

# 情感计算与方言交互在 AI 陪护机器人中的实现与伦理思考

江祥玲, 王旭, 涂程昊

温州商学院 浙江 温州 325035

DOI: 10.61369/ETR.2026070043

**摘要：** 本文针对老龄化社会下老年群体情感陪护需求，提出情感计算与方言交互融合的智能陪护方案。通过构建多模态情感识别、高精度方言交互及记忆强化对话系统，实现机器人从功能工具向情感伴侣转型。技术上采用迁移学习与联邦学习破解方言数据稀缺难题，提升交互精度与隐私保护能力。同时，深入剖析情感欺骗、隐私安全及责任归属等伦理风险，并提出治理路径。研究为智能养老产业提供新方向，助力技术普惠与老龄化社会治理。

**关键词：** 情感计算；方言交互；多模态情感识别；适老化设计；伦理风险

## Implementation and Ethical Considerations of Affective Computing and Dialect Interaction in AI Companion Robots for the Elderly

Jiang Xiangling, Wang Xu, Tu Chenghao

Wenzhou Business College, Wenzhou, Zhejiang 325035

**Abstract:** This paper proposes an intelligent companionship solution that integrates affective computing and dialect interaction to address the emotional care needs of the elderly in an aging society. By constructing multimodal emotion recognition, high-precision dialect interaction, and memory-enhanced dialogue systems, it enables robots to transition from functional tools to emotional companions. Technologically, transfer learning and federated learning are employed to overcome dialect data scarcity challenges, enhancing interaction accuracy and privacy protection capabilities. Meanwhile, ethical risks such as emotional deception, privacy security, and accountability are thoroughly analyzed, with governance pathways proposed. This research offers a new direction for the intelligent elderly care industry, supporting technological inclusivity and aging society governance.

**Keywords:** affective computing; dialect interaction; multimodal emotion recognition; elderly-oriented design; ethical risks

## 引言

### （一）智能养老的核心痛点与情感陪护需求

21世纪人类社会正面临全球老龄化持续加深的问题。国家统计局数据显示，截至2024年末，中国65岁及以上老年人口占比达15.6%，人数超过2.203亿，该数据不断逼近中度老龄化<sup>[1]</sup>临界线。当下的家庭结构存在日趋小型化、空巢化的问题，传统思想下“养儿防老”的家庭照护模式逐渐消失，社会养老服务体系面临前所未有的压力。

2024年年末人口数及其构成

表1 2024年年末人口数及其构成

指标	年末数 (万人)	比重 (%)
全国人口	140828	100.0
其中，城镇	94350	67.0
其中，乡村	46478	33.0
其中，男性	71909	51.1
其中，女性	68919	48.9
其中，0-15岁 (含不满16周岁) <sup>[6]</sup>	23999	17.1
其中，16-59岁 (含不满60周岁)	85798	60.9
其中，60周岁及以上	31031	22.0
其中，65周岁及以上	22023	15.6

课题来源：国家级大学生创新创业训练项目

作者简介：

- 江祥玲 (2005—)，女，汉族，温州商学院，管理学院，会计学专业，本科生。
- 涂程昊 (2004—)，男，汉族，温州商学院，传媒与设计艺术学院，视觉传达设计专业，本科生。

老年群体对健康管理与情感陪护的需求飞速攀升与当前养老服务的供给能力存在较大矛盾。当下的养老机构护工与老人配比普遍高达1:10，人力资源极度匮乏；现有智能养老产品功能单一，不能适配老年人深层次的情感需求；个性化服务供给程度不足，定制化深度情感陪伴、高效应急响应等核心领域存在显著缺口。现有的多数产品仅支持使用标准普通话，与更习惯使用方言交流的基层老年人存在隔阂，这一现代化环境产生的数字鸿沟<sup>[2]</sup>加剧了他们的社会疏离感。

### （二）技术融合的突破潜力：情感计算与方言交互的协同

养老问题逐渐加剧的同时，人工智能、物联网等技术也在迅猛发展。为解决机器人的冰冷机械工具属性，情感计算<sup>[3]</sup>与自然语言处理技术<sup>[4]</sup>不断进行突破，使其成为兼具情感感知与自然交互能力的智能伴侣。在底层实现可识别、理解并响应人类情感，需要搭建人机情感沟通桥梁模块；顶端实现老年人对乡音的情感依赖深度契合，需要构建方言深度交互架构以此打破语言壁垒，也真正实现用技术破解养老难题。

但当前两大技术呈分离态势，情感计算研究采用标准普通话作为底层分析语言，难以适配地方语言文化语境；现有方言处理方式选用语音识别合成，输出模式机械，缺乏深度情感交互。为研发既能听懂乡音，亦能共情心理，为老年人提供更具温度、陪伴与人文关怀的智能陪护方案需将两种技术协同应用。这无疑为智能养老产业开辟了全新的发展方向。

