

DOI:10.3969/j.issn.1005-202X.2025.05.020

医学人工智能

基于语音分析的机器学习模型识别阿尔茨海默病

张雨曦^{1,2}, 孙炜¹, 朱国栋^{2,3}, 任之尧^{2,3}, 张瑞秋^{1,2}

1. 华南理工大学设计学院, 广东广州 510000; 2. 广州市民政科技协同创新中心, 广东广州 510000; 3. 广州市老年医院老年医学研究所, 广东广州 510000

【摘要】目的:通过语音分析结合机器学习及特征选择技术,筛选阿尔茨海默病(AD)发展进程中的关键声学特征,构建AD分类模型,为AD的早期筛查提供候选工具。**方法:**从NCMMSC2021 AD语音数据集中获取AD、轻度认知障碍(MCI)与正常老年人(HC)的语音样本,对语音样本进行数据预处理,通过OpenSmile工具包提取eGeMAPS特征集中的特征参数,利用随机森林和支持向量机(SVM)机器学习算法构建分类模型。通过Python进行特征显著性检验及特征重要性排序,运用序列前向搜索(SFS)进行特征选择,并对比特特征选择前后分类模型的效能。采用准确率和受试者工作特征曲线下面积(AUC)指标进行模型分类性能评价。**结果:**分类模型中重要声学特征主要来源于谱斜率、共振峰、基频、响度。模型最优分类效果均来自SFS特征选择后的SVM模型,AD/MCI组识别准确率为0.926(AUC=0.974),AD/HC组识别准确率为0.875(AUC=0.956),MCI/HC组识别准确率为0.879(AUC=0.904)。**结论:**SVM模型性能优于随机森林模型,采用SFS进行特征选择可以有效提高模型性能,语音分析可以成为未来AD快速评估筛查的有效补充工具。

【关键词】阿尔茨海默病;语音分析;随机森林;支持向量机;序列前向搜索

【中图分类号】R318;TP181;R749.16

【文献标志码】A

【文章编号】1005-202X(2025)05-0685-08

Voice analysis-based machine learning models to diagnose Alzheimer's disease

ZHANG Yuxi^{1,2}, SUN Wei¹, ZHU Guodong^{2,3}, REN Zhiyao^{2,3}, ZHANG Ruiqiu^{1,2}

1. School of Design, South China University of Technology, Guangzhou 510000, China; 2. Collaborative Innovation Center for Civil Affairs of Guangzhou, Guangzhou 510000, China; 3. Institute of Gerontology, Guangzhou Geriatric Hospital, Guangzhou 510000, China

Abstract: Objective To identify key acoustic features associated with the progression of Alzheimer's disease (AD) through voice analysis combined with machine learning and feature selection techniques, thereby constructing classification models that serve as candidate tools for the early screening of AD. **Methods** Voice samples from AD, mild cognitive impairment (MCI) and healthy (HC) elderly individuals were obtained from the NCMMSC2021 AD voice dataset. The voice samples underwent data preprocessing, followed by feature extraction from the eGeMAPS feature set via the OpenSmile toolkit. Classification models were obtained utilizing random forest and support vector machine (SVM) algorithms. Significance testing and feature importance ranking were conducted using Python, and the further selection of the optimal features was performed through sequential forward selection (SFS). The classification performance before and after feature selection was compared and evaluated using accuracy and the area under the receiver operating characteristic curve (AUC). **Results** The significant acoustic features in the classification models primarily derived from spectral slope, formant, fundamental frequency, and loudness. The optimal classification performance was achieved with the SVM model following SFS feature selection, with recognition accuracies of 0.926 (AUC=0.974) for AD/MCI group, 0.875 (AUC=0.956) for AD/HC group, and 0.879 (AUC=0.904) for MCI/HC group. **Conclusion** SVM model performs better than random forest model, and the use of SFS for feature selection can effectively enhance model performance. Voice analysis has the potential to serve as a valuable supplementary tool for the rapid AD assessment and screening.

Keywords: Alzheimer's disease; voice analysis; random forest; support vector machine; sequential forward selection

【收稿日期】2024-12-05

【基金项目】广东省重点领域研发计划(2023B0303010003);广州市科技计划重点研发项目(202206060005);广州市民政科技基金(GCAAL2022001)

【作者简介】张雨曦,硕士研究生,研究方向:阿尔茨海默病筛查及计算机应用,E-mail: 202321055540@mail.scut.edu.cn

【通信作者】张瑞秋,博士,教授,硕士生导师,研究方向:计算机软件及计算机应用,E-mail: rqzhang@scut.edu.cn

前言

全国老龄化进程加快,截至2022年末,全国60周岁及以上老年人口已达2.8亿,占总人口的19.8%^[1],老年人是阿尔茨海默病(Alzheimer's Disease, AD)的主要发生群体。据2022年《中国阿尔茨海默病报告》显示,AD已成为中国地区城乡居民的第五大死因,其高昂的照护成本给家庭造成严重的经济负担^[2-3]。AD是最常见的一种痴呆症,其本质是一种神经退行性疾病,主要临床表现包括记忆、语言、思维和行为能力在内的认知功能显著下降^[4]。轻度认知障碍(Mild Cognitive Impairment, MCI)是正常衰老到AD之间的过渡阶段,MCI患者相较于正常老年人会出现一定程度的语言能力障碍^[5]。MCI患者的疾病程度会随年龄增长以每年10%~15%的速度发展为AD^[6],随着疾病程度的加深,AD患者会表现出更严重的语言障碍^[7]。

