

基于 Agent Skill 的高校课程项目选题研究 ——以《数字图像处理与计算机视觉》为例

李智慧, 刘咏梅

哈尔滨工程大学计算机科学与技术学院, 黑龙江 哈尔滨 150001

DOI:10.61369/EIR.2026060032

摘 要 : 针对《数字图像处理与计算机视觉》课程中学生选题困难、易选现成项目的问题, 本文提出基于 Agent Skill 的自动化选题 Topic-select Skill。该 Skill 通过双策略方式选题, 策略 A 通过检索小众开源论文、分析论文中的缺陷从而发现改进空间; 策略 B 通过搜索 SOTA 方法的应用空白领域、评估跨域迁移可行性以生成创新选题。系统对候选项目进行难度可行性、代码量、可改进空间、可复现性四维度综合评分, 最后输出排序的推荐列表。实验以 5 道选题为验证案例, 验证了方法的有效性与实用性。

关 键 词 : Agent Skill; 课程项目选题; 计算机视觉; 深度学习教学

Research on Course Project Topic Selection Based on Agent Skill— A Case Study of "Digital Image Processing and Computer Vision"

Li Zhihui, Liu Yongmei

School of Computer Science and Technology, Harbin Engineering University, Harbin, Heilongjiang 150001

Abstract : In the "Digital Image Processing and Computer Vision" course, students often face difficulties in selecting topics and tend to choose existing projects. This paper proposes the Topic-select Skill based on Agent Skill to address these issues. This Skill selects topics through a dual-strategy approach. Strategy A discovers improvement spaces by searching for niche open-source papers and analyzing their flaws; Strategy B generates innovative topics by searching for application gaps of SOTA methods and evaluating the feasibility of cross-domain transfer. The system conducts a comprehensive scoring of candidate projects in four dimensions: difficulty feasibility, code volume, improvement space, and reproducibility, and finally outputs a sorted recommendation list. The experiment uses five selection questions as verification cases, verifying the effectiveness and practicality of the method.

Keywords : Agent Skill; course project topic selection; computer vision; deep learning teaching

引言

项目教学法是通过实施一个完整的项目而进行的教学活动, 其目的是在课堂教学中把理论与实践教学有机地结合起来, 充分发掘学生的创造潜能, 提高学生解决实际问题的综合能力。在项目中心式课程教学中, 选题环节是决定项目教学质量的第一个关键步骤。一个好的选题可以激发学生的学习兴趣、引导学生进行有深度的探索; 而选题不当则可能导致学生敷衍了事、学习效果大打折扣。《数字图像处理与计算机视觉》课程的大作业要求学生完成综合性项目, 但选题环节长期面临挑战: 领域发展迅速, 海量开源项目使学生难以甄别; 部分学生选现成项目作为自己的成果提交, 达不到锻炼目的^{[1][2]}。理想的选题应主要应该是难度适中、学生工作量饱满, 最好有改进, 然而仅凭师生手动筛选很难高效满足上述要求。近年来, 大语言模型在阅读文献与代码的能力大幅度提升, 大语言模型与 Agent 技术为解决该问题提供了新思路^{[3][4]}。本文提出一种基于 Agent Skill 的批量选题方法, 设计并实现选题 Skill, 通过 Agent 协作自动完成论文检索、代码评估、缺陷分析和评分排序。

基金项目: 2024 年度哈尔滨工程大学教学改革研究项目, 《项目中心式课程教学中的困难与解决方法研究》(JGYB20240607)。

作者简介: 李智慧 (1966.11-), 女, 汉族, 黑龙江哈尔滨, 副教授, 博士, 研究方向: 计算机视觉, 机器学习, 深度学习。

一、相关工作

现有选题方法可分五类，基于协同过滤的推荐系统利用学生历史数据与兴趣标签进行个性化推荐^[5]，由于每年选课的学生不一样，这种方法不适合单一课程的教学场景。基于知识图谱的方法构建课程知识点与开源项目关联图谱^[6]，语义关联更丰富但构建维护成本高。基于 LLM 的辅助工具在作业设计和代码审查方面展现潜力^{[7][8]}，但是不能直接用于批量选题。此外，PaperWithCode 等论文-代码关联平台提供了星标数、引用量、代码量等统计指标^[9]，但其定位为通用科研检索与 Benchmark 平台，对于课程大作业来说，题目难度太高，且平台仅提供题目信息，缺乏较深层次的难度、局限性等方面的分析功能。

