

# 基于神经网络的多特征轻度认知功能障碍检测模型\*

王欣<sup>1</sup>, 陈泽森<sup>2</sup>

1. 中山大学外国语学院, 广东 广州 510275
2. 中山大学航空航天学院, 广东 深圳 518107

**摘要:** 轻度认知功能障碍是介于正常衰老和老年痴呆之间的一种中间状态, 是老年痴呆诊疗的关键阶段。因此, 针对潜在 MCI 老年人群进行早期检测和干预, 有望延缓语言认知障碍及老年痴呆的发生。本文利用患者在语言学表现变化明显的特点, 提出了一种基于神经网络的多特征轻度认知障碍检测模型。在提取自然会话中的语言学特征的基础上, 融合 LDA 模型的 T-W 矩阵与受试者资料等多特征信息, 形成 TextCNN 网络的输入张量, 构建基于语言学特征的神经网络检测模型。该模型在 DementiaBank 数据集上达到了 0.93 的准确率、1.00 的灵敏度、0.8 的特异度和 0.9 的精度, 有效提高了利用自然会话对老年语言认知障碍检测的准确率。

**关键词:** 轻度认知功能障碍; 自然会话; 神经网络模型; 多特征分析; 会话分析

**中图分类号:** H030 **文献标志码:** A **文章编号:** 2097-0137(2023)06-0107-09

## A neural network-based multi-feature detection model for mild cognitive impairment

WANG Xin<sup>1</sup>, CHEN Zesen<sup>2</sup>

1. School of Foreign Languages, Sun Yat-sen University, Guangzhou 510275, China

2. School of Aeronautics and Astronautics, Sun Yat-sen University, Shenzhen 518107, China

**Abstract:** Mild cognitive impairment (MCI) is both an intermediate state between normal aging and Alzheimer's disease and the key stage in the diagnosis of Alzheimer's disease. Therefore, early detection and treatment for potential elderly can delay the occurrence of dementia. In this study, a neural network-based multi-feature detection model for mild cognitive impairment was proposed, which exploits the characteristics of patients with obvious changes in linguistic performance. The model is based on extracting the linguistic features in natural speech and integrating the T-W matrix of the LDA model with the subject data and other multi-feature information as the input tensor of the TextCNN network. It achieved an accuracy of 0.93, a sensitivity of 1.00, a specificity of 0.8, and a precision of 0.9 on the DementiaBank dataset, which effectively improved the accuracy of cognitive impairment detection in the elderly by using natural speech.

**Key words:** mild cognitive impairment; natural speech; neural network model; multi-feature detection; speech analysis

\* 收稿日期: 2023-07-18

录用日期: 2023-07-30

网络首发日期: 2023-09-21

基金项目: 教育部人文社会科学基金 (22YJCZH179); 中国科协科技智库青年人才计划 (20220615ZZ07110400); 中央高校基本科研业务费重点培育项目 (23ptpy32)

作者简介: 王欣(1991年生), 女; 研究方向: 应用语言学; E-mail: wangx736@mail.sysu.edu.cn

轻度认知障碍(MCI, mild cognitive impairment)是一种神经系统慢性退行性疾病,也是阿尔茨海默病(AD, Alzheimer's disease)的早期关键阶段。研究发现, MCI患者每年约有10%~15%的概率转化为AD(Buckner, 2004)。国际AD协会报告显示,全球AD患者已超过5 500万人,到2050年将增加到1.31亿人(Prince et al., 2015)。AD已成为继癌症、心脏病、脑血管疾病之后,引起老年人死亡的第四大病因,是不可忽视的重大健康问题。

提高MCI的早期诊断准确率是延缓、防治AD的关键(Association, 2019)。但由于发病机制尚未明确,早期诊断十分困难,漏诊率高达76.8%(田金洲等, 2019)。检测手段中使用基于脑电图等各类影像学检查结果准确率高(Duan et al., 2020),对专业医生的经验依赖性强,成本高、效率低(黄立鹤, 2022);生物标志物检查对患者具有创伤性及昂贵的缺点,不适用于大规模筛查(李妍等, 2019)。此外,对于未出现明显影像学特征的MCI患者,上述方法则表现出较低的准确率和特异性。

