

• 综 述 •

自动语音分析在老年认知障碍疾病管理中的应用进展*

王颖臻¹, 孟 玲¹, 徐小盼²综述, 莫蓓蓉^{1,3△} 审校

(1. 遵义医科大学珠海校区, 广东 珠海 519041; 2. 南昌大学护理学院, 江西 南昌 330036;

3. 华中科技大学协和深圳医院, 广东 深圳 518052)

[摘 要] 认知障碍患者在早期甚至前驱期即可出现语言功能问题。以人工智能技术为核心的自动语音分析能对丰富的语音数据进行快速处理, 实现认知障碍的自动化识别、评估及持续监测。该文将对自动语音分析的相关概念、发展进程及在认知障碍领域的应用现状进行综述, 以期为今后的临床实践与研究提供参考信息。

[关键词] 自动语音分析; 老年人; 认知障碍; 综述

DOI:10.3969/j.issn.1009-5519.2023.15.019

中图法分类号: R592; R741

文章编号: 1009-5519(2023)15-2619-05

文献标识码: A

Progress in the application of automatic speech analysis in the management of cognitive impairment diseases in the elderly*

WANG Yingzhen¹, MENG Ling¹, XU Xiaopan², MO Beirong^{1,3△}

(1. Zhuhai Campus of Zunyi Medical University, Zhuhai, Guangdong 519041, China; 2. School of Nursing, Nanchang University, Nanchang, Jiangxi 330036, China; 3. Huazhong University of Science and Technology Union Shenzhen Hospital, Shenzhen, Guangdong 518052, China)

[Abstract] Patients with cognitive impairment can have problems with language function in the early or even prodromal period. Automatic speech analysis with artificial intelligence technology as the core can quickly process rich speech data to realize automatic recognition, assessment and continuous monitoring of cognitive impairment. This article will review the related concepts, development process and application status of automatic speech analysis in the field of cognitive impairment, in order to provide reference information for future clinical practice and research.

[Key words] Automatic speech analysis; Elderly; Cognitive impairment; Review

认知障碍是一类以认知功能损害为特征、造成患者日常生活能力减退甚至精神行为异常的脑病综合征^[1], 按严重程度分为轻度认知障碍(MCI)和痴呆 2 个阶段。据统计, 中国 60 岁以上痴呆患者已超过 1 500 万, MCI 患者高达 3 877 万^[2], 这一庞大的患病群体, 将会产生高额的医疗照护费用, 带来沉重的社会和经济负担。因此, 通过早期筛查、评估与护理的综合管理, 对提高患者生活质量和延长生存时间至关重要^[3]。

有研究表明, 语言功能障碍在认知障碍早期甚至前驱期即可出现, 并随疾病发展呈渐进式损害^[4-6], 主要表现为语义和语用处理水平的下降^[7]。在疾病初期患者自发语言基本流畅、无发音错误、遵循基本语法原则, 但句法复杂度明显降低, 语义障碍凸显, 出现找词困难^[8]。中后期时语言产出和理解能力进一步减退, 语音错误增加, 口语流利性显著下降, 停顿次数

增加且时间延长, 语速缓慢, 话语重复, 有效信息量大幅减少, 会话和话题控制能力逐渐降低, 交际能力逐步下降, 甚至丧失交流能力^[9-10]。因此, 语音和语言分析对测量认知功能变化具有重要意义。

但以往的语音数据多由人工分析完成, 如今人工智能的发展已使分析过程进入自动化阶段。自动语音分析不仅解决了人工分析的高成本、低效率等问题, 而且能实现海量数据的快速分析^[11], 捕捉细微语言改变, 早期识别认知障碍, 监测健康状态, 实现对疾病的持续追踪和动态评估^[12]。现将对自动语音分析的概念及发展、应用过程及在老年认知障碍管理中的应用效果综述如下, 以期为老年认知障碍患者的全程管理提供参考和信息支持。

1 自动语音分析

1.1 相关概念 自动语音分析是一种以自动语音识别(ASR)和自然语言处理(NLP)为核心的人工智能

* 基金项目: 广东省深圳市科技创新委员会科技计划项目(JCYJ20210324112207021); 广东省深圳市南山区卫生健康局技术研发和创意设计项目分项资金教育(卫生)科技项目(NS2021057)。