截至目前,AD尚无根治性治疗药物,当前最有效的治疗手段在于早期识别患者,并通过针对性的康复治疗来延缓AD患者大脑衰退,减缓疾病进程^[8]。量表测试、磁共振成像等传统AD早期筛查方法耗时耗力,受硬件设备等条件限制,应用广泛性差。近年来,语音分析技术以其简单快速、低成本、高效无创等特点在疾病识别分类上得到广泛应用,然而基于语音分析的AD早期筛查与风险评估研究还处于起步阶段。本研究基于NCMMSC2021 AD语音数据集,通过语音分析技术,结合随机森林(Random Forest, RF)和支持向量机(Support Vector Machine, SVM)算法,应用序列向前搜索(Sequential Forward Selection, SFS)进行特征选择,构建AD分类模型,验证SFS进行特征选择的有效性,为AD的早期筛查和识别提供潜在的工具。

1 数据来源与方法

1.1 数据来源

数据来源于2021年公开的NCMMSC2021 AD语音数据集,该语音数据集由第十六届全国人机语音通讯学术会议所组织的AD识别竞赛提供,该竞赛由江苏师范大学、清华大学和海天瑞声联合举办。数据集收集123位老年受试者的语音样本,分别为健康对照组(Healthy Control, HC)44人、MCI组53人、AD组26人。

受试者需完成3项预设的语音任务以收集样本,语音任务为图片描述、流畅性测试和自由谈话。其中,图片描述任务来源于波士顿痴呆症检测,任务要求受试者对《饼干盗贼》图片进行口头描述(图1);流畅性测试广泛应用于多种认知、智力评估量表,任务

要求受试者根据工作人员播放的图片内容尽可能说出你联想到的类别名称,如动物名称;自由谈话要求受试者与工作人员进行非特定主题的自由对话。

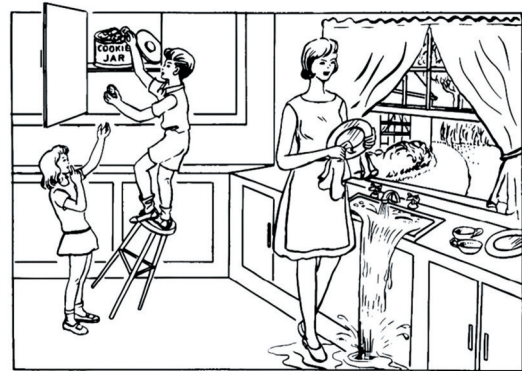


图1 饼干盗贼图片

Figure 1 Picture of the cookie thief

1.2 研究方法

本研究针对NCMMSC2021 AD语音数据集中AD、MCI和HC的3个受试者类别,两两组合为AD/MCI、AD/HC、MCI/HC组,以便后续建立模型并展开研究。首先对原始语音样本进行预处理;随后提取语音特征,将其按照分类组别输入模型进行训练,获得初始模型结果;通过评估特征重要性及特征显著性差异进行初筛,采用SFS进行特征选择,以识别最优特征子集;基于所选特征,重新训练模型,并与初始模型结果进行比较分析,以验证特征选择的有效性并优化模型性能。

1.3 数据预处理和特征提取

采用Python对原始语音样本进行预处理。数据预处理包括对语音数据集中的原始语音样本进行降噪、预加重、端点检测等操作,以提升语音信号的质量,确保后续分析的准确性。在特征提取过程中,通过Python结合OpenSmile工具包从语音样本提取eGeMAPS特征集,共获得88个特征参数。依据受试者标签(AD、MCI、HC)对提取的特征进行分类整理,以便为后续模型的建立提供结构化数据。

1.4 统计学方法

采用Python的scipy库对声学特征进行统计学分析。首先通过Shapiro-Wilk检验各类特征参数的正态分布,其次通过Levene's检验特征参数的方差齐性。对于符合正态分布且满足方差齐性的特征参数,使用独立样本 t 检验;对于不符合正态分布或不满足方差齐性的特征参数,使用Mann-Whitney U 检验。当 $P < 0.05$ 时,认为该特征参数具有统计学意义。

1.5 分类建模

采用RF和SVM两种机器学习方法,分别针对AD/HC、AD/MCI、MCI/HC组建立分类模型。使用10倍交叉验证对每个模型的超参数进行网格搜索,寻找模型的最优参数。SVM的非线性核函数在处理高维特征上表现更优,但非线性SVM并不直接提供特征重要性的度量,因此,本研究中特征重要性将基于由RF模型得出的特征重要性进行评估。

1.6 特征选择

采用SFS进行特征选择,SFS是一种基于贪心算法,逐步向空特征集添加新特征,并评估模型性能的特征选择方法^[9]。相较于其他特征选择方法,SFS能更好地考虑特征之间的相互关系。由于每次添加新特征后都需要重新评估模型性能,因此在特征参数较多时,模型运算负荷会显著增加,为此,本研究首先通过特征显著性检验及RF模型评估的特征重要性对特征进行初步筛选,移除部分冗余特征,随后使用SFS进行特征选择,旨在降低计算负荷,同时确保所选特征对模型性能有显著提升。

