词或字符序列的出现频率来预测下一个词的概率, 结构相对简单、计算效

率高, 能够获取局部依赖; 然而, N-gram 模型最大的弊端是存在数据稀疏问题, 且难以捕捉长距离依赖关系。

2.2.2 基于神经网络的语言模型

随着深度学习的发展, 基于神经网络语言模型逐渐取代传统基于统计的语言模型。

Jacob Devlin 等人首次提出 BERT 模型, 利用双向自注意力机制, 从未标记文本中预训练深度双向表示, 有效提取了长距离依赖关系, 在多项自然语言处理任务中取得了出色效果^[10]; 陈蕾等人将 BERT 模型用于电网调度语音识别任务, 提出基于 BERT 的电网调度语音识别语言模型^[11], 通过对调度语句进行语义特征、关键字特征和命名实体特征提取, 提高了模型对电网调度语言的适应性; 闫圣学进一步提出 NE-BERT 语言模型, 通过改进 BERT 的掩码策略和语言模型训练方式, 融合电力调度语义特征, 显著提升了电力调度语音识别任务的准确率。实验结果显示, NE-BERT+BiLSTM+CRF 模型相比于 BERT+BiLSTM+CRF 模型, 提升了 5.38% 的精确度和 2.45% 的 F1 分数^[12]。

韩亚旭等人在电力调度语音识别语言模型术语模块中加入了专业术语模块, 该模块由文本编码器和交叉注意力机制层构成, 结合 CTC 和注意力损失进行联合训练, 旨在提高 DRformer 声学模型对电力调度专业术语的识别准确率。实验结果显示, 融合专业术语模块后, 字错误率相对降低 4%^[13]。

2.3 语音识别模型对比

为更直观地展示不同模型在电力调度场景下的性能和适用性, 本节对文中重点提及的各类模型进行横向对比, 主要从技术特点、适用场景及性能优势维度进行分析。

如表 1 所示, GMM-HMM 模型在小规模词汇量和简单场景下能达到较高的识别精度, 为语音识别奠定了基础。然而, 随着电力调度场景复杂度的提升, DNN-HMM 等深度学习模型的引入, 通过强大的非线性建模能力, 显著提升了识别准确率, 并能利用上下文信息增强泛化能力。

针对电力调度场景特有的复杂噪声环境, DRformer 模型通过融合 DRSN, 在降噪方面表现出突出优势, 有效降低了电力调度场景的字错误率。

CTDNN-FSA 和 CT-LSTM 模型则通过结合 CNN、TDNN-F 和 LSTM,增强了模型对语音信号局部和全局特征的表征和还原能力,尤其擅长处理语音信号的时域频域偏移或畸变。

通用语言模型 BERT 在电力调度专业领域识别性能不佳,NE-BERT 模型通过对 BERT 的改进和领域语义特征的融合,显著提升了对电力调度专业词汇的识别准确率,体现了领域定制化的重要性。RNN-T 模型因其天然的流式化特性,成为满足电力调度高实时性需求的理想选择。

表 1 各语音识别模型在调度领域性能对比

模型名称	技术特点	适用场景及性能优势
GMM-HMM	基于统计建模,高斯混合模型与隐马尔可夫链	适用小规模词汇、简单声学环境环节,CER 约 4.24%
DNN-HMM	用深度神经网络替代 GMM,结合 HMM	利用上下文信息增强泛化能力,擅长处理大词汇量连续语音识别,CER 约 5.37%
DeepRNNs/LSTM	循环神经网络,长短期记忆网络	解决 RNN 梯度问题,TIMIT CER 17.7%
Conformer	CNN 与 Transformer 结合	有效平衡局部与全局信息,LibriSpeech WER 2.1%/4.3%
DRformer	DRSN 与 Conformer 结合	噪声鲁棒性强,适用于复杂噪声环境,电力调度场景 CER 15.9%
CTDNN-FSA	CNN 与 TDNN-F 结合,多头自注意力	增强局部特征提取,捕捉上下文关联,擅长处理电力调度语音信号的时域频域偏移或畸变
CT-LSTM	CNN、TDNN-F 和 LSTM 结合,兼顾局部与全局特征	相比单一 LSTM CER 降低 5%,相比 TDNN-F CER 降低 4.64%,提升整体识别效果
NE-BERT	改进 BERT,动态掩码策略,融合语义特征	解决通用 BERT 在电力调度领域适应性不佳问题,相比 Transformer-LM CER 降低 0.87%
RNN-T	改进 CTC、预测网络,在线实时解码	克服 CTC 语言建模能力弱问题,适用于实时识别任务

