

语音重复任务在轻度认知功能障碍检测中的应用[☆]殷潇潇* 王思文* 王贺* 高琳琳** 任智* 王钦文***[⊗]

【摘要】轻度认知功能障碍(mild cognitive impairment, MCI)通常被视为痴呆的前驱阶段,其主要特征为认知功能轻度下降。研究表明,MCI患者中语言变化可能先于其他认知功能症状,这为早期识别和干预提供了机会。MCI患者语言特点包括语速、发音和语调等异常。五个单词测验、数字延迟匹配测试和句子重复测试等语音重复任务,是评估MCI患者语言特点的有效方法,这些任务要求患者重复特定内容,分析重复准确性,从而评估其语言功能。机器学习和深度学习技术的应用,能自动提取语音重复任务数据中的MCI相关特征,提高诊断准确性。这些技术的结合应用有助于早期发现MCI,为及时干预提供依据。

【关键词】认知功能障碍 任务重复 阿尔茨海默病 语言 语音识别 机器学习 深度学习

【中图分类号】R749.1

【文献标识码】A

Application of speech repetition task in the detection of mild cognitive impairment. YIN Xiaoxiao, WANG Siwen, WANG He, GAO Linlin, REN Zhi, WANG Qinwen. Ningbo University School of Medicine, Ningbo 315211, China. Tel: 0574-87608922.

【Abstract】 Mild cognitive impairment (MCI) is often regarded as a prodromal stage of dementia, primarily characterized by mild decline in cognitive function. Due to the mild nature of its symptoms, many MCI cases miss the opportunity for intervention. Research indicates that linguistic changes in MCI may precede other cognitive symptoms, providing an opportunity for early identification and intervention. These linguistic changes include abnormalities in speech rate, pronunciation, and intonation. Speech repetition tasks, such as the five-word test, delayed digit matching, and sentence repetition tests, are effective methods for assessing the speech characteristics of MCI patients. These tasks involve asking patients to repeat specific content, and analyzing the accuracy of their repetitions to assess their speech function. The application of machine learning and deep learning techniques enables automatic extraction of MCI-related features from speech data, improving diagnostic accuracy. The combined application of these techniques contributes to the early detection of MCI, providing a basis for timely intervention.

【Keywords】 Mild cognitive impairment Task repetition Alzheimer disease Language Speech repetition Machine learning Deep learning

轻度认知功能障碍(mild cognitive impairment, MCI)是痴呆前的过渡状态,表现为记忆、语言等认知功能改变。随着我国老龄化,MCI和阿尔茨海默病(Alzheimer disease,

AD)的患病率不断上升,早期筛查至关重要。开发便捷、高效的MCI早期筛查方法成为迫切需求。语言作为认知功能缺陷的早期迹象,为MCI的识别提供了新的途径。近年来,机器学习和深度学习在MCI和AD检测中发挥了重要作用。语音重复任务记录患者语言数据,提取特征,将其作为机器学习模型的输入,可实现对疾病的早期识别。本文旨在回顾近年来利用语音重复任务检测MCI的研究,总结MCI患者的语言特点、常用语音重复任务类型及机器学习和深度学习的应用方法。

1 MCI语言特点

MCI的语言特点主要表现为多方面的语言功能下降。

doi:10.3969/j.issn.1002-0152.2024.04.010

* 国家自然科学基金项目(编号:32171035);浙江省“尖兵”“领雁”研发攻关计划(编号:2024C03101);宁波市“科技创新2025”重大专项(编号:2019B10034)

* 宁波大学教师教育学院(宁波315211)

** 宁波大学信息科学与工程学院

*** 宁波大学医学院

[⊗] 通信作者(E-mail:wangqinwen@nbu.edu.cn)