### （三）研究创新点：技术集成与应用模式的双重突破

本研究旨在提出情感计算+方言交互+适老化设计三位一体技术方案，以此来实现技术与应用模式的双重创新突破。首次构建整合了多模态情感识别、高精度方言交互及记忆强化对话的系统，实现完整技术闭环。此方案所构建出的健康管家+方言陪聊+情感锚点三维关怀模式，成功的把机器人从功能工具到情感伴侣的转型。

## 一、核心理论与技术基础

### （一）情感计算理论体系

情感计算的核心宗旨在于赋予计算机辨识、领会、传递及回应人类情感的能力，搭建起人机情感互通的纽带。为适配不同场景的情感解析需求，情感建模领域存在两类经典模型可供参考，根据 Ekman 提出的离散情感模型<sup>[5]</sup>，人类在面临愉悦、悲伤等基础情感类型时，适用于面部表情识别、基础语音情感解析等场景；维度情感模型以效价-唤醒度<sup>[6]</sup>二维模型为核心，效价维度代表情感正负倾向，唤醒度维度则反映情感生理上的激活水平，维度情感模型可精准描摹面临复杂细腻状态下的情感状态，更适配于解析连续动态的情感变化，可以为深层次情感交互提供更优的理论框架。

### （二）方言语音处理关键技术

选用方言 ASR 技术<sup>[7]</sup>的核心难题在于声学模型的适配程度，由于方言与普通话发音存在较大的系统性差异，且不同的方言之间也是大相径庭，若直接套用普通话的模型会大幅降低识别精度和范围。因此本项目选择采用声学模型进行迁移学习的方案，其本质在于先采用大规模普通话语音数据训练得到成熟的基础模型，再逐渐通过加入少量方言语料进行微调，使其快速适配目标方言的声学特性。在此基础上，采用端到端语音识别模型<sup>[8]</sup>，凭借该模型具备的结构简练、无需强制对齐的优势，来实现在低资源方言识别任务的适配性。

方言 TTS 技术则是支撑机器人用方言完成富情感表达的重要工具，也是保障实现机器人高质量情感陪伴的关键。基于深度神经网络的端到端 TTS 模型已成为方言合成的主流路径，在其核心技术上聚焦两点，一方面通过每季度联邦学习更新持续扩充语料库，为破解方言数据稀缺困境提供数据基础；另一方面注重深入

挖掘方言特有的语调、重音等韵律特征，以此切实提升语音自然度与情感感染力。

### （三）人机交互与适老化设计框架

结合老年群体特性，机器人在交互设计上需要严守核心适配原则，即在交互方式上，整合语音、手势、动作等多元形式，重点优化语音识别能力与降低按键操作复杂性，最大限度降低学习和使用门槛；在交互容错性上，优化对于不完整、有歧义或带口音的指令的识别和执行，并同步给予友好引导反馈；坚守隐私保护底线，事先清晰告知用户数据采集、使用及存储规则，以保障用户始终掌控交互主动权。对于项目研发的记忆强化对话系统集中彰显上述设计思路，通过融合用户健康数据与长期交互记录精准构建动态用户画像，进而实现从单次交互到持续情感联结的跃升。

## 二、全域方言交互系统的实现

### （一）基于迁移学习与联邦学习的方言 ASR 模型

系统通过搭载全域方言神经网络<sup>[9]</sup>，来做到20余种方言的实时双向交互，可以实现其识别准确率 $\geq 95\%$ 。技术实现上采用深度迁移学习方法，即首先利用大规模标准普通话语音数据训练模型，使该模型建立起一套丰富的声学知识和语言模式，随后针对每种目标方言，收集数小时纯净标准语音数据，对上述训练模型的底层声学特征提取层和高层语言模型层进行微调，以极低的数据成本实现对方言发音特点精准适配。

系统引入联邦学习框架实现持续优化。使得在已经部署在用户家中的机器人可调用本地使用脱敏后的交互语音数据进行微调模型，此过程仅将模型参数更新量加密后上传至云端服务器进行聚合，生成更加适配用户使用需求的全局模型。通过这种更新机制既保护用户隐私，又实现了方言语料库的持续扩增和模型效果的迭代优化。

## （二）情感化方言 TTS 合成实现

应用端到端语音合成模型开发的情感化方言 TTS 系统可实现机器人的方言回应更具亲切感。在开发过程需要聘请方言发言人录制包含高兴、焦急等不同情感的语音语料，用于构建带情感标签的方言语料库。进入模型训练阶段，引入情感标签和韵律信息作为控制变量，让模型学习情感与语音特征的映射关系。最后合成阶段，需要对话通过管理系统生成文本来响应附带相应情感标签，TTS 模型根据情感标签和方言类型分析自动生成具备对应情感的方言韵律的语音波形。

## 三、多模态情感识别与融合机制

### （一）多源情感信息采集

为整合多元情感线索，现有的 AI 智护机器人采用多模态感知策略。语音模态可以从方言语音信号中提取基频、语速等声学特征用于情感解析；生理信号模态通过采集心率等实时客观生理指标，真实反映用户情绪波动情况；视觉模态、通过捕捉面部表现的微表情、肢体动作等线索来分析用户心情。利用决策层融合方法整合上述信息，并借助深度学习模型进行联合评判，输出对用户情感状态更精准、全面的洞悉结果。