3 系统实现与应用研究

电力调度语音识别系统一般包括语音数据采集、信号预处理与语音特征提取、声学模型、语言模型等组件,各组件功能描述如表 2 所示。为提升 ASR 技术在电力调度场景的应用效果,研究人员在系统层面进行了适应性改进。

表 2 电力调度语音识别系统组件及功能

组件名称	功能描述	典型技术/原型
语音数据采集	采集电力调度通话语音数据,通常为 8 kHz 采样率	/
信号预处理与特征提取	对原始语音信号进行降噪、静音检测、端点切割,提取 MFCC 等声学特征	DRSN
声学模型	将语音特征序列映射为音素或字符序列的概率分布。	GMM-HMM、DNN-HMM、DRformer、CTDNN-FSA、CT-LSTM
语言模型	评估词序列出现的概率,处理领域特定语义和语法。	N-gram、BERT、NE-BERT、TransformerLM、LLM
解码器	结合声学模型、语言模型和发音词典,搜索最优文本序列。	WFST 解码、CTC、RNN-T、AED
专用数据集与词典	包含电力调度专业术语、特殊发音规则的语音和文本数据。	Pd-Data、Pd-Data1、Pd_E

3.1 实时性流式语音识别

电力调度系统对实时性要求极高,传统非流式技术难以满足电力调度对实时性的要求。文献[14]选择具有流式化特性的循环神经网络转换器(RNN-transducer, RNN-T)模型作为研究对象,并对其网络框架进行改进优化。通过引入预测网络实现声学模型和语言模型的端到端联合优化,能实现在线实时解码。

3.2 数据稀缺和地域口音适配

针对电力调度数据稀缺问题,文献[1]提出基于迁移学习的电力调度语音识别系统设计方法。该方法通过在通用数据集上预训练模型,再利用小规模电力调度专有数据集进行微调,从而提高模型对专有词汇的识别率,并解决了训练数据匮乏的问题,提升了模型优化升级效率。

为提升语音识别模型对地域口音的识别准确率,文献[9]提出多地区口音语音识别方法,在 CT-LSTM 声学模型基础上,针对特定方言训练输出层,对不同地区的语料进行训练,共享隐藏层神经网络,有效适应地域口音。实验结果显示,该方法在西北地区数据集的字错误率降低了 5.32%。

3.3 本地化与安全部署

针对电力调度数据的敏感性和安全性要求,本地化部署是关键。将语音识别服务部署在本地,无需外部网络连接,可有效保障了数据安全。王邦鸿设计并实现了基于国网电力调度场景的本地化智能语音识别系统,强调了本地化部署以满足数据安全性和隐私性要求^[15]。该系统实现了终端侧的实时数据接收、多级数据预处理(包括静音检测、端点切割、降噪、数据增强)和关键词唤醒等功能。通过改造 Kaldi 解码框架,实现了并行多路数据解码,提升了实时响应速率并降低了功耗。

3.4 拓展应用

ASR 技术在电力调度中不仅是效率工具,更是主动安全监督、人员行为回溯与智能人机交互的关键支撑。

1) 主动安全监督:高卫恒提出基于智能语音识别技术的电力调度主动安全监督策略,该策略利用 ASR 技术将调度录音转译为文本,并以此作为安全监督分析判据的来源^[16]。通过对转译结果进行分词、

实体提取(设备名称、时间关键字)并与规程制度比对,可实现对调度联系汇报中复诵情况的验证和对“禁止约时停送电”等违规行为的程序化判据分析。目前,商业引擎在电力行业的字识别率约85%,因此需要足够的容错机制,并为电力生产环境建立专属语音数据集,以提高准确率。

2) 人员声纹特征 AI 回溯提取:郭萌提出基于卷积神经网络(CNN)的声纹特征 AI 回溯提取方法^[7]。该方法通过对调度人员语音信号进行预处理(傅氏变换、短时能量与过零率均值分析),提取声纹特征参数,并利用 CNN 训练语音数据,引入连续片段说话人识别模型确定目标说话人身份,实现声纹特征的高精度、快速回溯提取,这为电力调度风险识别和管理规范性提供了重要支持。