随着衰老过程,老年人喉和声道的解剖结构以及生理功能发生变化,导致发音困难,如音域缩小、基频变化等^[1]。此外,频率、分贝和振幅的变化更为明显,共振减少,语音停顿更为频繁。认知功能障碍程度的发展过程中,语言的多个层面受到影响,通常先从语义学和其应用处理水平上开始,逐渐波及语法和音韵学^[2]。患者早期命名困难是最明显的缺陷,难以回忆不常见的物品和专有名称,动词的命名相较于名词更加困难^[3]。此外,语言的抽象性、隐喻、讽刺等高级语言功能也受损,语言流畅性困难^[4]、停顿^[5]在MCI早期也很明显。一般来说,阅读理解和书面表达比听觉理解和言语能力的退化更快^[6],当发展为AD时语句简短、不完整,自言自语减少^[7]。进一步研究认知功能障碍患者在隐喻和谚语处理方面的表现,结果表明,患者对隐喻所要传递的意义存在误解,并且难以理解谚语所要表达的信息^[8]。还有研究表明在MCI阶段时,患者的推论总结性语言变少^[9]。

在MCI和AD的语言检测研究中,大多数研究对词性变化、词汇丰富度、句法复杂性和信息量等加以区分,只有少数研究从心理语言学角度研究了MCI语言,发现其语义障碍可能表现为对非常熟悉单词的依赖增加。也有研究者对MCI的语言特点进行研究时,忽略掉语音长度和沉默片段,专注于完成任务时语音韵律的变化^[10]。

综上所述,MCI的语言特点是多维度的,涵盖了发音、命名、语法、音韵学以及高级语言功能等多个方面。这些特点不仅提供了早期识别和评估认知功能障碍的重要线索,也为深入理解认知功能障碍的病理机制和制定针对性的干预措施提供了重要参考依据,有关研究中常用的语言特征归纳总结见表1。

2 语言检测任务

语言变化是认知功能缺陷的早期迹象,多种语言检测任务广泛应用于区分MCI与AD,其中语音重复任务在筛查中尤为重要。目前常用的语音重复任务:五个单词测验

(five word test)和数字延迟匹配测试(digit span test),分别通过回忆和重复词语或数字,评估短期记忆和工作记忆能力;句子重复测试(sentence repetition test),通过复述难度递增的句子,考察短期工作记忆;Wechsler逻辑记忆测验(Wechsler logical memory I/II),要求受试者复述简短故事,其即时和延迟复述分数反映逻辑认知能力和短期工作记忆,该测试有多个变式,如回忆短片内容^[11]、回忆梦境等。以上语音重复任务不仅反映中老年人对于信息的记忆和加工能力,还可揭示思维方式和认知能力的变化,甚至体现情感上的反应,并有研究进一步表明采用多项语言任务融合进行认知功能的检测,准确率更高^[13]。

3 语言分析技术

语言的分析可以为认知功能筛查提供有价值的信息。语音分析初期依赖于手动转录,依据一定规则来区分话语,但对于零碎的话语则依赖前后语义内容和是否停顿来判断。在细节标注上,需区分受试者声音与其他声音,记录单词和停顿的时间,以计算相关指标。在句子重复任务中,还需去除填充词并对齐语句,为确保转录一致性,通常由同一研究者完成转录工作。为了克服复杂的工作和评分者差异,采用自动化技术和定量指标分析复杂语言成为大规模语言筛查的必要需求。为了应对这一挑战,自然语言处理技术应运而生,并已有研究表明计算机分析语言与临床医生评估具有很大的相似性^[14],这也证明了自然语言处理技术的可靠性。

3.1 机器学习 机器学习是人工智能的一个重要分支,专注于从数据中自动构建统计模型。这一过程中,模型的迭代优化是提高其整体准确性的关键。在众多分类器中,线性分类器因其简洁性和有效性而在自然语言处理中受到广泛应用,常见的有支持向量机(support vector machines, SVM)和随机森林分类器(random forest classifier, RF),其中SVM尤为突出。