△ 通信作者, E-mail: limo1997@163.com。

技术,通过提取语音数据中的声学 and 语言学特征,结合机器学习分类模型,实现对目标的自动化分析与识别^[11,13]。

1.2 发展过程 自动语音分析主要以 ASR 和 NLP 技术为核心。1956 年,MARVIN MINSKEY、JOHN MCCARTHY 等提出了人工智能的概念,希望“机器可以模仿人类的学习并在其他方面变得智能”^[14]。此后,ASR 和 NLP 作为人工智能的主要研究领域逐步发展。但初期的 ASR 系统仅能通过模板匹配,即将待识别的语音特征与训练中的模板相匹配,处理少量孤立单词,无法应对大规模的词汇识别^[15]。NLP 则通过规则方法处理自然语言,但对文本中的歧义性、模糊性等问题的处理能力较为欠缺,难以应对复杂的语言现象^[16]。20 世纪 70 年代后 ASR 和 NLP 进入以统计建模为基础的时期。通过隐马尔科夫模型

(HMM),ASR 可处理大量词汇和连续语音^[15],但识别速率和准确性尚未达到实际应用水平。NLP 利用 HMM 解决了词性标注中一词多义的问题,令其能根据上下文进行正确的选择^[17]。但这些技术并未广泛用于医疗领域。21 世纪以来,深度学习的出现使 ASR 和 NLP 等人工智能技术进入了全新的发展阶段,推动了其在医疗领域的深度结合和广泛应用,大量学者利用这些技术对临床实践的不同步骤进行了自动化研究,如自动筛查阿尔茨海默病(AD)^[18]、帕金森病^[19]、心力衰竭^[20]等。

2 自动语音分析在老年认知障碍人群中的应用方法
自动语音分析现暂无标准化应用方法,根据现有研究^[11-12,21-24]、总结自动语音分析的应用方法,共分为准备、实施和目标识别 3 个阶段。见表 1。

表 1 自动语音分析的应用方法

阶段	具体内容
准备阶段	• 将受试者安排在隔音或噪声小于 100 dB 的较为安静的环境内 • 将语音输入终端(如麦克风)置于受试者嘴部正前方 5~10 cm 处 • 设置语音数据采集频率为 20~16 000 Hz • 由操作医师或护士给予受试者发音指导并示意测试开始
实施阶段	• 由操作医师或护士指导受试者以舒适的发音状态完成语音任务,并进行语音数据采集 • 语音任务主要为图片描述测试和语言流畅性测试 • 图片描述测试要求受试者在固定时间(多为 1 min 内)描述图片中看到的所有内容,最常用的图片描述测试为波士顿诊断性失语症测验中的“偷吃饼干图” • 语言流畅性测试包括语义流畅性和音位流畅性任务,要求受试者在 1 min 内尽可能多说属于同一类别或以同一字母开头的词语
目标识别阶段	• 使用自动语音分析系统提取语音特征,通过系统内部机器学习模型分类器,自动完成目标识别 • 语音特征提取主要为声学、语言学或其他副语言特征 • 声学特征,如梅尔频率倒谱系数、梅尔滤波器组和感知线性预测系数等 • 语言学特征,如总词汇量、词性数量与比例(名词、动词、形容词等)、语法和句法等 • 副语言特征,如语速、停顿频率、沉默持续时间等

3 自动语音分析在老年认知障碍管理中的应用效果
3.1 健康状态监测 ROBIN 等^[25]基于移动应用程序开发了一款自动语音分析系统 Winterlight Assessment,通过自动记录和分析参与者完成图片描述测试的自发语音数据,结果显示,无认知障碍组、MCI 组和 AD 组参与者语音数据在语言连贯性、信息丰富度、句法复杂度、找词能力方面得分比较,差异均有统计学意义($P<0.05$)。随后在 6 个月的认知功能监测中发现,与无认知障碍组比较,MCI 组和 AD 组患者语言连贯性得分随疾病进程大幅下降,表明自动语音分析有可能是老年认知功能长期监测的一个重要手段。LAURAITIS 等^[26]则基于自我管理老年认知检查开发了一款移动应用程序以监测神经退行性疾病患者的健康状态。通过提取音高、梅尔频率倒谱系数、