MCI患者在话语表现方面与健康老年人有着较大的差异(刘红艳, 2014;刘建鹏等, 2017),存在词汇量减少、语法结构受损、交流能力下降等特征。这使得基于语言学测量指标的诊断手段成为可能。与传统的神经心理学量表、脑电图、磁共振成像等检测方法相比,话语表现具有显著的外显特征,侵入性低,更经济、扩展性更强,引起了研究人员的广泛关注。研究结果显示, MCI患者不仅在语义流畅性测验任务上的表现受到病理性老化和正常老化的双重影响(刘红艳, 2014),而且使用话语标记的频率和变化呈减少倾向(冉永平等, 2017;杨军, 2004)。但由于指标差异等原因,单纯地利用语言学理论难以对自然会话的多个维度进行量化,研究成果对病理性老化的区分度不高。

在人工智能技术发展的推动下,计算机辅助话语分析技术有了明显的改善(鄂海红等, 2019)。基于卷积神经网络开展的研究已经成为了AD和MCI诊断准确率的热点趋势之一(Chen, 2015)。Shi et al.(2017)将一种多模态堆叠深度极性网络模型应用在多模态的神经影像学数据中,并对AD进行特征学习和分类预测。König et al.(2015)使用语音信号处理技术,对MCI和AD患者的语音进行标记,并使用机器学习方法训练检测

的自动分类器,测试检测的准确性。结果表明,对照组和轻度认知障碍之间识别准确率为(79%±5%);对照组和实验组之间识别准确率为(87%±3%);MCI和AD之间识别准确率为(80%±5%)。该研究证明自动语音分析是客观评估老年人认知能力下降与否的补充工具。

利用MCI患者日常会话中的异常特征,多个研究团队使用神经网络区分患者是否罹患MCI。针对DementiaBank数据集中的473个样本的语音和转录文本,Fraser et al.(2016)选取了35个声学特征和语言学特征组成的特征集合训练逻辑回归分类器,识别的最高准确率为81.92%。Chen et al.(2016)提出更高效的层次化Transformer模型,提高了AD等神经性认知障碍检测任务的运行效率;Zhang et al.(2022)从全局和局部的音频和文本中提取4个特征,提出融合粗粒度和细粒度的复合模型,提高了检测的准确性;董瑞等(2020)提出了一种融合多种语言学特征的Bi-LSTM-CNN-CRF神经网络模型进行命名实体识别,温健(2022)针对AD患者的自发语音构建神经网络模型,搭建了认知衰退检测系统。总体来说,基于神经网络的研究方法在AD早期诊断和发展趋势检测方面具有较高的可行性,但由于此类研究多通过人工智能数据仿真技术获得会话数据,忽视了数据的语言学特征,其检测准确率逐渐陷入瓶颈(赵俊海, 2012)。

综上,如何提取患者自然会话中的语言学特征是提升语言认知障碍的检测准确率的关键。因此,本文采用真实受试者的自然会话数据取代仿真数据,融合语言学多特征分析与神经网络算法,构建基于神经网络的多特征MCI模型,提高了语言认知障碍检测的准确率。

## 1 数据概述

### 1.1 数据来源

本文使用的数据集来源于TalkBank中的DementiaBank—English Protocol Delaware Corpus数据集(Lanzi et al., 2023)。DementiaBank是一个开源的、面向老年人痴呆症状的语料库,也是世界最知名的AD患者多模态语料库之一。在数据采集过程中,所有受试者要求在44岁以上,至少受过7年的教育,没有精神失常的历史或正在服用精神药物。研究人员首先把图片呈现给受调查者,并给每个受调查者1 min时间熟悉图片的内容,然

后要求受试者将自己的图片内容尽可能完整地讲述出来。在语料收集的过程上, 患者组和健康组均遵循相同的步骤。每个被采集的音频文件由人工转出文本, 依据 TalkBank 的 CHAT (codes for the human analysis of transcripts) 协议写成标准的 CHAT 格式。文本被按句分段并标注出填充停顿、起止时间、错误的词和无法理解的词。

表1为访谈调查员和受试者06 (MCI患者) 访谈会话转录文本示例。其中, INV表示访谈调查员; PAR表示访谈参与者; eng表示访谈语言为英语; audio表示访谈使用的媒介为音声; ID为访谈

调查员与访谈参与者的详细信息。@G: Cookie表示围绕《偷饼干图》(见图1) 进行的访谈会话。

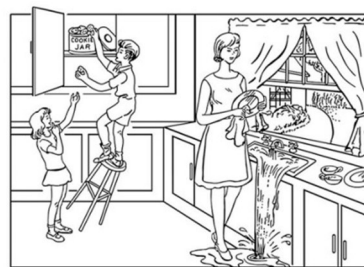


图1 偷饼干图(de la Fuente Garcia et al., 2020)

Fig. 1 Cookie theft diagram

表1 转录文本示例

Table 1 Transcription example

受试者06

@Languages: eng

@Participants: INV Investigator, PAR Participant

@Options: multi

@ID: eng|Delaware|INV||||Investigator||

@ID: eng|Delaware|PAR|82; 00. |female|MCI||Participant||24|

@Media: 06, audio

INV: is I [//] we're gonna look at some different pictures .