Tab.1 Summary of common language features

表1 常用语言特征的归纳总结

特征类型	语言特征
词汇特征	各类词的占比、词频、重复的次数、删除的词数、主题密度、特定标签使用频率、填充暂停
声学特征	周期段和非周期段的持续时间、发声速率、暂停率、总暂停时间、平均暂停时间、犹豫比、总发声时间、总语音时间、最大音速、Mel频率倒谱系数(MFCC)特性
句法特征	词汇多样性、词汇密度、句子长度、话语标记物、命题密度、主题密度、局部相关性、整体相关性、平均深度、总深度、句法错误、语言的有效性指数、信息量
心理语言学特征	具体性、熟悉度、可想象性、词汇习得年龄

SVM的基本思想是在特征空间中找到一个超平面,将不同类别的样本分开,并使同一类别的样本尽可能聚集在一起。N折交叉验证法常被用来调整SVM分类器的参数^[12-13,15-17],例如FRASER等^[12]的经典研究使用SVM分类器,并用十折交叉验证调整分类器对AD和健康个体进行区分,准确率达到81%。GOSZTOLYA等^[18]采用SVM基于声学特征区分MCI与健康对照,实现了78%的检测准确率,对AD的检测准确率更是高达82%,而在AD与MCI混合样本中检测正确率也达到了76%。另外,JIANG等^[17]利用五折交叉验证调整了SVM,在区分MCI与健康个体时,准确率达到88.8%±3.59%,展现出良好的性能。

RF是基于多个决策树的集成,也是可用于自然语言处理的强大工具。POMPILI等^[19]结合随机森林和基于序列前向选择(sequential forward selection, SFS)的方法,在区分MCI、AD患者和健康对照时准确率达到82.6%±5.1%。SHIMODA等^[20]通过电话访谈记录AD和MCI患者语音样本中的韵律、沉默等声学特征,使用基于梯度提升(eXtreme gradient boosting, XGBoost)、RF和逻辑回归(logistic regression, LR)的三种机器学习模型联合检测,准确率为90%。LIU等^[21]使用CMUSpinx和Mozilla工具包制作的语音识别方法,对AD检测的准确率可达92.72%。

3.2 深度学习 近年来,深度学习中的卷积网络在成像、视频、语音和音频处理方面取得很大进步。目前,机器学习和深度学习相结合的方法成为老年人认知功能水平评估分析的主流方法。基于深度学习的数据分类模型不断涌现,但需要注意的是,通过使用不同的算法,相同的向量特征可能得到不同的分类结果。

与机器学习相比,深度学习具有高度自适应的优点。深度学习模型通过大量数据训练,具有对不同语言和语境高度自适应能力,从而提高模型性能和泛化能力。鉴于老年人语言特点,深度学习能够有效捕捉到其语言中的隐含模式和规律。此外,深度学习在处理文本中上下文信息,理解语义和语境关系方面表现出色,这对检测MCI患者的语言特征至关重要,因为患者可能在表达中出现断片化、缺乏连贯性的情况。例如Transformer模型应用多头注意机制(multi-head attention),在处理这类问题时显示出良好的效果^[22]。GULATI等^[23]将擅长捕捉全局特征的Transformer和擅长捕捉局部特征的卷积神经网络(convolution neural networks, CNNs)结合,提出算法模型Conformer,将其应用于检测AD,准确率达91.7%。

在老年人语音研究中,声学特征与语义特征同等关键。

除了机器学习,还有多种语音特征提取工具。如Kaldi^[24]是自动语音识别工具包,但对非技术用户而言操作困难;Praat^[25]是广受欢迎的语音分析软件,适用于多种分析;OpenSMILE^[26]专注于音频信号,用法简单但配置难;Shen-nong^[27]是基于Kaldi的Python工具箱,提供多种算法且易于使用。这些工具在不同应用中取得良好结果,如王荃^[15]用Praat处理语音,MCI和健康人群分类准确率达95%,LI等^[28]用OpenSMILE检测MCI,显示良好的检测结果。