Gammatone 倒谱系数和小波散射变换等声学特征,结合支持向量机和双向长短期记忆神经网络分类器进行自动分析,评估认知障碍患者大脑神经受损程度和语言功能状态,实现对其健康状态的日常监测,便于医务人员及时了解疾病信息,随时掌握疾病进展,为治疗和护理提供全面的信息支持。
3.2 识别与划分认知障碍严重程度
3.2.1 MCI 的识别 MCI 患者认知功能改变较轻微,极易被忽视,故仍有相当大比例的患者未被发现。自动语音分析可捕捉因认知功能异常引起的语言细微改变,及早识别 MCI 患者。TOTH 等^[27]将 ASR 与机器学习的随机森林模型分类器结合,自动识别 MCI,获得了 78.8% 的准确率,并通过对语音数据的自动分析发现,健康组和 MCI 组受试者自发语音在

语速和回答问题时的停顿次数比较,差异均有统计学意义($P < 0.05$)。KÁLMÁN 等^[28]基于语音的时间特征研发了匈牙利语版 Speech-Gap Test[®] (S-GAP Test[®])筛查工具,并进行跨语言检测,结果显示,该工具也同样适用于英语语种,且对健康组和 MCI 组受试者语音速度和发音速度的区分灵敏度达到 100%,但该研究暂未拓展至其他语种。此外,自动语音分析还可配合神经心理学量表使用,将部分评分过程自动化,减少由人为因素造成的结果偏差,以提高结果的可靠性。泰国朱拉隆功大学的学者使用基于时延神经网络隐马尔可夫模型的 ASR 技术,开发了蒙特利尔认知评估量表中语言流利度测试的自动评估系统,结果显示,与传统手动评分比较,其正确率达到了 93%,表明了该系统在临床应用中的巨大潜力^[29]。但由于泰语以声调区分词义的特殊性,特定语境以外的独立单词发音增加了音调检测的难度,可能会改变单词含义并降低评分准确性。

3.2.2 AD 的识别 大量研究表明,自动语音分析已广泛用于 AD 患者的识别,且具有良好的识别准确率^[18,21,30]。CLARKE 等^[31]通过图片描述、对话、叙述性回忆、程序性回忆和小说叙事重述 5 种语音任务识别 AD 患者,共提取 286 种语言学特征,结合支持向量机模型,获得了 90% 的识别准确率。SANZ 等^[32]则着重于对语义的分析,将自发语音数据和机器学习的梯度提升模型分类器结合,自动评估参与者的语义粒度和持续语义变异性,正确识别了 74% 的 AD 患者。同样,研究人员将自动语音分析与经典痴呆筛查量表结合,获得了更高的筛查准确率。SADEGHIAN 等^[33]收集 72 例参与者的资料,从中提取声学、语言学、人口统计学及精神状态评估表(MMSE)评分共 236 个特征,运用 ASR 和多层感知神经网络自动识别 AD 患者,获得 94.4% 的识别准确率,将 MMSE 评分的准确率提高了 23.6%。

3.2.3 划分认知障碍严重程度 GOSZTOLYA 等^[22]采用 ASR 和机器学习区分 MCI 和轻度 AD 患者,自动分析即刻回忆与延迟回忆语音数据中的声学特征,其准确率为 74%~82%,在融合了语言学特征后获得了 80%~86% 的识别准确率。上海瑞金医院的学者 QIAO 等^[34]自行研发了 ASR 软件,分析“偷吃饼干图”描述任务的自发语音数据,确定了 7 个与认知功能相关的语音参数,结果显示,静音段的平均持续时间最有助于区分 AD、MCI 患者和健康者,但其未报道区分的准确率。哥伦比亚大学的研究小组基于跨任务的多模态语音和眼动数据集结合机器学习,区分了轻中度 AD、MCI 及主观记忆抱怨患者,从图片描述、阅读、记忆描述和瞳孔校准任务中提取声学、语言学及眼球运动特征,使用 3 种机器学习算法训练分类准确度,得到总体最高的平均受试者工作特征曲线下面积为 0.83 ± 0.01 ^[35]。