2 结果

2.1 RF和SVM初始模型性能

本研究以eGeMAPS语音特征集中的88个特征作为输入特征,受试者的分类情况(AD、MCI、HC)作为标签,通过网格搜索寻找RF和SVM的最优超参数进行分类建模。RF模型及SVM模型各组准确度和AUC值见表1。结果表明,RF模型在3个组别中AD/MCI与AD/HC组的整体分类效果较好,分类准确度和AUC值均在0.750以上,其中AD/HC组表现最佳,准确率为0.792,AUC值为0.896;相较之下,MCI/HC组的分类效果较差,准确率仅为0.697。SVM模型在3个组别中均表现出较好的性能,准确率均超过0.750,AUC值均超过0.800,其中AD/MCI组表现最优,准确率为0.770,AUC值为0.914。相较

之下,MCI/HC组的准确度和AUC值均最低,准确率为0.762,AUC值为0.835。

表1 基于eGeMAPS特征集的RF模型和SVM模型各组准确度和AUC值

Table 1 Accuracy and AUC of each RF and SVM model on the eGeMAPS feature set

组别	准确度	AUC
RF		
AD/MCI	0.778	0.780
AD/HC	0.792	0.896
MCI/HC	0.697	0.798
SVM		
AD/MCI	0.770	0.914
AD/HC	0.775	0.871
MCI/HC	0.762	0.835

综合比较,SVM模型在3个组别的准确度均值(0.769±0.007)及AUC值均值(0.873±0.040)均高于RF模型的准确度均值(0.756±0.051)及AUC值均值(0.825±0.062),SVM模型的整体性能优于RF模型。

2.2 特征显著性结果及重要性排序

通过Python对eGeMAPS特征集中的88个特征在AD/MCI、AD/HC、MCI/HC组中进行显著性检验。分析结果显示,在AD/MCI组中,有32个特征参数具有统计学意义;在AD/HC组中,有29个特征参数具有统计学意义;在MCI/HC组中,有43个特征参数具有统计学意义。图2展示了在3个组别均表现出统计学差异的特征的离散分布情况。图2a及图2b显示特征 slopeUV500-1500_SMA3NZ_AMEAN 及 loudness_sma3_meanFallingSlope 在不同组别之间表现出明显的差异性,且图2a呈现出明显的递减趋势,可能反映AD患者语音在无声段谱斜率的加剧,这表明这两个特征可能与疾病状态的变化存在较强的关联性,是潜在的重要分类指标。

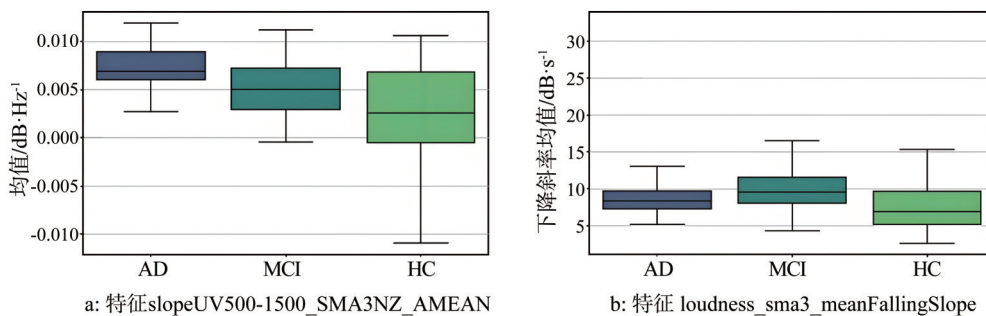


图2 两个特征在AD、MCI和HC组中的箱型图

Figure 2 Box plots of 2 features in AD, MCI and HC groups

通过RF模型中进行特征重要性计算及排序,按照重要性排序的88个特征如图3所示。该分析衡量了各特征对于模型分类能力的贡献,排名最高的特征重要性超过0.060,而排名最低的特征重要性低于0.001;前10个特征的重要性急剧下降,随后逐步趋于平稳,该趋势表明这一小部分特征在分类模型中占据更重要的地位。

由于特征参数较多,本研究仅展示在RF模型中重要性排名前10特征参数的特征来源、重要性及显著性检验结果,见表2。结果显示,在AD/MCI、AD/HC、MCI/HC组中,排名前10的特征重要性均高于0.019。在重要性排名前10的特征中有76.67%的特征表现出统计学差异,具有统计学差异的特征中有52.17%的特征P值小于0.001,即它们在各自组别内存在极其显著的差异。这些特征重要性