INV: and I'm gonna have you describe them to me or tell me some different &-um stories related to them .

INV: okay ?

PAR: say that again .

PAR: &\*INV: so do you want me to +/-?

PAR: you're gonna +/-.

INV: <I'm gonna> [//] &+sh &\*PAR: xxx I'm gonna show you a few pictures and have you describe them or tell me a story .

PAR: okay .

INV: but I'll explain it as I show you each picture .

INV: so ‡ for this first one is I would like you to just +/-.

INV: are you able to see it okay ?

PAR: yes .

@G: Cookie

INV: can you just tell me everything that you see going on in this picture ?

PAR: &-um the kids &-um horsing around &-um trying to get cookies, falling down .

PAR: running water spilling over .

PAR: mother working in the kitchen out in a window <in the> [//] looking outside .

PAR: but she's not looking at the kids .

PAR: &-um mother working, kids playing, getting in trouble .

PAR: okay . [+ exc]

INV: alright .

## 1.2 数据结构

本文使用 DementiaBank 数据集的转录文本进行语言学特征的提取和神经网络的训练, 受试者的采访内容包括看图说话(偷饼干图等)、故事叙述(灰姑娘的故事)、自由谈话(生命中的重要事件、自豪的事件)以及程序性描述话语(制作花生酱及三明治的过程)。为了确保模型的泛化能力, 将训练集设置为每一个文档中随机分配的 80% 文本, 其他的 20% 文本作为测试集, 以期通过神经网络的训练预估受试者患有 MCI 的概率, 数据规模如表 2 所示。其中, MCI 为轻度认知障碍; WNL(within normal limits)为正常范围, 即对照者。如表 3 所示, 数据集中包括每位受试者的人口学信息, 如年龄、性别、第一语言等。以第一语言为例, 因为采访要求受试者运用英语进行回答, 所以第一语言为英语的受试者在表达流畅程度方面会好于第一语言为汉语或者其他语种的受试者。因此, 将人口学信息作为输入张量的一部分对于 MCI 检测模型的训练是必要的。

表 2 训练集和测试集规模  
Table 2 Size of training and test sets min

类型	训练集	测试集
MCI	315.43	78.85
WNL	200.70	50.17
Sum	516.13	129.02

表 3 健康组及患者组人口学信息示例  
Table 3 Demographic information of health and patient groups

测试序号	类型	年龄/岁	性别
01	MCI	84	女
45	MCI	78	男
04	WNL	72	女
25	WNL	70	男

## 2 基于语言学的多特征分析

### 2.1 特征提取

AD 患者在早期轻度认知障碍阶段就已经表现出了一定的语言学特征, 并且话语表现在 AD 的不同阶段表现出不同特征, 其价值不受疾病阶段限制(de la Fuente Garcia et al., 2020)。因此, 话语表现作为 MCI 早期评估和诊断的重要证据(Croot et al., 2000), 成为了老年语言学和神经网络算法等

交叉领域的热点。但一方面, 由于单纯地利用语言学理论难以对语言学测量指标进行量化, 另一方面神经网络算法相关研究多通过数据仿真扩大数据集规模, 虽然使得实验效率有所提高, 但同时简化了真实会话数据特征, 模拟结果与自然会话数据之间存在一定的差距, 致使其检测准确率逐渐陷入瓶颈。语言学测量指标是精确区分病理老化和正常老化的重要筛选工具, 也是构建语言障碍诊断模型的基础。因此, 本文将真实的自然会话语料作为数据集, 通过将语言学定性分析和神经网络算法定量分析结合起来相互验证, 而全面提取这些特征将进一步提高神经网络模型在 MCI 检测上的准确率。

语言学测量指标主要分为语音、语义、语用三个层面。由于 MCI 患者在患病初期语音和句法相对完好(Tsantali et al., 2013), 其受损通常出现在语义和语用层面(Araujo et al., 2011), 因此本模型分别提取了自然会话中的语义和语用层面的特征。在语义指标中, 由于 MCI 患者最常见的障碍是语义流畅性受损导致的命名不能(anomia), 因此语义流畅度可作为客观研究和量化 MCI 语言障碍的工具(Bucks et al., 2000)。另外, 相比较于健康老年人对照组, MCI 患者的形容词比例和动词比例相对较高(黄立鹤, 2022)。因此, 本文选择形容词比例和动词比例, 作为模型的语义特征。进一步考虑 MCI 患者在填塞语和语义密度上表现出显著缺陷, 本文总共提取了以下 5 个特征: 命名不能、形容词比例、动词比例、语义密度、填塞语。

