

多模态生成式 AI 工具融入高职课堂教学的探索与研究

——以我校人工智能通识课为例

何雪梅

(南充职业技术学院信息工程学院,四川南充 537100)

[摘要]随着多模态生成式人工智能的爆发式发展,以 ChatGPT、文心一言、通义千问等为代表的工具正在重塑知识生产与传播的方式。高职院校作为技术技能人才培养的主阵地,如何将这类工具从“防控对象”转变为“教学资源”,是当前教学改革亟待回应的现实命题。本研究以南充职业技术学院信息工程学院为实践场域,在分析高职学生学情与多模态 AI 教学价值的基础上,设计并实施了“概念奠基—应用体验—综合实践”三阶段融合教学方案,将文本对话、图像生成、视频创作等工具嵌入项目式学习全过程。通过一学期的行动研究,结合问卷调查、访谈与成绩分析,发现多模态生成式 AI 的融入能显著激发学习动机、降低抽象概念的理解门槛,并有效提升学生 AI 工具迁移应用能力;但同时也暴露出工具依赖、过程评价困难及教师角色转型滞后等问题。论文据此提出明确工具使用边界、构建 AI 素养微课程、强化过程性评价等改进策略,以期同类高职院校人工智能通识教育的数字化转型提供参考。

[关键词]多模态生成式 AI;高职教育;人工智能通识课;生成式人工智能;教学改革

[作者简介]何雪梅(1994—),女,四川南充人,南充职业技术学院信息工程学院助教,工学硕士,研究方向:计算机应用技术、图像处理。

[基金项目]本文系南充职业技术学院 2025 年度校级课题“多模态生成式 AI 工具融入高职课堂教学的探索与研究——以人工智能通识课为例”(项目编号:RWB2502)。

[DOI] <https://doi.org/10.62662/kxwxz0308016>

[中图分类号] G434

[本刊网址] www.oacj.net

[投稿邮箱] jkw1966@163.com

一、问题的提出

2022 年底 ChatGPT 横空出世后,生成式人工智能以月为单位持续迭代。到了 2025 年,模型的多模态能力已从“锦上添花”变为“基础配置”:GPT-4o 能像人一样实时对话、识别表情、读懂屏幕信息;文心一言 4.0 和通义千问均已全面支持文本、图像、音频、视频的联合理解与生成;可灵、即梦等视频生成工具则让“一句话出片”从梗变成现实。可以说,AI 工具已经不再是一个奢侈品,而是像搜索引擎一样触手可及的日常基础设施。

产业端的反应同样迅速。据悉,在 2024 年底南充本地几家信息技术企业中,约超过 6 成将 AI 工具嵌入日常办公流程。反观高职课堂,我们面对的却是一群对 AI 既充满好奇又缺乏驾驭能力的学生,以及一套以知识讲授为主、与产业脱节较明显的课程体系。如何将多模态生成式 AI 从概念介绍转化为学生的真实能力,成为摆在我们面前的一道必答题。

更直接的触动来自课堂观察。在以往的人工智能通识课上,当讲到反向传播、梯度下降这些概念时,台下往往睡倒一片;但只要放一段 Sora 生成的音乐,整个教室的气氛立刻就不一样了。这也表明高职教学的关键矛盾——学生来学技术,却要先啃一大堆他们既不擅长也暂时用不上的数学。

基于上述背景,以南充职业技术学院信息工程学院 2025 级人工智能通识课为依托,开展了一学期的多模态生成式 AI 融入教学实践,本文即是这一探索的阶段性总结。

二、概念界定与研究现状

(一)多模态生成式 AI 的内涵

“多模态生成式 AI”是多模态技术与生成式技术的交汇产物。所谓多模态,指的是系统能同时处理文本、图像、语音、视频等多种形态的信息,并在不同形态之间建立语义关联。生成式 AI 则区别于以分类、判别为核心的传统 AI,其核心能力是学习

数据分布并生成新的、合理的内容。将二者叠加,就得到了既能“看懂”世界,又能“创造”世界的智能工具。这种能力不是简单的功能堆砌:当你给 GPT-4o 看一张手绘草图并对它说“帮我把这张图画成海报,再写一段宣传文案”,它需要同时完成图像理解、意图推理、视觉生成和文本创作,且整个过程要保持风格统一、语义连贯。这种跨模态的协同能力,正是多模态生成式 AI 与传统单模态工具的本质区别。