近些年的研究表明,声学特征和语音特征结合可以提高MCI检测的准确率。ROARK等^[29]将老年人语音中的语音声学特征、语义特征和不同认知测试分数结合起来进行认知功能障碍的检测,筛查性能提升4%。YAMADA等^[30]研究使用机器学习分析表征声学、韵律和语言方面的语音特征,对AD、MCI患者和认知功能正常的对照进行分类,结果表明,区分AD和对照的准确率可达91.2%,区分MCI和对照组的准确率可达87.6%。另外,其他深度学习技术也在检测语音性能方面表现良好。例如MAHAJAN等^[31]发现前人研究^[32-33]存在一定问题,并在其基础上做出改进,提出Speech-GRU,其检测AD的准确性提升6.25%。

认知功能障碍检测中常用的算法模型总结见表2。语言分析技术在MCI检测中的应用重要性日益凸显。机器学习与深度学习的结合为识别MCI患者的语言特征提供了强大支持。SVM、随机森林等机器学习算法,在结合声学、语义特征后,可达到更高的检测准确率。而深度学习则以其高度自适应性和处理复杂语境的能力,为捕捉老年人语言中的隐含模式提供了新途径。声学特征提取工具的发展也极大地推动了这一领域的研究进展。这些技术不仅简化了MCI的检测流程,还为老年人的内心状态和思维方式研究提供了新视角。未来,随着技术的不断完善,语言分析技术将在MCI检测及老年健康管理领域发挥更大作用。

4 总结

传统的MCI检测方法,如MRI和脑脊液分析昂贵或具有侵入性,量表评估耗时且主观。老年群体的语言特征为MCI检测提供了新线索。计算机自然语言识别和自动文本分析技术可深入探索语言特征,提高检测效率并提供客观标准。目前研究多关注声学、语义特征,但情绪情感、词汇习得年龄等心理语言学特征仍需深入。结合多维度语言特征,算法模型能更精确辨别MCI患者,预测认知功能水平,有助于早期干预。此技术简化了检测流程,为老年人的内心状态和思维方式研究提供了新视角,具有广泛的应用

Tab.2 Summary of commonly used algorithm models in cognitive impairment detection
表2 认知功能障碍检测中常用的算法模型总结

工具	区分人群	准确率	优点	缺点
SVM	MCI和AD	92.1% ^[34]	处理大型特征空间和非线性特征之间的相互作用表现好;无需依赖整个数据	计算复杂度高,核函数难找;对缺失数据、噪声敏感
	MCI和健康人群	95% ^[15]		
RF	MCI和AD	91.9% ^[34]	简单易实现;处理高维数据;非线性特征能力强;直观的决策规则	处理不平衡数据能力有限;在噪声较大的问题中易过拟合;处理不同取值的属性数据能力不强
	MCI和健康人群	96% ^[15]		
LR	MCI和AD	90.8% ^[34]	计算代价不高,易理解实现;对数据中小噪声的鲁棒性好	容易欠拟合,分类精度不高;对缺失数据敏感;特征空间大时影响其分类;对于非线性需要转换
	AD和健康人群	89.3% ^[20]		
	MCI和健康人群	95% ^[15]		
XGBoost	AD和健康人群	86.3% ^[20]	简单易用;高效可扩展性;鲁棒性强	无法对时空位置建模;捕捉图像语音文本等高维数据能力有限
BERT	AD和非AD人群	88% ^[28]	有效捕捉上下文的关系;可以在大规模数据上进行预训练;支持多种任务,如文本分类、命名实体识别等	计算资源消耗大;存在学习不足问题;处理超过512个标记的文本困难;预训练数据集有限