在语用指标中, 大致分为语篇衔接性、连贯性和简洁性(Bucks et al., 2000)。语篇的衔接性体现受试者在词汇和语法方面的衔接手段, 如有无滥用、误用代词及不定词, 空洞词汇及杜撰次词, 话语缺损出现频次等; 连贯性主要体现受试者话题维持的能力, 如是否偏题、答非所问, 多次重复某一主题, 维持话语连贯上是否具有严重缺陷等; 简洁性则体现在受试者话语的简洁程度, 常见的有话语是否冗长, 是否过分使用某一种特定的言语行为, 是否能够进行正常的话轮转换及话轮转换次数等。通过运用话语分析理论对数据集的语料做了细致的标记与分析后, 本文发现 MCI 患者在话语衔接与连贯方面与健康老年人的话语构建策略有着较大的差别。具体而言, MCI 患者在语言学的语篇衔接性、连贯性和简洁性等方面比健康组使用的话语构建策略较少, 且存在较多

的错误, 这和以往研究成果一致(黄立鹤等, 2022; 李妍等, 2019)。因此, 本文采用以下5个语用指标: 无指称代词数、信息单位、局部连贯性、省略、重复。综合以上分析结果, 本文中采用的语义和语用方面的具体特征如表4所示。

表4 本文采用的语言学特征

Table 4 The linguistic features used in this study

特征类型	特征名称	特征说明
语义指标	命名不能	anomia
	形容词比例	A-rate
	动词比例	N-rate
	语义密度	idea density
	填塞语	filler
语用指标	无指称代词数	pronoun without referent
	信息单位	Information unit
	局部连贯性	local coherence
	省略	ellipsis
	重复	palilalia

## 2.2 基于LDA的T-W矩阵提取

隐含狄利克雷分布(LDA, latent Dirichlet allocation)是自然语言处理领域中文本建模中的重要工具。通过每篇文档的主题的概率分布, 构建每篇文档(document)中主题(topic)和词语(word)的关系。采用LDA对测试文本进行预处理, 提取会话文本中的主题词、表征语义密度、语篇连贯性、简洁性等语言学特征, 构建神经网络的输入张量。

**2.2.1 文本清洁** 在构建神经网络模型之前, 对自然会话数据进行文本清洁(见表5)。一方面, 文本清洁可以有效减少待处理的输入变量和文本随机性, 提高了总体性能, 避免误报; 另一方面, 文本清洁可以在LDA和神经网络的构建、测试和训练时, 删除明显的干扰数据, 降低载入数据所需的处理量。文本清洁操作通过Python中的自然语言处理包(NLTK, natural language toolkit)实现。其中,  $\text{text}_{i,j}$  表示第*i*条文本中的第*j*个词语; blanks表示多余空白; tags表示词性标注; stopwords表示除去常见定语等停用词表, 避免对MCI的检测准确率造成影响; digits表示数字, symbols表示特殊符号, seg表示分词函数, low\_freq\_words表示出现低频词。

在文本清洁后, 进一步制作基于目标数据集的词典, 词典可以提供训练模型所需的语言背景和文本数据。对清洗后的文本进行词频统计, 参照2.1中提取的语言学指标进行有监督的深度学

表5 文本清洁算法

Table 5 Text cleaning algorithm

算法 文本清洁
1 $\text{text} = []$
2 $i = 0$
3 当 $i < \text{len}(\text{doc})$ :
4 $\text{text}_{i,j} = \text{text}_{i,j} - \text{blanks} - \text{tags}$ 去除多余空白、词性标注
5 $\text{text}_{i,j} = \text{text}_{i,j} - \text{stopwords}$ 去除符号等停用词
6 $\text{text}_{i,j} = \text{text}_{i,j} - \text{digits} - \text{symbols}$ 去除数字和特殊符号
7 $\text{text}_{i,j} = \text{seg}(\text{text}_{i,j})$ 进行分词处理
8 $\text{text}_{i,j} = \text{text}_{i,j} - \text{low\_freq\_words}$ 去除低频词
9 end
10 返回 $\text{text}$