3.3 识别认知障碍相关神经精神症状(NPS) NPS 可见于认知障碍的不同时期^[36],常表现为情绪/情感症状(焦虑、抑郁、淡漠等)、精神症状(激越、幻觉、妄想等)和行为改变^[37],可能会加速认知功能的恶化。据统计,98% 的认知障碍患者曾出现过 NPS^[38],若能尽早识别并干预可有效减轻 NPS 严重程度,甚至延缓认知功能下降^[39]。法国国家信息与自动化研究所的学者 KÖNIG 等^[39]采集了 141 例 65 岁以上认知障碍患者的语音数据,采用自动语音分析工具,自动化提取其中的时间、韵律、共振峰、声源及频谱等特征,结合支持向量和 Lasso 2 种回归模型实现 NPS 的自动评估,结果显示,自动语音分析在评估焦虑、淡漠和抑郁得分方面表现良好。淡漠的言语比率较低,焦虑的声噪比较高,但由于样本量较小,暂未发现抑郁的显著语音特征。

3.4 其他 自动语音分析除用于上述领域外,也被用于构建基于语音特征的 AD 风险预测算法、辅助痴呆患者的语言沟通等方面。GREGORY 等^[40]开展了一项前瞻性观察研究,利用电话收集有 AD 风险的参与者的语音数据,通过自动语音分析提取语音特征,结合 AD 生物标志物,寻找能预测 AD 风险的语音特征,最终开发了一个基于语音特征预测 AD 风险的算法,目前,该研究正在进行中。有学者开发了一款人工智能通信工具,基于记录患者的个人语音数据,使用人工智能相关技术自动分析和识别唯一的语言标识符,帮助用户将被认为是难以理解的语音翻译成可理解的对话,辅助痴呆患者的语言沟通^[41-42]。

4 小结与展望

自动语音分析能有效减少传统量表因受试者主观因素造成的影响,使结果更加准确、客观,实现快速分析和应对大规模人群筛查,为早期干预提供机会。然而,现阶段研究均为横断面设计,缺乏能反映在认知障碍进展过程中患者语音特征变化的长期追踪数据。故后续研究有必要考虑采用前瞻性队列设计,以观察认知功能的衰退对语言功能的影响^[43]。目前的机器学习分类模型尚未达到理想性能,因此,亟须开发更具鲁棒性的深度学习语音分析模型,并进一步优化模型分类能力,以实现血管性痴呆、额颞叶痴呆、路易体痴呆等认知障碍亚型的临床鉴别。此外,针对汉语言特征的语音分析研究尚不多见,将会影响语音分析技术在我国临床适用性。要求今后的研究不仅应该对已有的自动语音分析工具进行跨语种调试,还需结合汉语言特征构建适合我国人群的自动语音分析算法。需建立自动语音分析的应用标准和评价体系,构建规范化管理的语音大数据库,完善人工智能技术的监管和个人隐私保护,为实现自动语音分析技术的市场化奠定制度基础。