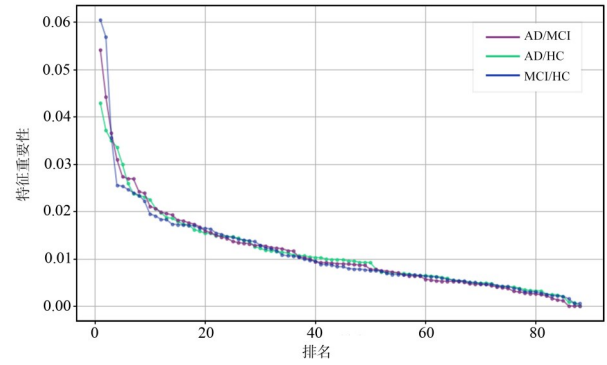


图3 特征重要性

Figure 3 Feature importance

排名靠前的特征主要来源于不同频率范围的谱斜率、共振峰、基频、响度等。

表2 重要性前10特征及其显著性
Table 2 The top 10 features and their significance

组别	特征	来源	排名	重要性	检验方法	P值
AD/MCI	HNRdBACF_sma3nz_stddevNorm	谐波噪声比	1	0.054 1	Mann-Whitney U	0.001 3**
	slopeUV0-500_sma3nz_amean	0~500 Hz 频率范围内的谱斜率(无声段)	2	0.044 2	Mann-Whitney U	0.000 1***
	F3bandwidth_sma3nz_amean	第三共振峰的带宽	3	0.036 5	Mann-Whitney U	0.000 6***
	F0semitoneFrom27.5Hz_sma3nz_mean-RisingSlope	基频相对于27.5 Hz的半音数	4	0.031 0	t-test	0.034 8*
	loudnessPeaksPerSec	响度峰值数(每秒)	5	0.027 3	Mann-Whitney U	0.015 3*
	mfcc4_sma3_amean	梅尔频率倒谱系数	6	0.026 9	t-test	4.69e-06***
	slopeV0-500_sma3nz_stddevNorm	0~500 Hz 频率范围内的谱斜率(有声段)	7	0.026 9	Mann-Whitney U	0.668 8
	slopeV0-500_sma3nz_amean	0~500 Hz 频率范围内的谱斜率(有声段)	8	0.024 2	Mann-Whitney U	0.000 2***
	mfcc3_sma3_stddevNorm	梅尔频率倒谱系数	9	0.023 8	Mann-Whitney U	0.012 3*
	spectralFluxV_sma3nz_stddevNorm	谱流量(有声段)	10	0.021 1	t-test	0.560 1
AD/HC	slopeUV500-1500_sma3nz_amean	500~1 500 Hz 频率范围内的谱斜率(无声段)	1	0.042 9	Mann-Whitney U	9.55e-05***
	shimmerLocaldB_sma3nz_amean	局部的振幅抖动(dB)	2	0.037 1	t-test	0.005 2**
	slopeV500-1500_sma3nz_amean	500~1 500 Hz 频率范围内的谱斜率(有声段)	3	0.035 0	t-test	1.83e-06***
	slopeV0-500_sma3nz_stddevNorm	0~500 Hz 频率范围内的谱斜率(有声段)	4	0.033 6	Mann-Whitney U	0.327 8
	alphaRatioV_sma3nz_amean	α 比率(有声段)	5	0.029 9	t-test	0.000 6***
MCI/HC	F0semitoneFrom27.5Hz_sma3nz_mean-RisingSlope	基频相对于27.5 Hz的半音数	6	0.025 9	t-test	0.168 8
	mfcc4_sma3_amean	梅尔频率倒谱系数	7	0.023 8	t-test	0.001 4**
	loudness_sma3_percentile50.0	响度	8	0.023 3	t-test	0.001 7**
	F1bandwidth_sma3nz_amean	第一共振峰的带宽	9	0.023 1	t-test	0.024 0*
	jitterLocal_sma3nz_amean	局部抖动	10	0.022 5	t-test	0.074 0
	loudnessPeaksPerSec	响度峰值数(每秒)	1	0.060 5	Mann-Whitney U	2.72e-05***
	loudness_sma3_meanFallingSlope	响度	2	0.056 8	Mann-Whitney U	0.000 1***
	F2amplitudeLogRelF0_sma3nz_stddev-Norm	第二共振峰的振幅与基频的对数相对值	3	0.035 6	Mann-Whitney U	5.66e-05***
	alphaRatioV_sma3nz_amean	α 比率(有声段)	4	0.025 6	t-test	2.16e-05***
	spectralFluxV_sma3nz_stddevNorm	谱流量(有声段)	5	0.025 3	Mann-Whitney U	0.009 0**
MCI/HC	slopeV500-1500_sma3nz_amean	500~1 500 Hz 频率范围内的谱斜率(有声段)	6	0.024 6	Mann-Whitney U	0.250 7
	alphaRatioUV_sma3nz_amean	α 比率(无声段)	7	0.024 0	Mann-Whitney U	0.002 2**
	slopeV0-500_sma3nz_stddevNorm	0~500 Hz 频率范围内的谱斜率(有声段)	8	0.023 4	Mann-Whitney U	0.152 4
	F1amplitudeLogRelF0_sma3nz_stddev-Norm	第一共振峰的振幅相对于基频的对数相对值	9	0.022 2	Mann-Whitney U	0.000 4***
	hammarbergIndexUV_sma3nz_amean	Hammarberg 指数(无声段)	10	0.019 4	t-test	0.005 2**