习, 进一步提取对病理性老化敏感度高的前*N*个语言学指标共同作为模型的词典。

**2.2.2 求解输入张量** 在LDA主题模型中, 每篇文档都由多个主题组成, 每份文档的主题分布由Dirichlet分布产生, 每份文档的主题分布都有差异, 是长度为主题个数的向量。单词的主题分布 $z$ 由参数为 $\theta$ 的多项式分布产生。因此, 将文本集中的每篇文档表示成一组主题的概率分布, 则LDA的概率分布函数为

$$p(\theta, z, w|\alpha, \beta) = p(\theta|\alpha) \prod_{n=1}^N p(z_n|\theta) p(w_n|z_n, \beta). \quad (1)$$

其中 $\theta$ 表示文档的主题分布,  $z_n$ 表示文档中第*n*个单词的主题,  $w_n$ 表示文档中第*n*个单词,  $\alpha$ 和 $\beta$ 是LDA模型的超参数。 $\alpha$ 用于控制每个文档中主题分布情况,  $\beta$ 则用于控制每个主题中词语的分布情况。模型求解目标为得到每份文档的主题分布 $\theta$ 和每个单词的主题分布 $z$ 。

从公式(1)可以看出,  $\alpha$ 和 $\beta$ 的选取是LDA分布的重要参数。为了确保数据的正则性, 本文将 $\alpha$ 和 $\beta$ 设置为相等。根据DementiaBank的话题设定, 本文引入模型复杂度和主题一致性两个指标, 如表6所示。由此确定最佳主题数量和主题词数分别为10。

给定文本的主题数和单词数之后, 通过LDA模型可以得到主题-词语(T-W, topic-word)的二维矩阵。T-W矩阵的每一行表示1个主题, 每一列表示1个词语, 矩阵中的每个元素表示该主题中包含该词语的概率。下面是一个示例T-W矩阵, 包含3个文档和3个主题, 如表7所示。

表6 LDA模型指标与主题数

$\alpha(\beta)$	主题一致性	模型复杂度
10	0.360	-6.016
4	0.353	-5.675
12	0.350	-6.097
18	0.343	-6.290
8	0.341	-6.392
22	0.341	-6.233

表7 T-W矩阵

Table 7 The T-W matrix

主题	文档1	文档2	文档3
1	0.1	0.4	0.5
2	0.7	0.1	0.2
3	0.3	0.3	0.4

从表7可得在各个主题下受试者转录文本中占比最多的词语,其融合而成的T-W矩阵表征了语言学特征,是神经网络的输入张量,可作为检测受试者是否为MCI患者的原始依据。

### 3 基于语言学特征的神经网络检测模型

#### 3.1 考虑多特征的分类模型结构

本文将人工定性分析提取的语言学特征与LDA定量分析提取的语言学特征相结合,构建基于语言学特征的卷积神经网络模型(TextCNN, text convolutional neural network)。TextCNN是卷积神经网络在文本分类任务中的一种应用。与传统模型相比,TextCNN能够有效地处理不同长度的文本,通过在数据集上提取多个高敏感度的语言学特征,并导入到神经网络模型中以达到更高的检测准确率。同时,由于模型引入了CNN,可设置卷积核的宽度和不同的卷积核大小,灵活性高、训练速度快,可以应用于临床大规模文本分类任务中。此模型既可以避免语言学指标无法客观量化等难题,又能够弥补利用数据仿真构建语料库的单一性,从根本上保证了MCI早期检测准确性。

本文的分类检测模型主要分为两部分:一是提取语言学特征,构建神经网络的输入。它主要包括语义和语用特征的提取,文本清洁和应用LDA模型提取制作T-W矩阵。二是构建TextCNN模型进行训练和测试,预估罹患MCI的概率。本文的分类模型如图2所示。

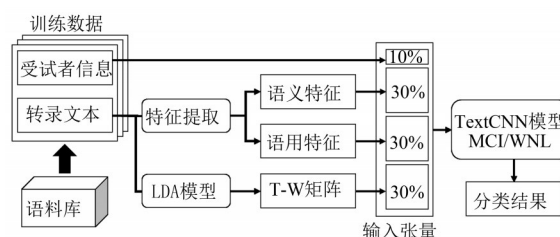


图2 考虑多特征的分类模型

Fig. 2 A classification model that consider multiple features

#### 3.2 模型参数选择

为确保训练模型的稳定性和准确率,研究对神经网络的输入张量进行了归一化、标准化处理。LDA中的主题数和单词数的选择是一个重要的超参数调节问题。合适的主题数和单词数可以提高模型的效果和稳定性。主题数和单词数的选择通常是在较小的范围内进行调整,以避免模型过于复杂和计算量过大。本文的两组主题数和单词数分别设置为5和5,以及10和10。对于主题数和单词数为5的情况,考虑到文本中的采访者,在引导采访时就设计了5个主题“Cookie”、“Cat”、“Rockwell”、“Cinderella”、“Sandwich”,模型能够提取出比较明显的主题特征,同时避免过拟合。通过2.2.2中对于LDA的最佳主题数和单词数实验,研究将主题数和单词数设置为10,可以更细致地提取主题特征并保持模型的准确率。