3) 智能人机交互:ASR 技术将传统的人机交互方式(如键盘鼠标)转变为更自然的语音交互,大大提升了调度员的工作效率和电网指挥效率。通过语音指令,调度员可以快速完成指令传达、调度日志记录和信息查询,有效解决信息拥堵问题。

4 结语

电力调度语音识别技术经历了从模板匹配、统计模型到深度学习的演进,已在声学模型、语言模型和系统部署三个维度取得显著进展。

在声学模型层面,以 Conformer 为代表的混合架构有效融合了 CNN 的局部特征提取能力与 Transformer 的全局建模优势,DRSN-Conformer 通过动态软阈值降噪机制显著改善了噪声场景下的识别性能;CTDNN-FSA 和 CT-LSTM 等模型则从不同路径强化了对电力调度语音时频特征的捕获能力,推动字错误率持续降低。

在语言模型层面,从传统 N-gram 到 NE-BERT 等领域定制化预训练模型的引入,有效缓解了通用模型对电力专业术语识别不足的问题;大语言模型(LLM)通过 LoRA 等参数高效微调方法与声学模型的融合解码,进一步提高了电力调度场景的识别准确率。

在系统实现层面,迁移学习策略有效应对了电力调度数据稀缺与地域口音适配难题,RNN-Transducer

等流式架构满足了实时调度的低延迟要求,本地化离线部署方案与声纹识别、主动安全监督等拓展应用则充分契合了电力调度专业对数据安全和运行规范的特定需求。

尽管如此,电力调度语音识别技术仍面临若干亟待突破的瓶颈。展望未来,以下方向值得关注:

1) 数据扩充与多模态融合:当前电力调度专用语料库规模有限,应积极探索数据增强、合成语音生成等手段扩充训练数据,同时融合视频监控、操作票文本等多模态信息,增强模型在复杂场景下的泛化能力。

2) 鲁棒性持续提升:针对强噪声、多人同时说话、复杂口音叠加等极端场景,进一步研究噪声自适应、说话人分离与口音无关建模技术,提升系统在全场景下的稳定表现。

3) 轻量化与边缘部署:推进模型压缩、知识蒸馏与高效推理算法研究,实现高性能 ASR 模型在低功耗、低算力的调度终端设备上的稳定运行,兼顾识别精度与部署效率。

4) 安全与可信体系建设:在本地化部署基础上,进一步完善声纹认证、指令合规性校验等安全机制,构建覆盖语音采集、传输、识别、存储全链路的可信语音交互体系,为电力系统稳定运行提供安全可靠技术保障。

参考文献:

- [1] 卢蓬媛.基于迁移学习的电力调度语音识别系统设计[D].北京:华北电力大学,2023.
- [2] 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.
- [3] HINTON G E, DENG L, YU D, et al. Deep neural networks for acoustic modeling in speech recognition[Z]. Modern Speech Recognition, 2012, 29: 82-97.
- [4] 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(03):16-22.
- [5] GRAVES A, MOHAMED A rahman, HINTON G. Speech recognition with deep recurrent neural networks[C/OL]// IEEE International Conference on Acoustics, Speech and Signal Processing, 2013: 6645-6649. (下转第 28 页)

- [8] 杨毅,高翔,朱海兵,等.智能变电站SCD应用模型实例化研究[J].电力系统保护与控制,2015,43(22):107-113.
- [9] 张灿勇.一种智能变电站二次设备配置的高可视化展示方法:CN104239406A[P].2014-07-31.
- [10] 高志远,黄海峰,徐昊亮,等.IEC 61850应用剖析及其发展探讨[J].电力系统保护与控制,2018,46(01):162-169.
- [11] 梅德冬,周斌,黄树帮,等.基于IEC 61850第二版变电站配置描述的集成配置解耦[J].电力系统自动化,2016,40(11):132-136.
- [12] 梅德冬,樊瑞,周斌.IEC 61850模型信息的规则表达与校核研究[J].电力系统保护与控制,2015,43(03):131-136.
- [13] 王智东,王钢,许志恒,等.一种改进的GOOSE报文HMAC认证方法(英文)[J].电网技术,2015,39(12):3627-3634.
- [14] Anwar Adnan, Abdun Naser Mahmood. Cyber security of smart grid infrastructure [M]. The State of the Art in Intrusion Prevention and Detection, CRC Press, Taylor & Fran-

cis Group, USA, 2014: 449-472.