最新的一类是通过 Agent 辅助选题，Hermes Edu Skills^[10]通过 Agent Skill 机制完成教育任务，其中包括项目选题推荐功能，强调模块化与 JSON 配置驱动，但其选题推荐偏向类似大语言模型的对话式推荐。现有方法的共同不足在于：缺乏系统性领域覆盖策略、缺乏代码可运行性验证、缺乏多维度量化评估等步骤，纯 LLM 方案缺乏与真实数据源的校验闭环。本文方法通过多 Agent 协作、API 调用和 Python 程序处理的多重校验，从而弥补了上述不足。

二、方法：Topic-select Skill 设计与构建

（一）整体架构

系统采用“双策略混合+四维度评分”设计，包含五步：参数确认、选题、逐题评价、综合评分排序、结果输出。通过 topicConfig.json 集中管理选题数量、策略比例、星标范围、评分权重等参数。选题前，先查看 topicConfig.json 确认参数是否修改，最终的选题保存在一个 JSON 文件中。

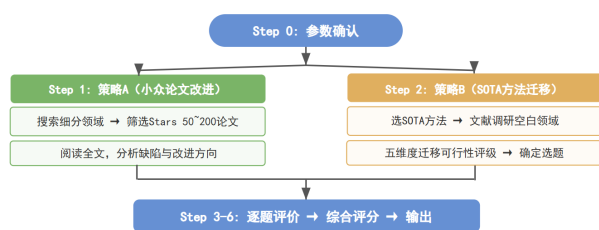


图1 Topic-select Skill 选题流程图

（二）选题策略

本文方法包含两种选题策略，选题时根据两种策略各选一部分题目。

1. 策略 A——小众论文改进选题

策略 A 核心思路为分析现有方法的缺陷并提出改进，培养批判性思维。由 Skill 自动完成以下步骤：（1）从数字图像处理与计算机视觉 N1 个一级领域中选择若干个领域，N 个选题覆盖 $\geq T$ 个一级领域保证多样性；（2）在每个一级领域下搜索 N 个小众细分方向；（3）筛选满足条件的论文——发表时间在距当前时间 10~36 个月之间，有 GitHub 仓库、星标数 50~200；（4）阅读全文全文分析缺陷，给出 1~2 个具体改进方向。

学生得到题目后，先根据未使用过的新数据来测试 Skill 给出

的缺陷和改进方向是否合理，如果不合理，自己也可以根据自己的测试结果分析方法的缺陷，Skill 给出的缺陷和改进方向可帮助学生快速定位潜在改进方向，节约探索时间。

2. 策略 B——SOTA 方法迁移选题

策略 B 以培养学生迁移学习能力、进而锻炼创新能力为目标。

步骤：（1）选择成熟开源 SOTA 方法，条件为星标大于 50、月均引用次数大于 8 次、有预训练权重和调优文档优先；（2）文献调研，下载并阅读引用文献，标注已应用细分领域，再通过搜索 SOTA 方法的细分领域来识别“空白领域”——尚未有该方法应用记录的方向，这些方向可能蕴藏最大创新空间；（3）对候选空白领域进行五维度评级（域特征匹配度、模型适配难度、数据获取难度、预期提升、文献空白度），仅平均分大于 3 者进入后续。

学生得到题目后，可先用细分领域数据集对现有模型进行微调，观察效果。还可以对细分领域重新训练模型，根据新模型测试的效果，决定进一步的改进方法。由于 SOTA 方法是精心设计的高水平方法，取得好的应用效果的概率较高，学生再进一步开动脑筋、查阅文献，实现改进的概率也较高。