本文的最优解求解器为SGD优化器,其训练速度快且在小样本的最优解求解中稳定性较高。优化器的误差函数为

$$L = \frac{1}{n} \sum_{i=1}^n (y \ln \hat{y} + (1 - y) \ln (1 - \hat{y})), \quad (2)$$

其中 $y$ 是受试者的真实类型(MCI为1, WNL为0), $\hat{y}$ 是检测的概率值, $L$ 表征真实样本标签和检测概率之间的差值。嵌入层输入为经过处理之后的二维张量;卷积层卷积核大小设置为 $3 \times 3$ ;卷积层激活函数为ReLU;全连接层激活函数为ReLU/Sigmoid;全连接层输出为one-hot编码;在CPU服务器上完成训练。

#### 3.3 检测模型的构建

本文模型的第一步是构建多特征输入张量,包括切分、标注自然会话语料,运用语言学话语分析理论,提取自然会话中的语义密度、形容词比例等语义特征以及填充语、重复等语用特征;以及清洁转录文本,应用LDA模型制作词典获取T-W矩阵及语言学特征,表征其语言学特征;进一步读取受试者性别、第一学历、年龄等信息,共

同融合形成神经网络输入张量。第二步是构建基于多语言学特征的Text-CNN模型, 模型中包含两个隐藏层, 以热点分布输出表示分类结果, 并在

数据集上对模型进行实证训练和检测。检测模型如图3所示。

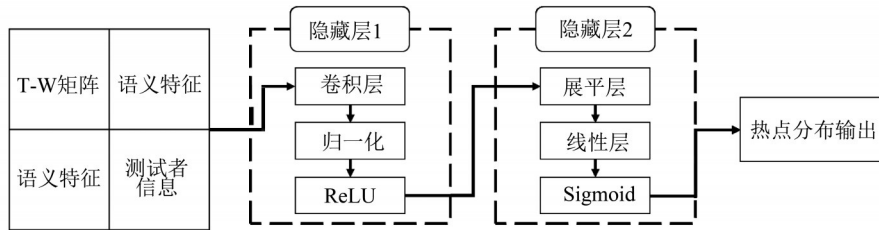


图3 基于多语言学特征的CNN模型

Fig. 3 CNN model based on multilingualism features

图3中输入张量由T-W矩阵、语义特征、语用特征和受试者信息组成。首先对于每个卷积核大小 $k$ , 使用多个宽度为 $k$ 的卷积核对输入进行卷积操作, 得到一系列的特征图, 然后将所有的特征拼接在一起, 作为全连接层的输入, 最终输出分类结果。具体来说, 本文使用了一组 $3 \times 3$ 的卷积核, 通过ReLU激活函数的作用。接下来使用全连接层对这个张量进行处理, 即将其展开为一维张量, 然后使用ReLU、Sigmoid激活函数进行激活, 最后使用one-hot编码输出分类结果。

样本数量。

#### 4.2 结果分析

本文应用LDA和TextCNN模型对Dementia-Bank数据集进行训练。其中MCI患者的测试总时长为394.28 min, WNL的测试时长则为250.87 min, 全部数据时长为645.15 min。对同一受试者的转录文本, 使用随机函数按照4:1的比例选择训练和校验数据, 训练误差如图4所示。

### 4 结果评价与分析

#### 4.1 评价指标

为了保证检测的准确性, 研究结合语言学的语篇衔接性、连贯性和简洁性, 设置了灵敏度、特异度、精度、准确性相关指标展示神经网络模型的测试效果。4项指标及计算公式如表8所示。