### （二）多模态情感融合决策

为实现多源信息智能整合，系统采用注意力机制融合方法，即通过分析动态环境变化对不同模态特征权重进行重新分配。例如，在环境嘈杂时，降低语音模态权重，加强依托生理信号；用户沉默时，强化生理与视觉模态作用。根据融合网络输出综合性情感状态数据，为后续对话决策提供核心数据支撑，这种方式有效的提升了面对复杂场景下机器人对于情感识别的准确性。

## 四、伦理风险分析与思考

### （一）情感欺骗与真实性困境

机器人与老年人建立的拟社会关系，其本质是基于算法模拟的虚假情感。长期依赖这类陪伴，容易出现疏离现实人际交往的问题，加剧社会隔离程度。对此，仍需明确机器人底层的工具属性，避免过度拟人化的设计，通过透明化机制披露其运作逻辑，始终坚持机器人仅作为辅助陪伴而非情感替代的伦理底线。

## 参考文献

- [1] 赵林海, 甘筱航. 人口老龄化与经济增长——基于 GaR 模型的实证研究 [J]. 中国管理科学, 2025, 33 (12): 57-70.
- [2] 董焕晴, 曹高辉, 王靖坤. 从数字鸿沟到数字融入: 老年人数字生活融入实践逻辑与优化路径研究 [J/OL]. 情报科学.
- [3] 张林. 情感人工智能的崛起: 实现基础与主要挑战 [J]. 上海广播电视研究, 2025, (04): 11-20.
- [4] 蔡丽芸. 计算机应用中人工智能自然语言处理技术的优化与实践 [J]. 中国信息界, 2025, (10): 16-18.
- [5] 程钢, 陈秀明, 于翔. 基于 TextCNN 融合模型的离散情感分析 [J]. 科学技术创新, 2023, (21): 124-127.
- [6] 马珊, 张文曦, 陈伟. 刺激的效价与唤醒度对陈述性记忆再巩固干预的影响: 基于 ERP 证据的一项研究 [A]. 中国心理学会, 第二十六届全国心理学学术会议摘要集 (六) [C]. 华南师范大学心理学院, : 中国心理学会, 2025: 11-12.
- [7] 赵泽彬. 基于迁移学习的小样本方言语音识别技术研究 [D]. 导师: 杨彦红; 姜丹. 北京印刷学院, 2024.
- [8] 徐凡, 杨剑峰, 颜为之, 王明文. 基于自注意力的端到端方言语音识别模型 [J]. 信号处理, 2021, 37 (10): 1860-1871.
- [9] 王慧琴. 基于深度学习的关中方言语音情感识别研究 [D]. 导师: 张留美; 李小和. 西安石油大学, 2025.
- [10] 李伟恒, 陶媛. 隐私政策下知情同意原则的适用困境与完善路径 [J]. 行政与法, 2025, (03): 104-115.

## （二）隐私侵蚀与数据安全挑战

由于机器人在使用过程中需持续采集老年人健康数据、语音对话等敏感信息，面临的数据隐私与安全隐忧也不容忽视。应贯彻隐私原则<sup>[10]</sup>，严格执行数据最小化的采集标准，使用机器人的过程中充分赋予用户知情同意权与精细化控制权，并结合联邦学习等技术手段，建立健全数据治理与安全防护规范。

## （三）责任追溯与社会公平困境

机器人的发展和使用一直面临着双重伦理难题，一方面，AI 系统决策或行为引发不良后果时，责任界定模糊不定，难以问责真正的过错方和保护受害者一方。另一方面，机器人研发与售价偏高，容易沦为少数群体专属，加剧社会不公平的程度，引发社会问题。解决路径上，不仅需保障算法可解释性，更需要完善法规以明确责任边界，同时推动技术普惠，以模块化设计降低成本，公共参与政策干预，通过采购补贴、纳入保险等方式，确保保障不同群体平等享有技术红利。

## 五、结论与展望

### （一）研究总结

本研究搭建基于多模态融合的情感交互技术框架，深度整合情感计算、方言语音处理等技术，构建多环节闭环情感交互模块。核心创新体现在借迁移学习与联邦学习破解方言数据稀缺难题，实现多方言高精度交互；构建三源融合决策模型提升情感识别精度；研发记忆强化对话系统促成角色跃升。同时，深入剖析情感欺骗、隐私安全、责任归属与社会公平伦理风险并给出治理路径。

### （二）未来展望

未来需要拓展研究边界，在现有基础上进一步深耕技术与应用的双向赋能。在技术层面上，应用大型语言模型以此优化方言对话的质感，完善多模态情感识别体系，新增更多方言品类适配，提升复杂场景交互稳定性。在应用层面上，探索机器人+社区+家庭协同养老模式，搭建共享服务平台，让技术惠及更多老年群体。在伦理层面上，推动行业标准化落地，细化责任界定与隐私保护规范。通过技术迭代与伦理护航，让 AI 智护机器人真正成为懂乡音、知人心、守边界的温暖伴侣，为老龄化社会治理提供更坚实的支撑。