参考文献

[1] 中国老年学和老年医学学会脑认知与健康分会,

- 中国老年医学学会认知障碍分会,中国医师协会神经内科医师分会认知障碍疾病专业委员会,等.中国数字化记忆门诊临床应用专家共识(2021)[J].中华医学杂志,2021,101(45):3712-3717.
- [2] JIA L, DU Y, CHU L, et al. Prevalence, risk factors, and management of dementia and mild cognitive impairment in adults aged 60 years or older in China: A cross-sectional study[J]. *Lancet Public health*, 2020, 5(12): e661-e671.
- [3] 汪凯,董强,郁金泰,等.卒中后认知障碍管理专家共识 2021[J].中国卒中杂志,2021,16(4):376-389.
- [4] MARTÍNEZ-NICOLÁS I, LLORENTE T E, IVANOVA O, et al. Many changes in speech through aging are actually a consequence of cognitive changes[J]. *Int J Environ Res Public Health*, 2022, 19(4): 2137.
- [5] AHMED S, HAIGH A M, DE JAGER C A, et al. Connected speech as a marker of disease progression in autopsy-proven Alzheimer's disease[J]. *Brain*, 2013, 136(Pt 12): 3727-3737.
- [6] FORBES-MCKAY K E, VENNERI A. Detecting subtle spontaneous language decline in early Alzheimer's disease with a picture description task[J]. *Neurol Sci*, 2005, 26(4): 243-254.
- [7] FERRIS S H, FARLOW M. Language impairment in Alzheimer's disease and benefits of acetylcholinesterase inhibitors[J]. *Clin Interv Aging*, 2013, 8: 1007-1014.
- [8] 李妍,周爱红,赵丽娜,等.轻度阿尔茨海默病患者自发语言特点研究[J].中华神经科杂志,2019,52(3):177-183.
- [9] 刘红艳.阿尔茨海默症患者语言障碍研究现状和进展:基于病理语言学的实验研究综述[J].外语电化教学,2020(5):72-78.
- [10] 黄立鹤,杨晶晶.汉语阿尔茨海默病患者口语非流利性研究[J].当代语言学,2022,24(2):192-207.
- [11] PETTI U, BAKER S, KORHONEN A. A systematic literature review of automatic Alzheimer's disease detection from speech and language[J]. *J Am Med Inform Assoc*, 2020, 27(11): 1784-1797.
- [12] KÖNIG A, SATT A, SORIN A, et al. Automatic speech analysis for the assessment of patients with predementia and Alzheimer's disease[J]. *Alzheimers Dementia (Amst)*, 2015, 1(1): 112-124.
- [13] 王结群,张卫,郑晓然,等.认知评估新技术在痴呆诊断中的探索与进展[J].重庆医科大学学报,2021,46(11):1296-1301.
- [14] 李晓理,张博,王康,等.人工智能的发展及应用[J].北京工业大学学报,2020,46(6):583-590.
- [15] 马晗,唐柔冰,张义,等.语音识别研究综述[J].计算机系统应用,2022,31(1):1-10.
- [16] 程美,王力华.医疗智能语音技术与应用综述[J].中国数字医学,2021,16(8):1-6.
- [17] 韩普,姜杰. HMM 在自然语言处理领域中的应用研究[J].计算机技术与发展,2010,20(2):245-248.
- [18] LIU N, LUO K, YUAN Z, et al. A transfer learning method for detecting Alzheimer's disease based on speech and natural language processing[J]. *Front Public Health*, 2022, 10: 772592.
- [19] TAI Y C, BRYAN P G, LOAYZA F, et al. A voice analysis approach for recognizing Parkinson's disease patterns[J]. *IFAC Papers On Line*, 2021, 54(15): 382-387.
- [20] AMIR O, ABRAHAM W T, AZZAM Z S, et al. Remote speech analysis in the evaluation of hospitalized patients with acute decompensated heart failure[J]. *JACC Heart Fail*, 2021, 10(1): 41-49.
- [21] LI R, WANG X, LAWLER K, et al. Applications of artificial intelligence to aid early detection of dementia: A scoping review on current capabilities and future directions[J]. *J Biomed Inform*, 2022, 127: 104030.
- [22] GOSZTOLYA G, VINCZE V, TÓTH L, et al. Identifying mild cognitive impairment and mild Alzheimer's disease based on spontaneous speech using ASR and linguistic features[J]. *Comput Speech Lang*, 2019, 53: 181-197.
- [23] 宋歌,王淼,高中宝,等.人工智能语音分析系统在帕金森病诊断中的一项探索性临床研究[J].中华老年心脑血管病杂志,2020,22(5):514-519.
- [24] WANG T, HONG Y, WANG Q, et al. Identification of mild cognitive impairment among Chinese based on multiple spoken tasks[J]. *J Alzheimers Dis*, 2021, 82(1): 185-204.
- [25] ROBIN J, XU M, KAUFMAN L D, et al. Using digital speech assessments to detect early signs of cognitive impairment[J]. *Front Digit Health*, 2021, 3: 749758.