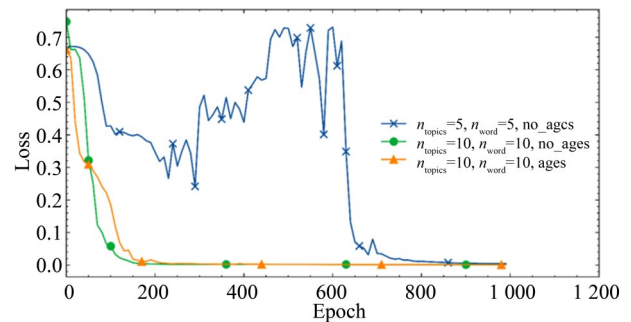


图4 训练误差曲线

Fig. 4 Training error curve

表8 4项指标及计算公式

Table 8 Four kinds of indicators and calculation formulas

指标	计算公式
灵敏度	$TP/(TP + FN)$
特异度	$TN/(FP + TN)$
精度	$TP/(TP + FP)$
准确性	$(TP + TN)/(TP + FP + FN + TN)$

表8中, TP表示真阳性, 即实际为正例且被检测为正例的样本数量; FN表示假阴性, 即实际为正例但被检测为负例的样本数量; TN表示真阴性, 即实际为负例且被检测为负例的样本数量; FP表示假阳性, 即实际为负例但被检测为正例的

图4中 $n_{topics} = 5, n_{word} = 5, no\_ages$ 表示LDA中主题数目为5, 单词数目为5, 输入张量不额外添加年龄等受试者信息;  $n_{topics} = 10, n_{word} = 10, no\_ages$ 表示LDA中主题数目为10, 单词数目为10, 输入张量不额外添加年龄等受试者信息;  $n_{topics} = 10, n_{word} = 10, ages$ 表示LDA中主题数目为10, 单词数目为10, 输入张量额外添加年龄等受试者信息。

比较3条曲线的走势, 可以发现: (1)主题数和单词数分别设置为5和5时, 误差曲线收敛效果明显较差。(2)主题数和单词数分别设置为10和10时, 收敛速度明显加快, 训练步数达200时基本完

成收敛。(3)三种模型在训练步数到达1 000时,误差依次递减,分别为0.004 021、0.001 050、0.000 874,并且最终在测试集中的准确率达到1.00。因此,本文在后续调整模型参数时选择第3种模型。

为了进一步提升模型的检测准确率,本文对模型进行了3种不同的参数调优操作,包括增加语言学输入张量信息(模型1)、加密神经网络结构(模型2)和修改激活函数(模型3)。如表9所示,随着参数的不断优化,模型的性能逐步提升,表现为灵敏度、特异度、精度和准确性逐步增加。

计算结果表明,本文所构建的神经网络模型能够对MCI和WNL进行有效识别,在测试集上的

表9 参数调后4项指标的变化  
Table 9 Variation of the four kinds of indicators after parameter adjustment

模型	灵敏度	特异性	精度	准确性
基础	0.89	0.20	0.67	0.64
1	0.89	0.40	0.73	0.71
2	1.00	0.40	0.75	0.79
3	1.00	0.80	0.90	0.93

## 参考文献:

- 董瑞,杨雅婷,蒋同海,2020.融合多种语言学特征的维吾尔语神经网络命名实体识别[J].计算机应用与软件,37(5):183-188.
- 鄂海红,张文静,肖思琪,等,2019.深度学习实体关系抽取研究综述[J].软件学报,30(6):1793-1818.
- 黄立鹤,杨晶晶,2022.阿尔茨海默病老年人篇章语用障碍指标构建及测定问题[J].外语教学,43(2):16-22.
- 黄立鹤,2022.老龄化与老年语言学引论[M].上海:上海外语教育出版社.
- 李妍,周爱红,赵丽娜,等,2019.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,52(3):177-183.
- 刘红艳,2014.基于语料库的老年性痴呆患者找词困难研究[J].解放军外国语学院学报,37(1):42-52.
- 刘建鹏,赵俊海,杜慧芳,2017.基于语料挖掘的阿尔茨海默症患者话语深层正式度研究[J].解放军外国语学院学报,40(3):36-44.
- 冉永平,李欣芳,2017.临床语用学视角下语用障碍的交叉研究[J].外国语(上海外国语大学学报),40(2):28-38.
- 田金洲,王鲁宁,王荫华,等,2019.适用于中国人群的阿尔茨海默病筛查和诊断框架[J].中华内科杂志,28(2):91-101.