（三）逐题评价

下面是题目评价步骤，包括四方面。

（1）方法难度评价与缺陷分析（权重 0.30）：难度采用倒 U 型评分，适中 = 5 分，过易 / 过难 = 1~3 分；

（2）代码量评估（权重 0.25）：实际分析核心代码文件数和行数，少量：小于 3000 行 = 3 分，中量：3000~4000 行 = 5 分；大量：大于 8000 行 = 1 分；

（3）改进空间（权重 0.25）：仔细阅读文章，并分析改进空间，“很有可能” = 5 分，一般 = 3 分，很难 = 1 分；

（4）代码运行验证（权重 0.20）：理想模式下克隆仓库并运行推理；无 GPU 时降级为 Git 克隆、语法检查、依赖审查 + Issues 分析（降级验证结果会在推荐报告中明确标注“未经实际运行验证”警告），代码完全可运行 = 5 分，可运行但需调整 = 3 分，无法运行 = 1 分并置底。

四项加权计算后，得到总分。

（四）Skill 的建立

系统运行于 Windows Server 2022 (PowerShell 5.1, Python 3.11)，OpenCode Agent 框架，底层模型 DeepSeek-V4-Pro。

（1）建立通过策略 A 选题部分，为 AI 提供“图像处理与计算机视觉的选题领域”文档，“选题要求与步骤”文档，告诉 AI 按文档要求选题，如有必要编写 Python 程序放在“选题要求与步骤”文档的同文件夹下，选题策略与步骤保存于“topicSelect.md”，参数保存于 topicConfig.json，选题结果保存于一个 JSON 文档，选题后给出步骤与结果的分析保存于一个“步骤与结果分析.md”文档中；根据“步骤与结果分析.md”及选题的结果，给出指令让 AI 对选题步骤进行修正，并修正相关文档。

（2）建立通过策略 B 选题部分，除“选题要求与步骤”文档与策略 A 不一样，其他都一样，选题过程与策略 A 基本一样。

（3）告诉 AI 把策略 A 与策略 B 的任务合到一起，形成 Skill，AI 会询问 Skill 选题触发词汇，为 AI 提供若干个词如“选

题”“计算机视觉”“数字图像处理”等，可以自动创建 SKILL.md 文档，修改其他文档，完成了测试，并建立好 Skill 的文件夹，在 OpenCode 中注册该 Skill。

（4）应用 Skill 进行选题，并分析结果。

目前该 Skill 已经开源，链接为 <https://github.com/lizhihui9941/topic-select>。

需要指出的是，Topic-select Skill 设计目标不是取代教师选题决策，而是当辅助工具扩大候选项目池并提高选题下限，其有效性高度依赖底层大模型推理能力以及 arXiv、GitHub 等外部数据源可访问性，若 API 调用失败或者论文 PDF 解析不完整时，Skill 可能跳过部分候选题目进而影响覆盖率。此外，策略 A 的‘缺陷分析’实际上是模型对论文进行文本理解，这种操作存在产生幻觉改进方向的风险，所以系统在输出推荐列表时会同时返回原始论文链接以及代码仓库地址，并且要求学生在最终确定题目之前必须进行人工复核。策略 B 的五维度评级里，‘文献空白度’和‘域特征匹配度’这两项依赖于模型对领域知识的掌握程度，如果底层模型训练数据中相关领域语料不足，评级结果就可能出现偏差。上述局限性提示，Agent Skill 在教育场景里应用得坚持“人机协同”原则，自动化选题结果得经过教师或助教二次筛选与适配调整后才可发布给学生，另外不同学期不同班级学生能力分布存在差异，固定权重配置很难适配所有情况，建议教师每次使用前依据实际教学对象调整 topicConfig.json 中的评分权重参数。

三、实验：选题实例展示

（一）实验过程

实验环境：Windows Server 2022，OpenCode Agent 框架，底层模型 DeepSeek-V4-Pro。选题数 N=5，按等比例划分策略 A 与策略 B 的题目数量，最后策略 A 三道题，策略 B 二道题。