注:SVM,支持向量机;RF,随机森林分类器;LR,逻辑回归;XGBoost,基于梯度提升;BERT,双向变换器模型;MCI,轻度认知功能障碍;AD,阿尔茨海默病。

前景。

尽管自动语音识别技术在检测MCI方面具有潜力,但在国内发展仍面临多重挑战。首先,深度学习对数据量的需求庞大,目前可用的中文语音数据库规模有限,例如针对普通话的Mandarin Lu数据库(<https://dementia.talkbank.org/access/>)仅包含52例样本,这限制了模型的训练效果。其次,现有的网络模型主要基于英语数据训练,在识别包括普通话在内的其他语言时性能可能受限。考虑到国内老年人语言中的方言特点,模型在本地语言识别方面面临额外挑战。此外,将检测程序转化为实用且普及率高的产品也是一项艰巨任务。值得注意的是,不同文化背景下认知功能障碍患者的语言特征可能存在差异,这为研究提供了丰富的对比材料和深入探索的空间。未来研究应致力于构建大规模的认知功能障碍中文语音数据库,并探索多模态MCI检测方法,结合语义、声学及医学图像特征,以实现更精准的MCI早期筛查,并进一步对不同文化背景下的患者语言特征进行对比分析。

参 考 文 献

- [1] ALONSO J B, DELENO J, ALONSO I, et al. Automatic detection of pathologies in the voice by HOS based parameters[J]. EURASIP J Appl Sig P, 2001, 4: 275-284.
- [2] KIRSHNER H S. Primary progressive aphasia and Alzheimer's disease: Brief history, recent evidence[J]. Curr Neurol Neurosci Rep, 2012, 12: 709-714.
- [3] ROBINSON K M, GROSSMAN M, WHITE-DEVINE T, et al. Category-specific difficulty naming with verbs in Alzheimer's disease[J]. Neurology, 1996, 47(1): 178-182.
- [4] WRIGHT L M, DE MARCO M, VENNERI A. Verbal fluency discrepancies as a marker of the prehippocampal stages of Alzheimer's disease[J]. Neuropsychology, 2023, 37(7): 790-800.
- [5] ANDRADE E I N, MANXHARI C, SMITH K M. Pausing before verb production is associated with mild cognitive impairment in Parkinson's disease[J]. Front Hum Neurosci, 2023, 17: 1102024.
- [6] APPEL J, KERTESZ A, FISMAN M. A study of language functioning in Alzheimer patients[J]. Brain Lang, 1982, 17(1): 73-91.
- [7] 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: 243-254.
- [8] PAPAGNO C. Comprehension of metaphors and idioms in patients with Alzheimer's disease: A longitudinal study[J]. Brain, 2001, 124(7): 1450-1460.
- [9] MAZIERO M P, BELAN A F R, CAMARGO M Z A, et al. Textual inference comprehension in mild cognitive impairment: The influence of semantic processing and verbal episodic memory[J]. Front Aging Neurosci, 2021, 13: 735633.
- [10] HIGUCHI M, NAKAMURA M, OMIYA Y, et al. Discrimination of mild cognitive impairment based on involuntary changes caused in voice elements[J]. Front Neurol, 2023, 14: 1197840.
- [11] 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.
- [12] FRASER K C, MELTZER J A, RUDZICZ F. Linguistic features identify Alzheimer's disease in narrative speech[J]. J Alzheimers Dis, 2016, 49(2): 407-422.
- [13] WANG T, HONG Y, WANG Q, et al. Identification of mild cogni-