***表示 $P < 0.001$; **表示 $P < 0.01$; *表示 $P < 0.05$

2.3 RF和SVM特征选择后模型性能

为避免进行特征选择时特征过多造成计算负荷过大,本研究首先对特征进行初步筛选,筛选标准为具有统计学差异($P<0.05$)且特征重要性大于0.010的特征;在初步筛选的基础上,采用SFS进行特征选择,筛选出最优特征子集。表3为RF模型及SVM模型中各组别经过SFS特征选择后的特征,图4为RF和SVM各组别经过SFS特征选择后的模型性能的变化趋势,通过逐步增加优势特征,模型性能逐步上升。

结果表明,经过SFS筛选后,AD/MCI组的SVM模型保留10个特征,RF模型保留9个特征,其中6个特征重合;AD/HC组及MCI/HC组的SVM模型和RF模型均保留6个特征,其中AD/HC组2个特征重合,MCI/HC组4个特征重合。高重合率显示这些特征可能直接关联AD神经退行机制,另外,基频相关参数在RF模型及SVM模型的3个组别中均被保留,表明基频在AD分类识别中的重要影响。

表3 经过SFS特征选择后的RF模型和SVM模型各组特征
Table 3 Features of the RF model and SVM model after SFS feature selection

RF组别	特征	来源	SVM组别	特征	来源
AD/MCI	F0semitoneFrom27.5Hz_sma3nz_meanRisingSlope	基频相对于27.5 Hz的半音数	AD/MCI	F0semitoneFrom27.5Hz_sma3nz_meanRisingSlope	基频相对于27.5 Hz的半音数
	F0semitoneFrom27.5Hz_sma3nz_stddevRisingSlope	基频相对于27.5 Hz的半音数		mfcc1_sma3_amean	梅尔频率倒谱系数
	mfcc3_sma3_stddevNorm	梅尔频率倒谱系数		mfcc2_sma3_amean	梅尔频率倒谱系数
	mfcc4_sma3_amean	梅尔频率倒谱系数		mfcc3_sma3_stddevNorm	梅尔频率倒谱系数
	F1amplitudeLogRelF0_sma3nz_amean	第一共振峰的振幅的相对基频		mfcc4_sma3_amean	梅尔频率倒谱系数
	F2bandwidth_sma3nz_amean	第二共振峰的带宽		F3bandwidth_sma3nz_amean	第三共振峰的带宽
	slopeV500-1500_sma3nz_amean	500~1500 Hz频率范围内的谱斜率(有声段)		slopeV500-1500_sma3nz_amean	500~1500 Hz频率范围内的谱斜率(有声段)
	mfcc1V_sma3nz_amean	梅尔频率倒谱系数		mfcc1V_sma3nz_amean	梅尔频率倒谱系数
mfcc4V_sma3nz_amean	梅尔频率倒谱系数	mfcc3V_sma3nz_amean	梅尔频率倒谱系数		
—	—	—	mfcc4V_sma3nz_amean	梅尔频率倒谱系数	
AD/HC	F0semitoneFrom27.5Hz_sma3nz_stddevNorm	基频相对于27.5 Hz的半音数	AD/HC	F0semitoneFrom27.5Hz_sma3nz_stddevNorm	基频相对于27.5 Hz的半音数
	F0semitoneFrom27.5Hz_sma3nz_pctlrange0-2	基频相对于27.5 Hz的半音数		loudness_sma3_amean	响度
	loudness_sma3_percentile50.0	响度		loudness_sma3_stddevNorm	响度
	mfcc1_sma3_amean	梅尔频率倒谱系数		hammarbergIndexV_sma3nz_amean	Hammarberg指数
	F1bandwidth_sma3nz_stddevNorm	第一共振峰的带宽		slopeV500-1500_sma3nz_amean	500~1500 Hz频率范围内的谱斜率(有声段)
	slopeV500-1500_sma3nz_amean	500~1500 Hz频率范围内的谱斜率(有声段)		slopeV500-1500_sma3nz_stddevNorm	500~1500 Hz频率范围内的谱斜率(有声段)
MCI/HC	equivalentSoundLevel_dBp	等效声级	MCI/HC	equivalentSoundLevel_dBp	等效声级
	spectralFluxV_sma3nz_stddevNorm	谱流量(有声段)		mfcc3_sma3_amean	梅尔频率倒谱系数
	logRelF0-H1-H2_sma3nz_amean	基于对数相对基频在H1和H2之间的特征		slopeUV500-1500_sma3nz_amean	500~1500 Hz频率范围内的谱斜率(无声段)
	spectralFlux_sma3_stddevNorm	谱流量		spectralFluxV_sma3nz_stddevNorm	谱流量(有声段)
	logRelF0-H1-A3_sma3nz_amean	基于对数相对基频在H1和A3之间的特征		logRelF0-H1-H2_sma3nz_amean	基于对数相对基频在H1和H2之间的特征
	loudness_sma3_percentile80.0	响度		spectralFlux_sma3_stddevNorm	谱流量

特征选择后的RF和SVM模型结果如表4所示。结果显示,经过特征选择,RF模型在3个组别的准确度及AUC值较初始模型均有显著提升,准确度均超过0.840,AUC值均超过0.900;其中AD/MCI组分

类性能最优(准确度=0.926、AUC值=0.934);3个组别平均准确度为0.883±0.040、平均AUC值为0.922±0.011,整体分类效果良好。SVM模型分类性能经特征选择后性能同样有较大提升,3个组别的准