准确率为0.93。分类结果的灵敏度和特异度分别达到了1.00和0.80,表明该模型对于MCI患者的识别能力很强,但是在识别WNL时会出现一定误差;在增加受试者的年龄、性别、第一语言等信息后有明显改进。同时,这些指标值也表明该模型具有较好的稳定性和泛化能力。

## 5 结 论

本文利用MCI患者语言学话语表现变化显著的特点,提出了一种基于神经网络的多特征轻度认知障碍检测模型。在提取自然会话中的语言学特征的基础上,融合LDA模型的T-W矩阵与受试者资料等多特征信息,形成TextCNN网络的输入张量,构建了基于语言学特征的神经网络检测模型。同时,利用世界公认的AD多模态语料库DemantiaBank进行了模型训练和测试。结果显示,该模型在MCI早期检测方面具有较高的准确率和灵敏度,并在阐明语言障碍机理具有一定的潜力。目前,研究侧重于会话文本这一单一模态。在未来的工作中,将增加语音、表情、手势等更多模态的语言特征,以期提高面向临床实践的AD早期检测的准确率。

- 温健,2022.基于自发语音的认知衰退预测模型的研究与应用[D].上海:东华大学.
- 杨军,2004.口语非流利产出研究述评[J].外语教学与研究,36(4):278-284+321.
- 赵俊海,2012.阿尔茨海默症患者话语的系统功能语言学研究[D].重庆:西南大学.
- ARAUJO N B, BARCA M L, ENGEDA K, et al, 2011. Verbal fluency in Alzheimer's disease, Parkinson's disease, and major depression[J]. Clinics, 66(4):623-627.
- ASSOCIATION A, 2019. Alzheimer's disease facts and figures [J]. Alzheimer's & Dement, 15(3): 321-387.
- BUCKNER R L, 2004. Memory and executive function in aging and AD: Multiple factors that cause decline and reserve factors that compensate [J]. Neuron, 44 (1) : 195-208.
- BUCKS R S, SINGH S, CUERDEN J M, et al, 2000. Analysis of spontaneous, conversational speech in dementia of Alzheimer type: Evaluation of an objective technique for analysing lexical performance [J]. Aphasiology, 14 (1) : 71-91.
- CHEN W, XING X, XU X, et al, 2016. SpeechFormer:

- A hierarchical efficient framework incorporating the characteristics of speech [J]. <https://arxiv.org/abs/2203.03812>.
- CHEN Y, 2015. Convolutional neural network for sentence classification[D]. Waterloo:University of Waterloo.
- CROOT K, HODGES J R, XUERE B J, et al, 2000. Phonological and articulatory impairment in alzheimer's disease: A case series[J]. *Brain Lang*, 75(2):277-309.
- de La FUENTE GARCIA S, RITCHIE C W, LUZ S, 2020. Artificial intelligence, speech and language processing approaches to monitoring Alzheimer's disease: A systematic review[J]. *J Alzheimers Dis*, 78(4):1547-1574.
- DUAN F, HUANG Z, SUN Z, et al, 2020. Topological network analysis of early Alzheimer's disease based on resting-state EEG[J]. *IEEE Trans Neural Syst Rehabil Eng*, 28(10):2164-2172.
- FRASER K C, MELTZER J A, RUDZICZ F, 2016. Linguistic features identify Alzheimer's disease in narrative speech[J]. *J Alzheimers Dis*, 49(2): 407-422.
- KÖNIG A, SATT A, SORIN A, et al, 2015. Automatic speech analysis for the assessment of patients with pre-dementia and Alzheimer's disease [J]. *Alzheimer's & Dementia: Diagn, Assess Dis Monit*, 1(1): 112-124.
- LANZI A M, SAYLOR A K, FROMM D, et al, 2023. DementiaBank: Theoretical rationale, protocol, and illustrative analyses [J]. *Am J Speech-Lang Pat*, 32(2): 426-438.
- PRINCE M J, WIMO A, GUERCHET M, et al, 2015. World Alzheimer report 2015—the global impact of dementia: An analysis of prevalence, incidence, cost and trends [J/OL]. <https://www.alzint.org/u/WorldAlzheimerReport2015.pdf>.
- SHI J, ZHENG X, LI Y, et al, 2017. Multimodal neuroimaging feature learning with multimodal stacked deep polynomial networks for diagnosis of Alzheimer's disease [J]. *IEEE J Biomed Health Inform*, 22(1): 173-183.
- TSANTALI E, ECONOMIDIS D, TSOLAKI M, 2013. Could language deficits really differentiate mild cognitive impairment (MCI) from mild Alzheimer's disease [J]. *Arch Gerontol Geriatr*, 57(3):263-270.
- ZHANG L, FAN Y, JIANG J, et al, 2022. Adolescent depression detection model based on multimodal data of interview audio and text[J]. *Int Journal Neural Syst*, 32(11): 2250045.

(责任编辑 王海蓉)