- [15] 盛海华,王德林,马伟,等.基于大数据的继电保护智能运行管控体系探索[J].电力系统保护与控制,2019,47(22):168-175.
- [16] 王智东,王钢,许志恒,等.结合域含义的GOOSE报文加解密方法[J].华南理工大学学报(自然科学版),2016(04):63-70.
- [17] 邓洁清,车勇,单强,等.基于标准中间过程文件的SCD版本比对优化研究[J].电力系统保护与控制,2016,44(14):95-99.
- [18] 高翔,杨漪俊,姜健宁,等.基于SCD的二次回路监测主要技术方案介绍与分析[J].电力系统保护与控制,2014,42(15):149-154.
- [19] 裘愉涛,周震宇,杨剑友,等.继电保护远程运维技术研究与应用[J].电力系统保护与控制,2018,46(18):7-24.
- [20] 文清丰.基于多信息融合的电网故障诊断技术研究[D].保定:华北电力大学,2014.
- [21] 张巧霞,贾伟华,叶海明,等.智能变电站虚拟二次回路监视方案设计及应用[J].电力系统保护与控制,2015,43(10):123-128.

(上接第11页)

1) 台账一致性图像识别校核:针对基础台账与现场实际不一致问题,提出基于YOLO-OBb的旋转目标检测方法,结合OCR文本识别技术,实现了设备铭牌中厂家名称、设备编号等关键信息的自动提取与系统比对,能够有效支撑大规模电网设备的台账自动化校核。

2) 跨系统数据融合校核:在确保台账准确性的基础上,融合PMS3.0系统与用电信息采集系统的电压数据,构建均方根误差、皮尔逊相关系数、标准差等多维度指标,结合时钟偏移校验与时间序列分割技术,实现了电压数据的快速校核与异常类型的精准识别。

参考文献:

- [1] 李颖秋.智能数字化技术对供电可靠性的影响分析[J].集成电路应用,2024,41(12):126-127.
- [2] 刘润霖.配电网电能质量监测与改善技术研究[J].电力设备管理,2026(01):34-36.
- [3] 刘希涛.基于用电信息采集数据的电能质量评估方法分析[J].集成电路应用,2025,42(10):334-335.
- [4] 刘雨,薛琪,刘伟琦,等.电能质量监测数据的多源融合与不确定性建模分析[J].电工技术,2025(09):99-102.
- [5] 王永强,王志强,廖思超,等.基于多模态优化的电力巡检图像识别技术研究[J].自动化与仪表,2026,41(02):160-164.

(上接第16页)

- [6] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[A/OL]. arXiv, 2017. <https://arxiv.org/abs/1706.03762>.
- [7] GULATI A, QIN J, CHIU C C, et al. Conformer: convolution-augmented transformer for speech recognition[A/OL]. arXiv, 2020. <https://arxiv.org/abs/2005.08100>.
- [8] 韩亚旭,高鹭,张飞,等.基于DRSN-Conformer的电力调度语音识别[J].现代电子技术,2026,49(06):112-119.
- [9] 金保含.基于迁移学习和多任务学习的电力调度声学模型研究[D].北京:华北电力大学,2024.
- [10] DEVLIN J, CHANG M W, LEE K, et al. BERT: pre-training of deep bidirectional transformers for language understanding[A/OL]. arXiv, 2018. <https://arxiv.org/abs/1810.04805>.

- [11] 陈蕾,郑伟彦,余慧华,等.基于BERT的电网调度语音识别模型研究[J].电网技术,2021,45(08):2955-2961.
- [12] 闫圣学.基于NE-BERT的电力调度语言模型研究[D].北京:华北电力大学,2024.
- [13] 韩亚旭.面向电力调度场景的语音识别技术研究与应[D].包头:内蒙古科技大学,2025.
- [14] 王亚豪.电力调度系统的流式语音识别技术研究[D].包头:内蒙古科技大学,2023.
- [15] 王邦鸿.基于国网电力调度场景下的智能语音识别系统的实现与优化[D].西安:西安电子科技大学,2024.
- [16] 高卫恒,杨子,王文林.基于智能语音识别技术的电力调度主动安全监督策略研究[J].山西电力,2025(06):30-36.
- [17] 郭萌,徐胜,陈锦龙,等.电力调度系统人员声纹特征AI回溯提取方法[J].微型电脑应用,2025,41(01):134-137.