(二)已有研究述要

国内学界对生成式 AI 教育应用的关注从 2023 年初开始迅速升温。祝智庭等人从教育数字化转型的视角,提出人工智能将从“辅助教学”走向“人机协同教育新生态”,并强调大模型将引发教学范式的深层变革。余胜泉进一步指出,生成式 AI 将倒逼教育评价体系从“结果导向”转向“过程导向”,因为当最终作品可能由 AI 代劳时,学习过程的记录与反思便成为评判学习真实发生的核心依据。焦建利等人系统梳理了生成式 AI 在教育中的典型应用场景,包括智能问答、个性化学习路径生成、创意辅助等,同时警示了学术诚信与“思维外包”的深层风险。

聚焦职业教育领域,刘邦奇等人的研究认为, AI 技术对职业教育的赋能不能停留在“用 AI 演示教学”的浅层应用,而应深入到“用 AI 重构工作流程”的层次,这与高职教育“做中学”的特质高度契合。赵国庆等人专门讨论了生成式 AI 融入课程教学的伦理困境与应对策略,主张将“负责任使用”的教育嵌入课程设计全流程。国际上, Mollick 在其 2024 年出版的著作中提出“永远邀请 AI 到桌边”的理念,倡导学生将 AI 视为工作伙伴而非万能答案机,这一视角对本研究有直接启发。

梳理已有研究可以发现:第一,生成式 AI 教育应用的理论探讨已较为丰富,但多集中在普通本科院校,以高职学生为对象的实证研究相对不足;第二,关于“多模态”维度的教学研究刚刚起步,多数实践仍停留在单一文本交互层面;第三,将多模态生成式 AI 融入一门完整的学期课程,并进行系统效果评估的案例尚不多见。本研究正是试图填补这一空白。

三、课程背景与学情分析

(一)课程基本情况

南充职业技术学院的人工智能通识课,是面向全院大一新生的公共基础课,共 32 学时,每周 2 学时。课程定位为“通识素养课”,目标是让学生建立

对人工智能的总体认知,了解主要技术方向和应用场景,培养基本的 AI 工具使用能力与信息伦理意识,不涉及底层算法的数学推导。

(二)学情调查

2025 年秋季学期信息工程学院学生共 691 人,来自计算机应用技术、大数据技术、数字媒体技术、人工智能技术应用等 7 个专业。课程开始前,通过在线问卷进行摸底,回收有效问卷 568 份。结果显示:97.9% 的学生听说过 ChatGPT 或其他国产 AI 对话产品,但实际使用过的比例仅为 14.4%;在认知层面,75.5% 的学生认同“AI 将是未来必备技能”;关于学习期待,“会用 AI 工具解决实际问题”以 81.4% 的选择率高居首位,远超“了解 AI 发展历史”(32.3%)和“掌握算法原理”(18.8%)。此外,约 42% 的学生自评计算机基础较弱,其中来自乡镇和农村的学生占比较高。

数据透露出几个关键信息:学生有强烈动机,但缺乏使用经验;他们渴望“即学即用”的技能,但对“理论学习”心存畏惧;数字素养的群体分化已经显现,教学设计需要兼顾不同起点的学生。

(三)课程痛点与改革方向

结合教学经验与相关数据,将课程的核心痛点归纳为三点:一是抽象概念难以具身化,学生听了“卷积”“注意力机制”却不知道它们长什么样、能做什么;二是学习产出过于单一,传统考核以期末考试为主,难以反映真实能力;三是与产业应用脱节,学生在课堂上接触不到企业实际使用的 AI 工具链。

针对这些痛点,我们确定了本轮教改的核心思路:以多模态生成式 AI 工具为支架,让学生在“用工具完成项目”的过程中,自然接触和理解人工智能的核心概念。

四、教学设计与实施

(一)整体框架与理论依据

本实践采用行动研究范式,在真实教学情境中“边行动、边观察、边调整”。教学设计以建构主义学习理论为底层逻辑,强调学习者在真实任务情境中通过工具中介完成知识建构。课程被重组为三个递进模块:概念奠基(1~4 周)、应用体验(5~12 周)、综合实践(13~16 周)。每个模块都嵌入了不同深度和广度的多模态 AI 工具应用,工具选择遵循“国产优先、免费可用、多模态协同”的原则,主要使用的平台包括文心一言、通义千问、通义万相、可灵、Kimi 等。