- tive impairment among Chinese based on multiple spoken tasks [J]. *J Alzheimers Dis*, 2021, 82(1): 185–204.
- [14] YEUNG A, IABONI A, ROCHON E, et al. Correlating natural language processing and automated speech analysis with clinician assessment to quantify speech–language changes in mild cognitive impairment and Alzheimer’s dementia[J]. *Alzheimers Res Ther*, 2021, 13(1): 109.
- [15] 王荃一. 基于语音识别的轻度认知障碍检测方法研究[D]. 兰州: 兰州交通大学, 2021.
- [16] 严泉雷. 基于语音识别技术的轻度认知障碍检测[D]. 深圳: 中国科学院大学(中国科学院深圳先进技术研究院), 2020.
- [17] JIANG J, ZHANG J, LI C, et al. Development of a machine learning model to discriminate mild cognitive impairment subjects from normal controls in community screening[J]. *Brain Sci*, 2022, 12(9): 1149.
- [18] GOSZTOLYA G, VINCZE V, TOTH 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.
- [19] POMPILI A, ABAD A, MATOS D M, et al. Pragmatic aspects of discourse production for the automatic identification of Alzheimer’s disease[J]. *IEEE J-STSP*, 2020, 14(2): 261–271.
- [20] SHIMODA A, LI Y, HAYASHI H, et al. Dementia risks identified by vocal features via telephone conversations: A novel machine learning prediction model[J]. *PloS One*, 2021, 16(7): e0253988.
- [21] LIU Z, PROCTOR L, CPLLIER P, et al. Machine learning of transcripts and audio recordings of spontaneous speech for diagnosis of Alzheimer’s disease[J]. *Alzheimers Dement*, 2021, 17: e057556.
- [22] ZHANG Q, LU H, SAK H, et al. Transformer transducer: A streamable speech recognition model with transformer encoders and mn-t loss[C]. *ICASSP*, 2020: 7829–7833.
- [23] GULATI A, QIN J, CHIU C C, et al. Conformer: Convolution–augmented transformer for speech recognition[J]. *ArXiv*, 2020: 2005.08100.
- [24] POVEY D, GHOSHAL A, BOULIANNE G, et al. The Kaldi Speech Recognition Toolkit[J]. *Idiap*, 2012.
- [25] BOSERSMA P. Praat, a system for doing phonetics by computer [J]. *Glott Int*, 2001, 5(9): 341–345.
- [26] EYBEN F, WOLLMER M, SCHULLER B. Opensmile: The munich versatile and fast open–source audio feature extractor[C]. *Proc 18th ACM Int Conf Multimedia*, 2010: 1459–1462.
- [27] BERNARD M, POLI M, KARADAYI J, et al. Shennong: A Python toolbox for audio speech features extraction[J]. *Behavi Res Methods*, 2023: 1–13.
- [28] LI J, YU J, YE Z, et al. A comparative study of acoustic and linguistic features classification for Alzheimer’s disease detection [C]. *ICASSP*, 2021: 6423–6427.
- [29] ROARK B, MITCHELL M, HOSOM J P, et al. Spoken language derived measures for detecting mild cognitive impairment[J]. *IEEE T Audio Speech*, 2011, 19(7): 2081–2090.
- [30] YAMADA Y, SHINKAWA K, NEMOTO M, et al. A mobile application using automatic speech analysis for classifying Alzheimer’s disease and mild cognitive impairment[J]. *Comput Speech Lang*, 2023, 81: 101514.
- [31] MAHAJAN P, BATHS V. Acoustic and language based deep learning approaches for Alzheimer’s dementia detection from spontaneous speech[J]. *Front Aging Neurosci*, 2021, 13: 623607.
- [32] KARLEKAR S, NIU T, BANSAL M. Detecting linguistic characteristics of Alzheimer’s dementia by interpreting neural models [J]. *ArXiv*, 2018: 1804.06440.
- [33] DI P F, PARDE N. Enriching neural models with targeted features for dementia detection[J]. *ArXiv*, 2019: 1906.05483.
- [34] PENG Z, YANG J, LEI N, et al. Connected multi–speech task for detecting Alzheimer’s disease using a two–layer model[C]. *2022 ISCSIC*, 2022: 83–88.

(收稿日期:2023–11–08 录用日期:2024–04–27)

(责任编辑:肖雅妮)