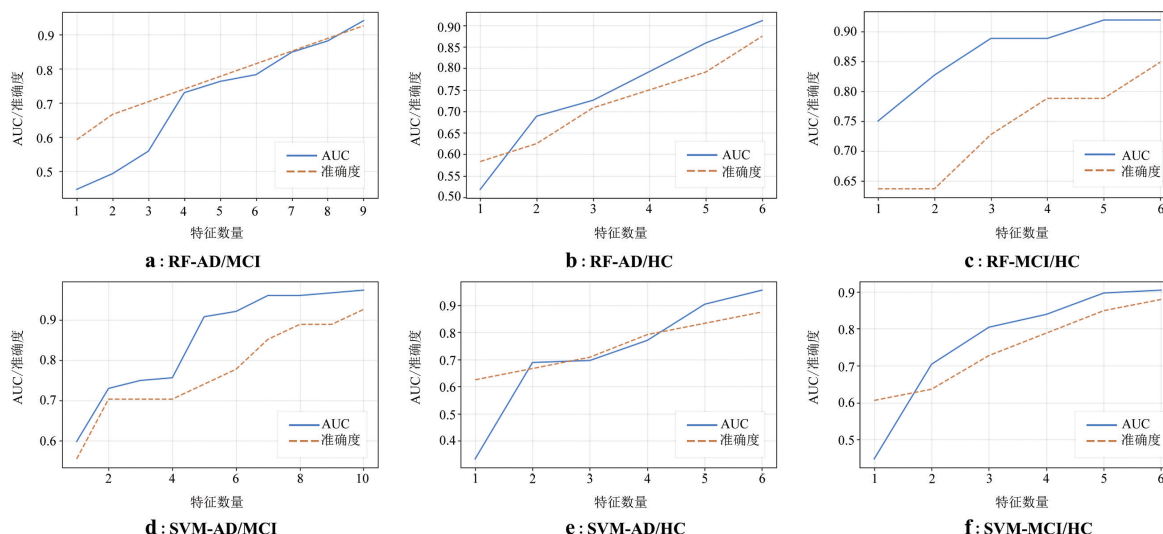


图4 RF,SVM模型经过SFS特征选择后的准确度与AUC

Figure 4 Accuracy and AUC of RF and SVM models after SFS feature selection

准确度及AUC值均超过0.870;其中AD/MCI组分类性能最优,准确度及AUC值均超过0.920,且AUC值高达0.974;3个组别平均准确度为 0.893 ± 0.028 、平均AUC值为 0.945 ± 0.036 ,整体分类性能优于RF模型。值得注意的是,AD/MCI组和AD/HC组的RF、SVM模型准确度相同,但在两个组别中,SVM模型的AUC值在这两个组别中均更高。综合模型分类效果和性能来看,SFS可以有效提高模型性能,且SVM模型的整体性能优于RF模型。

表4 基于重要语音特征的RF模型和SVM模型各组准确度和AUC值

Table 4 Accuracy and AUC of each RF and SVM model based on significant voice features

组别	准确度	AUC
RF		
AD/MCI	0.926	0.934
AD/HC	0.875	0.919
MCI/HC	0.848	0.912
SVM		
AD/MCI	0.926	0.974
AD/HC	0.875	0.956
MCI/HC	0.879	0.904

3 讨论

3.1 语音分析在疾病筛查的应用

语音分析技术在疾病筛查领域的应用日益广泛。Tracy等^[10]验证了应用于语音信号的传统机器学习模型可以有效区分帕金森病患者,其中AUC最高

达0.88;Rosen-Lang等^[11]验证了峰值音量比、停顿标准差等声学参数在高衰弱和低衰弱老年人组之间的统计学差异,进而证明使用语音分析进行衰弱分类的可行性。在AD筛查方面,König等^[12]通过自动语音分析及机器学习方法评估健康老年人、MCI和AD患者的分类可能性,3个组别平均识别精度为82.00%,相较之下,本研究模型准确度更高;García-Gutiérrez等^[13]验证了机器学习在基于自发语音任务的不同程度认知障碍(主观认知下降、AD、MCI)的区分能力,其中3个组别准确度均高于0.90,本研究结果显示3组平均准确率为0.893,AUC值均超过0.90。本研究结果分类性能良好,具有较强的可靠性。

3.2 重要声学特征分析

本研究结果显示在3个组别中重要性排名前10的特征主要来源于谱斜率、共振峰、基频、响度等。谱斜率反映音频信号在低频和高频之间的能量分布,AD患者大脑中神经元的损伤可能导致语音信号在高频部分的能量减少,从而影响谱斜率的变化。基频和共振峰与声道的解剖结构相关,基频决定声音的音高,共振峰影响声音的音色和稳定性;随着AD程度的加深,患者的声带和喉部肌肉受到神经退行性变化的影响,难以调整音高的稳定性,且伴有不自主的震颤,这进一步影响声音的基频和共振峰^[14];此外,AD导致的皮质退化可能会造成甲状腺样萎缩^[15],导致声带变薄和基频变高^[16]。响度是声音强度的主观感受,通常与声压级有关,且会受到包括记忆、语言功能在内的认知功能下降的明显影响,这与Bae等^[17]的研究一致。Szatloczki等^[18]在声学特征训练模型中发现AD患者在语音中表现出更低的响度。