提示词设计：“请在计算机视觉方面为我选5道大作业题，题目保存在“F:\opencode\选题”中的“topic_selection.json”中，为我显示所触发的 Skill。”

5个独立运行的 Agent 各携带策略 A 或策略 B 的详细约束，具体约束见 3.2.1 和 3.2.2。

选题过程：选题并非一次性命中。Agent 在 arXiv 返回论文中逐篇筛选：排除无仓库链接论文（约 40%）、星标不符合条件论文（约 20%）、热门方法变体论文（约 15%）。策略 B 还需确认引用量、识别空白领域并做五维度评级，仅评级高于 3 分者进入。最终 5 个 Agent 均找到完全符合条件的选题。

结果概况：5 道选题最高分 4.10，最低分 3.60，AI 给出了推荐排序。覆盖图像分割、目标检测、图像增强、异常检测、视频理解 5 个不同领域。下小节以最高分选题为代表展示详情。

（二）选题结果实例

1. 最佳选题与简介

基于 Depth Anything V2 的路面裂缝 3D 检测（并列最高分，4.10 分）

该题目通过 SOTA 方法：Depth Anything V2（NeurIPS

2024），星标=8,200，引用 1607 次。该方法是一个大规模预训练单目深度估计模型。Depth Anything V2 模型已在室外道路数据预训练，与路面裂缝检测任务的域特征匹配度很高；3D 检测方法能更准确地识别裂缝的严重程度，预期检测效果提升可能性较高；文献空白度很高，极少深度估计方法应用记录，创新空间最大。

2. 拟创新点

AI 阅读文献后给出以下建议创新点：

将大规模预训练深度模型用于编码裂缝 3D 几何特征（深度、宽度、走向）在路面裂缝检测领域较少应用先例，弥补了 2D 方法难以评估严重程度的本质缺陷。

3. Skill 最后评分

代码包括 24 个核心文件，约 2800 行（中量），依赖简单（Pytorch、Opencv 等），无重型框架。评分：难度 5、代码量 4、可改进性 4、可复现性 3，总分 4.10。

4. 人工选题评价

前面都是 AI 生成的结果，下面给出人工结果评价：

路面裂缝检测是智慧交通与道路养护的核心需求，从 2D 升级到 3D 具有明确工程落地前景。文献虽然不是完全空白，但已有文献非常少，如果认真完成，学生可真正做出原创性工作。AI 选题领域非常合适，题目难度稍高，因为三维裂缝数据集较少，需要通过某种方法制作或生成样本，这里有不少需要探索的方面，能诞生有价值的创新。但是 AI 给出的创新建议只能作为参考，学生选题后还需要进一步查阅文献，通过一些的实验来验证创新的可能性，即使创新失败，也比使用现成的项目跑跑代码能获得更多的经验。

四、分析与讨论

策略 A 和 B 各贡献 3 题和 2 题。策略 A 选题星标在 74-195 之间，适合中等规模项目；策略 B 方法星标高达 8,200 和 9,200，但只要找到合适的空白领域，学生无需从头复现整个 SOTA 方法，仅需专注于适配模块设计，工作量有限且聚焦于获得最优性能，可锻炼学生的创新能力。5 题覆盖图像分割、目标检测、图像增强、异常检测、视频理解 5 个不同领域，有效避免选题撞车。

从评价方面看，倒 U 型评分确保“适中”而非“最简单”或“最前沿”获得最高分。排名不唯总分——选题 4 与选题 2 同为 4.10 分，但选题 4 凭借代码量少（2800 vs 10000 行）、方法成熟度高（1607 次 vs 114 次引用）、文献空白度大（5 分）三项优势位列第一，体现了系统在综合衡量创新空间、工作量、基础可靠性方面的设计考量。这一案例表明，AI 选题的价值不在于直接提供“正确答案”，而在于为学生划定一个有价值的探索区间，降低选题的盲目性。