(二)模块一:概念奠基

这一阶段的表层目标是让学生掌握 AI 的基本

概念和发展脉络,其深层意图是帮学生建立与AI工具的“第一次友好接触”。

我们大概花两周时间带学生完成:注册账号、完成首次对话、学会追问。第一个正式任务叫“让AI给你当私人教师”。任务提交时,学生必须附上完整的对话截图,并总结自己的理解。该任务设计思路是:与其让学生读教材,不如让他和AI反复对话,在互动中建构理解。同时,对“追问”的硬性要求意在培养学生提问的深度——我们发现,很多学生第一次提问时只会说“讲讲XX”,第二次就知道加上“请举例”“请对比”“请帮我打个比方”,提问的质量提升非常明显。

第3—4周,我们引入了图像生成工具。任务要求是:“用通义万相生成一张能代表‘人工智能’的图像,然后写一段文字解释你为什么认为这张图能代表AI”。学生在尝试过程中自然就会发现:不同的提示词生成的图像风格差异巨大,画面中元素的布局、色彩、风格与提示词的措辞密切相关。教师趁势引出“特征”“向量”“语义空间”等概念——因为学生已经亲身体会过“提示词微调导致画面大变”的现象,再听到“AI将语义映射到高维特征空间”时,就不觉得那么玄乎了。这就是用体验为概念铺路的典型做法。

(三)模块二:应用体验

模块二是学时最长、活动最密集的主体阶段,围绕三个技术方向展开:自然语言处理(NLP)、计算机视觉(CV)与AIGC创作。此阶段的核心教学形式是“迷你项目+即时展示”。

NLP单元的任务是“做一个智能问答机器人原型”。学生使用Kimi或文心一言的对话功能,先自行设置角色背景(如“校园服务助手”),然后模拟至少10组问—答场景,记录AI的表现并标注回答错误或不当之处,最后形成测评报告。此过程中,学生自然会接触到知识库、幻觉等概念。例如,有学生发现AI对他提出的“某门课期末考试重点是什么”回答得言之凿凿却完全不对,由此引发了关于“AI会一本正经地胡说八道”的热烈讨论,教师顺势引入了大模型幻觉的成因与风险——这些本属“高级话题”的内容,因为有了切身体验,学生接受起来毫无障碍。

CV单元的任务是“用AI给校园设施做个智能诊断”。学生用手机拍摄校园内的设施照片,用通义千问的多模态识别功能让AI判断设施是否完好。记录AI的判断结果,并与自己实地查看的结果做对比,统计准确率。这个项目的巧妙之处在于,它让

学生不仅要“用”AI,还要“检验”AI,这正是信息素养的重要组成部分。有学生拍摄了一个生锈的消防栓,AI判断为“可能存在安全隐患”,而实际该消防栓功能完好,由此学生意识到图像识别的局限性——光照、角度、遮挡都会影响判断。这些发现比教师直接讲授“鲁棒性”概念要深刻得多。

AIGC创作单元是本模块的高潮,我们设计了“为南充设计一张城市名片”综合创作任务。任务要求是:以小组为单位,使用文本对话工具研究南充的历史文化、标志性元素,完成创意文案;然后使用通义万相或文心一格生成视觉主图,再用可灵将静态图转为动态短视频,最后撰写一段200字的设计说明。整个过程涉及文本、图像、视频三种模态的串联使用,模拟了真实世界中的内容生产流水线。值得记录的一个细节是:由于部分文生图工具对英文提示词的响应质量优于中文,很多学生主动去查找丝绸、嘉陵江、三国文化等关键词的英文表达,并相互交流“哪个提示词出图效果好”。这种自发性的跨学科学习行为,在设计任务之前是很难在传统课堂上看到的。

(四)模块三:综合实践

最后4周是综合项目阶段。学生自行组队,从教师提供的选题池中自选题目完成。每组除提交最终作品外,还必须附上“过程档案袋”,内容包括:任务分工表、每次与AI交互的完整记录、至少2次方案迭代的说明、成员个人反思。

考核采用“作品质量(40%)+过程档案(30%)+现场答辩(30%)”的结构。答辩环节特别设置为“无AI模式”,学生必须在不借助任何AI工具的情况下,回答教师关于项目细节和技术原理的提问。其旨在区分“会用AI”和“真的理解”,同时向学生传递一个明确信号:AI是工具,最终的思考和判断要靠人自己。