总之,AD患者受神经退行性变化的影响,通常会表现出记忆力衰退、语言障碍、注意力下降、思维缓慢等问题^[19],影响语言连贯性和说话清晰度,进而影响患者的语音质量。通过对这些重要特征的分析,有助于进一步理解AD对患者语言能力的影响,并为未来的早期筛查和干预提供参考。

3.3 SVM模型与RF模型性能比较

本研究结果显示SVM模型性能优于RF模型性能。RF具有高水平的预测性能,并且需要调整的参数相对较少,其不仅能评估特征的重要性,还可以处理混合变量类型(如连续变量和分类变量),并适用于小型和大型数据集。SVM模型相较RF模型能更有效地处理高维特征数据,可以选择不同的核函数来处理不同的决策边界,具有良好的稳定性和鲁棒性,且在小样本数据上表现更突出^[20]。SVM模型被广泛应用于疾病分类识别和预测中,性能表现良好。毕玉珍等^[21]运用逻辑回归、RF、SVM这3种机器学习方法构建孤立性纤维性肿瘤与血管瘤型脑膜瘤分类模型,SVM模型效能最高,测试集AUC为0.99。张娟等^[22]基于多种机器学习模型构建糖尿病足溃疡复发风险预测模型,其中SVM模型为相对最优模型。Trabassi等^[23]基于步态运用多种机器学习方法构建帕金森分类模型,在分类性能上,SVM模型表现最优,F1、准确率、AUC均大于0.85。

3.4 分类模型中特征选择的有效性

本研究显示特征选择可以有效提高模型性能,这与Tuncer等^[24]的研究结果一致。模型中冗杂过多的特征参数不仅会拉低模型性能,还会增加模型的计算量和运算负担。陶帅等^[25]在基于步态的遗忘型MCI和AD识别研究中通过递归特征消除进行特征选择,有效提高RF模型和梯度提升决策树模型分类准确率。Park等^[26]通过多种机器学习方法构建中风识别模型,并利用特征选择技术将模型平均性能提高15.3%。通过特征选择等方法可以去除无意义的特征,有效减少模型的复杂度,提高模型的实用性和性能。

4 结果

近年来,语音分析在疾病筛查上应用广泛,为AD的筛查提供一种新的方法。本研究将语音分析技术与机器学习及特征选择相结合,构建基于声学特征的机器学习分类模型,实现对AD的早期诊断和筛查,最终分类性能最优的分类模型为SVM模型,分类准确率最高为92.6%。因此,基于语音分析构建的机器学习分类模型可以作为AD未来的识别筛查工具,通过SFS进行特征选择可以有效提高模型性能,

验证语音分析技术在AD识别中的可行性,并为机器学习在其中的应用提供支撑和参考。

本研究仍存在一些局限性:(1)本研究仅应用RF和SVM,虽然分类效果良好,但未来应该考虑更多机器学习模型或深度学习模型,如XGBoost、前馈神经网络、卷积神经网络、BERT等,以确定分类性能更优的模型;(2)本研究所提取的特征来源于现有语音特征集,特征可解释性不高,未来应加入更多可解释性高的特征进行研究;(3)因非线性SVM本身并不直接提供特征重要性的度量,所以本研究中的SVM模型特征选择前的特征初筛是根据RF特征重要性进行筛选的,因此所纳入的特征可能并不完全适用于该SVM模型。

【参考文献】

- [1] 民政部,全国老龄办. 2022年度国家老龄事业发展公报[EB]. (2023-12-14) [2024-07-01]. <https://www.gov.cn/lianbo/bumen/202312/P020231214405906944856.pdf>. Ministry of Civil Affairs, National Aging Office. Communique on the development of the national cause for aging in 2022[EB]. (2023-12-14) [2024-07-01]. <https://www.gov.cn/lianbo/bumen/202312/P020231214405906944856.pdf>.
- [2] Ren RJ, Qi JL, Lin SH, et al. The China Alzheimer report 2022[J]. Gen Psychiatr, 2022, 35(1): e100751.
- [3] Jia LF, Du YF, 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.
- [4] 解恒革. 阿尔茨海默病认知障碍表现与量表检测[J]. 中国实用内科杂志, 2010, 30(10): 883-887. Xie HG. Cognitive impairment and neuropsychological assessment in Alzheimer's disease[J]. Chinese Journal of Practical Internal Medicine, 2010, 30(10): 883-887.
- [5] Lai YH. Language processing of seniors with Alzheimer's disease: from the perspective of temporal parameters[J]. World Acad Sci Eng Technol Int J Med Health Biomed Bioeng Pharm Eng, 2017, 11: 414-419.
- [6] Petersen RC. Mild cognitive impairment: transition between aging and Alzheimer's disease[J]. Neurologia (Engl Ed), 2000, 15(3): 93-101.
- [7] Chen LJ, Guo XB, Li ZT, et al. Relationship between long non-coding RNAs and Alzheimer's disease: a systematic review[J]. Pathol Res Pract, 2019, 215(1): 12-20.
- [8] Woo J. Brain health: the importance of recognizing cognitive impairment: an Asian viewpoint[J]. J Am Med Dir Assoc, 2015, 16(9): 722-723.
- [9] 施启军,潘峰,龙福海,等. 特征选择方法研究综述[J]. 微电子学与计算机, 2022, 39(3): 1-8. Shi QJ, Pan F, Long FH, et al. A review of feature selection methods[J]. Microelectronics & Computer, 2022, 39(3): 1-8.
- [10] Tracy JM, Özkanca Y, Atkins DC, et al. Investigating voice as a biomarker: deep phenotyping methods for early detection of Parkinson's disease[J]. J Biomed Inform, 2020, 104: 103362.
- [11] Rosen-Lang Y, Zoubi S, Cialic R, et al. Using voice biomarkers for frailty classification[J]. GeroScience, 2024, 46(1): 1175-1179.
- [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 Dement (Amst), 2015, 1(1): 112-124.
- [13] García-Gutiérrez F, Alegret M, Marquí M, et al. Unveiling the sound of the cognitive status: machine learning-based speech analysis in the Alzheimer's disease spectrum[J]. Alzheimers Res Ther, 2024, 16(1): 26.