- [26] LAURAITIS A, MASKELIÜNAS R, DAMA SEVICIUS R, et al. A mobile application for smart computer-aided self-administered testing of cognition, speech, and motor impairment [J]. *Sensors*(Basel), 2020, 20(11):3236.
- [27] TOTH L, HOFFMANN I, GOSZTOLYA G, et al. A speech recognition-based solution for the automatic detection of mild cognitive impairment from spontaneous speech[J]. *Curr Alzheimer Res*, 2018, 15(2):130-138.
- [28] KÁLMÁN J, DEVANAND D P, GOSZTOLYA G, et al. Temporal speech parameters detect mild cognitive impairment in different languages: Validation and comparison of the Speech-GAP Test® in English and Hungarian[J]. *Curr Alzheimer Res*, 2022, 19(5):373-386.
- [29] KANTITHAMMAKORN P, PUNYABUKKANA P, PRATANWANICH P N, et al. Using automatic speech recognition to assess thai speech language fluency in the montreal cognitive assessment (MoCA) [J]. *Sensors* (Basel), 2022, 22(4):1583.
- [30] LIU N, YUAN Z, TANG Q. Improving Alzheimer's disease detection for speech based on feature purification network [J]. *Front Public Health*, 2021, 9:835960.
- [31] CLARKE N, BARRICK T R, GARRARD P. A comparison of connected speech tasks for detecting early Alzheimer's disease and mild cognitive impairment using natural language processing and machine learning[J]. *Front Comput Sci*, 2021, 3:634360.
- [32] SANZ C, CARRILLO F, SLACHEVSKY A, et al. Automated text-level semantic markers of Alzheimer's disease [J]. *Alzheimers Dement* (Amst), 2022, 14(1):e12276.
- [33] SADEGHIAN R, SCHAFFER J D, ZAHORIAN S A. Towards an automatic speech-based diagnostic test for Alzheimer's disease [J]. *Front Comput Sci*, 2021, 3:624594.
- [34] QIAO Y, XIE X Y, LIN G Z, et al. Computer-assisted speech analysis in mild cognitive impairment and Alzheimer's disease: A pilot study from Shanghai, China[J]. *J Alzheimers Dis*, 2020, 75(1):211-221.
- [35] JANG H, SOROSKI T, RIZZO M, et al. Classification of Alzheimer's disease leveraging multi-task machine learning analysis of speech and eye-movement data[J]. *Front Hum Neurosci*, 2021, 15:716670.
- [36] 王华丽. 神经认知障碍精神行为症状群临床诊疗专家共识[J]. *中华精神科杂志*, 2017, 50(5):335-339.
- [37] 赵程, 于会艳, 秦斌. 轻度行为损害与痴呆相关研究的共识与争议[J]. *中华神经科杂志*, 2018, 51(7):551-554.
- [38] PHAN S V, OSAE S, MORGAN J C, et al. Neuropsychiatric symptoms in dementia: Considerations for pharmacotherapy in the USA [J]. *Drugs R D*, 2019, 19(2):93-115.
- [39] KÖNIG A, MALLICK E, TRÖGER J, et al. Measuring neuropsychiatric symptoms in patients with early cognitive decline using speech analysis[J]. *Eur Psychiatry*, 2021, 64(1):e64.
- [40] GREGORY S, LINZ N, KÖNIG A, et al. Remote data collection speech analysis and prediction of the identification of Alzheimer's disease biomarkers in people at risk for Alzheimer's disease dementia: The Speech on the Phone Assessment (SPeAk) prospective observational study protocol[J]. *BMJ Open*, 2022, 12(3):e052250.
- [41] SU Z, BENTLEY B L, MCDONNELL D, et al. 6G and artificial intelligence technologies for dementia care: Literature review and practical analysis[J]. *J Med Internet Res*, 2022, 24(4):e30503.
- [42] MCCOY K F, DEMASCO P W, JONES M A, et al. A communication tool for people with disabilities: Lexical semantics for filling in the pieces [C/OL]//. *Proceedings of the first annual ACM conference on Assistive technologies* October, 1994, California USA (1994-11-03) [2022-11-23]. <https://dl.acm.org/doi/abs/10.1145/191028.191058>.
- [43] 周雁, 梁玥, 袁晶, 等. 数字化语言标志物区分阿尔茨海默病和轻度认知功能障碍的疾病效应与老化效应初探[J]. *中华神经科杂志*, 2020, 53(5):341-347.

(收稿日期:2022-12-14 修回日期:2023-03-22)