与大语言模型比较方面看，Topic-select Skill 相比直接让大语言模型推荐选题具有以下优势：（1）可配置：所有筛选条件和评分权重可通过 topicConfig.json 调整，适应不同课程和班级的需求。（2）缺陷驱动：每道题附带基于论文和代码分析的缺陷与

改进方向，而非泛泛而谈；（3）可优化：未来在应用中把成功与失败的选题案例一起给 AI，可实现对 Skill 的自动优化。

虽然 Agent Skill 可批量选题，但 AI 的选题还是只能作为参考，学生需要在 AI 选题的基础上进行优化选题，但是 AI 的选题可有效地提升学生选题的底线，提升学生选题的质量。综上所述，通过 Agent Skill 批量选题，可快速建立有时效性的题库，为提升大作业完成质量起到重要作用。

基于上述分析，本文建议教师采用 Topic – select Skill 时按下面操作流程来做，首先要依据当学期学生平均编程能力和课程学时，对 topicConfig.json 里的难度权重和代码量阈值进行调整，比如针对研究生层次可适当高‘改进空间’权重，针对本科生层次就把‘代码运行验证’权重上调，以此保证评分导向和教学目标达成一致。其次，Skill 输出的推荐题目不要直接完全发布，建议教师按照一定比例进行人工筛选与改写，在保留 AI 挖掘的前沿方向的要兼顾本校实验环境、数据资源可获取性和学生先修知识基础，确保最终题库既具备创新

性又具备可操作性。最后，每学期结束之后，要把学生实际完成情况包括常见卡点、创新达成度、题目适配性反馈当作经验积累起来，用于下一轮使用时对配置参数和提示词约束予以调整优化，形成‘AI 生成，教师校准，教学反馈，参数迭代’的闭环。这一流程突出了自动化工具和教师专业判断的互补关系，既发挥 Agent 在海量文献中快速筛选的优势，又保留教师在教学经验、学情判断和伦理把关方面不可替代的作用。

五、结论

本文提出基于 Agent Skill 的选题系统，通过双策略混合架构生成差异化选题，经四维度评分输出推荐。实验表明该方法能有效筛选出难度适中、代码可控、有改进空间且可运行的课程项目。未来工作方向是收集高质量选题与失败选题案例，输入给 AI，实现对 Skill 的迭代优化，确定最优的 Skill 工作流程。

参考文献

[1] 张剑平, 陈仕品. 人工智能赋能教育变革: 机遇、挑战与路径 [J]. 中国电化教育, 2020, (1): 1-8.

[2] 余胜泉. 人工智能教师的未来角色 [J]. 开放教育研究, 2018, 24 (1): 16-28.

[3] Kasneci E, et al. ChatGPT for good? On opportunities and challenges of large language models for education[J]. Learning and Individual Differences, 2023, 103: 102274.

[4] Chu Z, Wang S, Xie J, et al. LLM agents for education: Advances and applications[J]. arXiv preprint arXiv:2503.11733, 2025, 2.

[5] 黄荣怀, 周伟, 杜静, 等. 面向智能教育的三个基本计算问题 [J]. 开放教育研究, 2019, 25 (5): 11-22.

[6] Chen P, et al. KnowEdu: A system to construct knowledge graph for education[J]. IEEE Access, 2018, 6: 31553-31563.

[7] Ghosh A, et al. The future of assessments in the age of generative AI[J]. Journal of Computer Assisted Learning, 2024, 40 (3): 1012-1025.

[8] Leinonen J, et al. Comparing code explanations created by students and large language models[C]//ITICSE 2023: 124-130.

[9] Papers with Code. Machine learning papers with code, evaluation tables and benchmarks[EB/OL]. <https://paperswithcode.com/>, 2024.

[10] Zhongwei V. Hermes Edu Skills: Agent Skills for Chinese education scenarios[EB/OL]. <https://github.com/zhongweiv/hermes-edu-Skills>, 2025.