- [14] Parlak MM, Saylam G, Babademez MA, et al. Voice analysis results in individuals with Alzheimer's disease: how do age and cognitive status affect voice parameters? [J]. *Brain Behav*, 2023, 13(11): e3271.
- [15] 闫岳铭. 阿尔茨海默症患者大脑微结构变化——基于表面的形态学研究[D]. 太原: 太原理工大学, 2022.
Yan YM. Brain micro-structural alteration of surface-based morphometry in Alzheimer's disease patients[D]. Taiyuan: Taiyuan University of Technology, 2022.
- [16] Çiyiltepe M, Şenkal ÖA. The ageing voice and voice therapy in geriatrics[J]. *Aging Clin Exp Res*, 2017, 29(3): 403-410.
- [17] Bae M, Seo MG, Ko H, et al. The efficacy of memory load on speech-based detection of Alzheimer's disease[J]. *Front Aging Neurosci*, 2023, 15: 1186786.
- [18] Sztatloczki G, Hoffmann I, Vincze V, et al. Speaking in Alzheimer's disease, is that an early sign? Importance of changes in language abilities in Alzheimer's disease[J]. *Front Aging Neurosci*, 2015, 7: 195.
- [19] 张燕, 郭晨晨, 张玉洁, 等. 阿尔茨海默病认知障碍神经心理学特征分析[J]. *阿尔茨海默病及相关病杂志*, 2021, 4(1): 69-72.
Zhang Y, Guo CC, Zhang YJ, et al. Analysis of neuropsychological characteristics of cognitive impairment in Alzheimer's disease[J]. *Chinese Journal of Alzheimer's Disease and Related Disorders*, 2021, 4(1): 69-72.
- [20] 张浩然, 韩正之, 李昌刚. 支持向量机[J]. *计算机科学*, 2002, 29(12): 135-137.
Zhang HR, Han ZZ, Li CG. Support vector machine[J]. *Computer Science*, 2002, 29(12): 135-137.
- [21] 毕玉珍, 白洁, 白培瑞, 等. 基于影像组学的机器学习模型鉴别孤立性纤维性肿瘤与血管瘤型脑膜瘤[J]. *磁共振成像*, 2023, 14(9): 50-55.
- Bi YZ, Bai J, Bai PR, et al. Machine learning models based on radiomics in differentiating solitary fibrous tumor from angiomatous meningioma[J]. *Chinese Journal of Magnetic Resonance Imaging*, 2023, 14(9): 50-55.
- [22] 张娟, 李海芬, 李小曼, 等. 糖尿病足溃疡复发风险预测模型的构建: 基于Logistic回归和支持向量机及BP神经网络模型[J]. *中国全科医学*, 2023, 26(32): 4013-4019.
Zhang J, Li HF, Li XM, et al. Construction of recurrence risk prediction model for diabetic foot ulcer on the basis of Logistic regression, support vector machine and BP neural network model [J]. *Chinese General Practice*, 2023, 26(32): 4013-4019.
- [23] Trabassi D, Serrao M, Varrecchia T, et al. Machine learning approach to support the detection of Parkinson's disease in IMU-based gait analysis[J]. *Sensors (Basel)*, 2022, 22(10): 3700.
- [24] Tuncer T, Dogan S, Acharya UR. Automated detection of Parkinson's disease using minimum average maximum tree and singular value decomposition method with vowels[J]. *Biocybern Biomed Eng*, 2020, 40(1): 211-220.
- [25] 陶帅, 韩星, 孔丽文, 等. 基于步态的机器学习模型识别遗忘型轻度认知障碍和阿尔茨海默病[J]. *中国全科医学*, 2022, 25(31): 3857-3865.
Tao S, Han X, Kong LW, et al. Machine learning-based gait analysis for recognition of amnesic mild cognitive impairment and Alzheimer's disease[J]. *Chinese General Practice*, 2022, 25(31): 3857-3865.
- [26] Park E, Chang HJ, Nam HS. Use of machine learning classifiers and sensor data to detect neurological deficit in stroke patients[J]. *J Med Internet Res*, 2017, 19(4): e120.

(编辑: 谭斯允)