五、教学效果分析

(一)学习动机与态度变化

课程结束后,我们仍使用问卷进行后测,同时增加了开放性问题 and 满意度量表。回收有效问卷502份。数据显示,对“人工智能很有趣”表示认同的比例从学期初的61.2%上升到88.8%;认同“我能学会使用AI工具”的比例从35.1%跃升至91.4%,增幅显著。在课程整体满意度方面,选择“满意”或“非常满意”的学生占比达82.4%,远高于上学期未融入AI工具时的53.6%。

(二)学业成绩分布

虽然采用了与以往不同的考核方式,我们仍对

比了其他学院未参与的一个班级的期末成绩分布(图1)。实验班优秀率(90分及以上)为18.2%,对照班为9.6%;实验班不及格率为4.8%,对照班为

12.3%。在项目答辩中,实验班学生普遍表现出更强的表达能力和技术解释力。

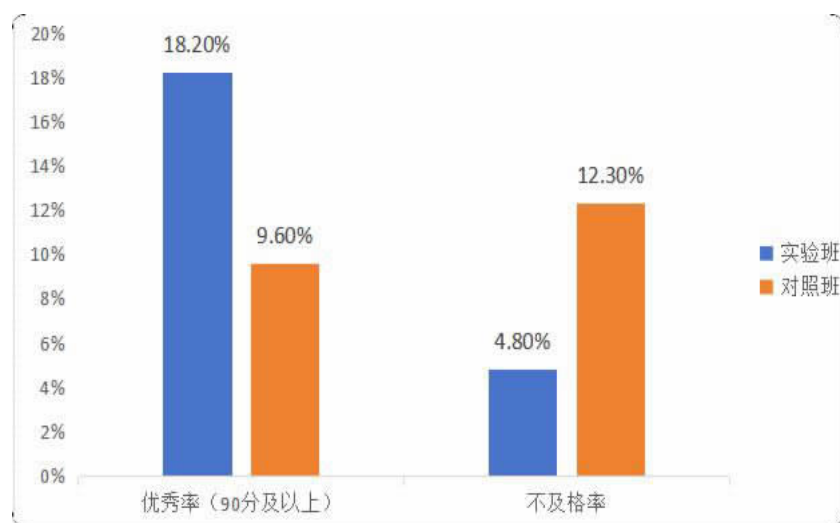


图1 期末成绩对比

六、问题与反思

任何教学改革都不可能一帆风顺,坦率地记录和分析问题,是行动研究的内在要求。本实践中暴露的主要问题有两个方面。

(一)“替代效应”与认知惰化

最突出的问题是部分学生将AI工具直接定位为“代写工具”。尽管在任务中都设计了过程记录和反思环节,但实际操作中,仍有学生先让AI生成完整内容,再花少量时间“反向写出”所谓的思考过程,使过程档案流于形式。据了解,大部分学生承认“大部分任务我先让AI生成,自己只做了微调”,一些学生甚至直言“如果没有AI,很多作业我可能根本不做”。这提醒我们,工具融入与学术诚信的博弈远未结束。

(二)教师角色的转型阵痛

AI工具引入后,教师角色从“知识来源”转向“学习设计师和引导者”,这在理论上是共识,在实践中却相当吃力。首先,备课时间大幅增加:以前准备一堂概念课,主要是梳理PPT;现在则需要提前测试多种AI工具在课堂任务场景下的表现,预判学生可能遇到的坑,准备备用方案。其次,课堂动态性增强:学生的AI生成结果五花八门,教师需要在短时间内做出判断和点评,对学科功底和应变能力要求更高。此外,部分资深教师对生成式AI工具本身存在抵触,认为“这是花里胡哨的东西,不如把底层原理讲透”。教师团队的理念统一和能力建设,

是教改能否持续推进的关键前提。

七、结语与展望

人工智能通识课在高职院校的使命,不是培养算法工程师,而是为每一位即将进入产业的学生装上“AI素养”的操作系统——让他们知道AI能做什么、不能做什么,学会驾驭它而不被它驾驭,始终保持批判性审视的能力。多模态生成式AI工具为这一目标的实现提供了前所未有的可能性:它让抽象的概念变得可视、可听、可对话,让课堂产出从试卷变成海报、视频、小程序,让学生从被动的知识接收者变成主动的内容创作者。

但工具只是工具,决定教育质量的始终是人。在此次实践中,更深切地认识到:技术越强大,教师在“为什么用、什么时候用、用多久”这些问题上的判断力就越重要。如果把AI工具比喻成一台加速器,教师就是那个紧握方向盘的人——必须清楚目的地在哪里,什么时候该踩油门,什么时候该点刹车。

未来,我们计划推进两件事:第一,编制一本“AI课堂使用公约”,明确不同任务环节的AI使用边界,让学生清楚“什么时候鼓励用、什么时候不能碰”;第二,开发8~10讲的AI素养微课,涵盖提示词工程、信息甄别、算法偏见、版权伦理等主题,作为通识课的配套资源。

技术仍在以惊人的速度向前演进。今天的多模态生成式AI或许只是未来更强大智能体教学伙

伴的雏形,但教育的基本命题不会改变——我们始终是在帮助一个个鲜活的学生,成为能够与技术共生、而非被技术裹挟的完整的人。

参考文献:

[1] 卢宇,余京蕾,王宇辰.生成式人工智能的教育应用与展望——以 ChatGPT 系统为例[J].中国远程教育,2023,43(4):24-31.

[2] 祝智庭,戴岭,赵红超. ChatGPT/生成式人工智能对教育的影响:数字化转型的新引擎[J].华东师范大学学报(教育科学版),2023,41(7):1-14.

[3] 余胜泉.生成式人工智能对教育评价的挑战与变革[J].开放教育研究,2023,29(4):4-14.

[4] 焦建利,钟洪蕊.生成式人工智能教育应用:潜力、挑

战与对策[J].现代远程教育研究,2023,35(3):3-11.

[5] 刘邦奇,李鑫.人工智能赋能职业教育的实践框架与推进策略[J].中国职业技术教育,2024(1):10-17.

[6] 王竹立.面向智能时代的知识观与学习观[J].电化教育研究,2023,44(2):5-12.

[7] 陈琳,李芒,郑旭东,等.多模态大模型的教育应用:场景、路径与实证[J].现代教育技术,2024,34(2):25-33.

[8] 赵国庆,黄荣怀.生成式人工智能融入课程教学:伦理风险与应对[J].中国电化教育,2024(3):18-26.

[9] Mollick E. *Co-intelligence: Living and Working with AI* [M]. New York: Portfolio, 2024: 45-102.

[10] 顾小清,郝祥军.从人机协同到人机共生:生成式人工智能教育应用的新进展[J].现代远程教育研究,2024,36(5):3-10.

Exploration and Research on Integrating Multimodal Generative AI Tools into Vocational College Classrooms: A Case Study of Nanchong Vocational and Technical College's General Artificial Intelligence Course

HE Xue-mei

(School of Information Engineering, Nanchong Vocational and Technical College, Nanchong Sichuan 537100, China)

Abstract: With the explosive development of multimodal generative artificial intelligence, tools represented by ChatGPT, ERNIE Bot, and Tongyi Qianwen are reshaping the way knowledge is produced and disseminated. As the main front for cultivating technical and skilled talents, how higher vocational colleges can transform such tools from “objects of prevention” into “teaching resources” is a pressing real-world issue that current teaching reform must address. This study takes the School of Information Engineering at Nanchong Vocational and Technical College as its practice field. Based on an analysis of vocational college students' learning situations and the pedagogical value of multimodal AI, a three-stage integrated teaching plan of “concept foundation—application experience—comprehensive practice” was designed and implemented, embedding tools for text dialogue, image generation, and video creation into the entire process of project-based learning. Through one semester of action research, combining questionnaire surveys, interviews, and grade analysis, it was found that the integration of multimodal generative AI can significantly stimulate learning motivation, lower the threshold for understanding abstract concepts, and effectively enhance students' ability to transfer AI tool skills. However, it also revealed problems such as tool dependence, difficulties in process evaluation, and lagging teacher role transformation. Accordingly, the paper proposes improvement strategies including clarifying the boundaries of tool use, constructing AI literacy micro-courses, and strengthening process-oriented assessment, with the aim of providing a reference for the digital transformation of general artificial intelligence education in similar higher vocational colleges.

Key words: multimodal generative AI; higher vocational education; general artificial intelligence course; generative artificial intelligence; teaching reform