Advances in Education 教育进展, 2026, 16(11), 917-924
Published Online November 2026 in Hans. https://www.hanspub.org/journal/ae
https://doi.org/10.12677/ae.2026.16112458

Hans汉斯

线上线下混合式教育学习画像构建方法探索

闫　龙[1]，徐　娜[2]，宋海草[1]，史俊伟[1]，黄力波[1]

[1]山东工商学院管理科学与工程学院，山东　烟台
[2]山东医药大学应用心理学教研室，山东　烟台

收稿日期：2026年10月6日；录用日期：2026年11月5日；发布日期：2026年11月12日

摘　要

学习画像基于大数据标签聚类系统生成，聚类不同群体的行为特征，并提供多维度的数据分析，从而置身于多种坐标空间，俯瞰学生群体特点，从而有利于管理者根据学生群体特征进行差异化管理。本文针对线上线下混合教育的过程和特点，对学生画像的指标体系、画像标签和画像构建展开研究，展示了学生画像的特点和作用，对学生分类指导和教育有一定的借鉴作用。

关键词

学习画像，混合式教育，大数据，人脸识别

An Exploration of Methods for Constructing Learning Profiles in Blended Online and Offline Education

Long Yan[1], Na Xu[2], Haicao Song[1], Junwei Shi[1], Libo Huang[1]

[1]School of Management Science and Engineering, Shandong Technology and Business University, Yantai Shandong
[2]Department of Applied Psychology, Shandong Medical and Pharmaceutical University, Yantai Shandong

Received: October 6, 2026; accepted: November 5, 2026; published: November 12, 2026

Abstract

Student profiles are generated using a big data-based tag clustering system that clusters the behavioral characteristics of different groups and provides multidimensional data analysis. This allows for a bird's-eye view of student group characteristics across multiple coordinate spaces, thereby enabling administrators to implement differentiated management based on these characteristics. This paper

文章引用: 闫龙, 徐娜, 宋海草, 史俊伟, 黄力波. 线上线下混合式教育学习画像构建方法探索[J]. 教育进展, 2026, 16(11): 917-924. DOI: 10.12677/ae.2026.16112458

examines the process and characteristics of blended online and offline education, focusing on the indicator system, profile tags, and construction of student profiles. It highlights the characteristics and functions of student profiles, offering valuable insights for providing tailored guidance and education to students.

Keywords

Learning Profiles, Blended Learning, Big Data, Facial Recognition

Copyright © 2026 by author(s) and Hans Publishers Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
http://creativecommons.org/licenses/by/4.0/
Open Access

1. 引言

在万物互联的时代，作为变革世界的新兴科技力量，大数据正以势不可挡之势深度融入社会各个行业和领域。大数据时代的到来，使得教育事业的发展发生了翻天覆地的变化，而对于课堂教学的冲击与影响尤甚，具体表现为课堂教学范式逐渐走向科学化、精准化、智能化与个性化[1]。在“互联网 + 教育”的背景下，随着移动智能设备的普及和数字化学习资源的极大丰富，网络学习、混合式教育逐渐成为主流的学习模式。学生在线下课堂、网络学习空间中的学习行为产生了大量的学习数据。利用基于大数据的用户画像技术对学生的线上线下学习数据进行挖掘分析、构建学生学习画像变得现实可行[2]。

2. 国内外研究现状

学习画像是用户画像技术在教育领域的应用，用户画像最初由 Cooper 提出，又被称为用户角色，是根据用户数据提炼出的描述用户属性及行为的标签集合，常作为勾画目标用户、产品定位、联系用户诉求与设计方向的有效工具，广泛应用于商业领域[3]。学习者画像借助用户画像的应用潜力，通过建立多维度画像标签体系，对教育大数据进行分析处理，反映学生的学习特征和学习行为，推荐学习资源提升学生的学习体验[4]，也可以帮助教师掌握教学实施情况，为制定新的教学、学习策略提供数据支撑[5] [6]。

学习画像的研究集中在指标体系划分、标签分类、聚类方法和数据采集与处理等方面。赵娅等从知识掌握能力、动机水平、行为特征方面划分学生画像模型维度，探究了虚拟仿真实验学习过程中学生的自主学习能力、水平及学习表现[7]。赵建伟以高校多源异构教学数据为基础，构建了以标签体系为核心的学生画像系统[8]。余明华等[9]将学生画像划分为能力属性、行为属性和兴趣属性，以数据分析和人工手段相结合的方式建立了学生画像的标签体系。肖君等根据在线大数据提出了在线学习者画像构建流程。张可佳等提出了一种课堂表现实时动态采集方法 LRD [10]，构建了学习状态判定指标体系。赖英旭等提出一种 SOR 标签体系构建与生成方法，生成学习画像并进行预测[11]。

总体来看，学习画像对数字化教育具有重要价值。现有成果大多聚焦纯在线教学场景，而国内大量教学过程发生于线下课堂，课堂行为数据应当纳入学习画像。另外，现有研究**完整教学实证相对缺乏，对指标局限、标签偏见、数据伦理隐私的系统性讨论不足**。针对以上不足，本文开展混合式教育学习画像构建研究，融合线上平台数据与计算机视觉采集的线下课堂数据，补充实证检验，优化标签体系，并专门论述伦理隐私保护方案。

3. 学习画像构建

本文以线上线下混合式教育过程为对象，以构建学生画像为目标展开研究，总体框架如图 1 所示。

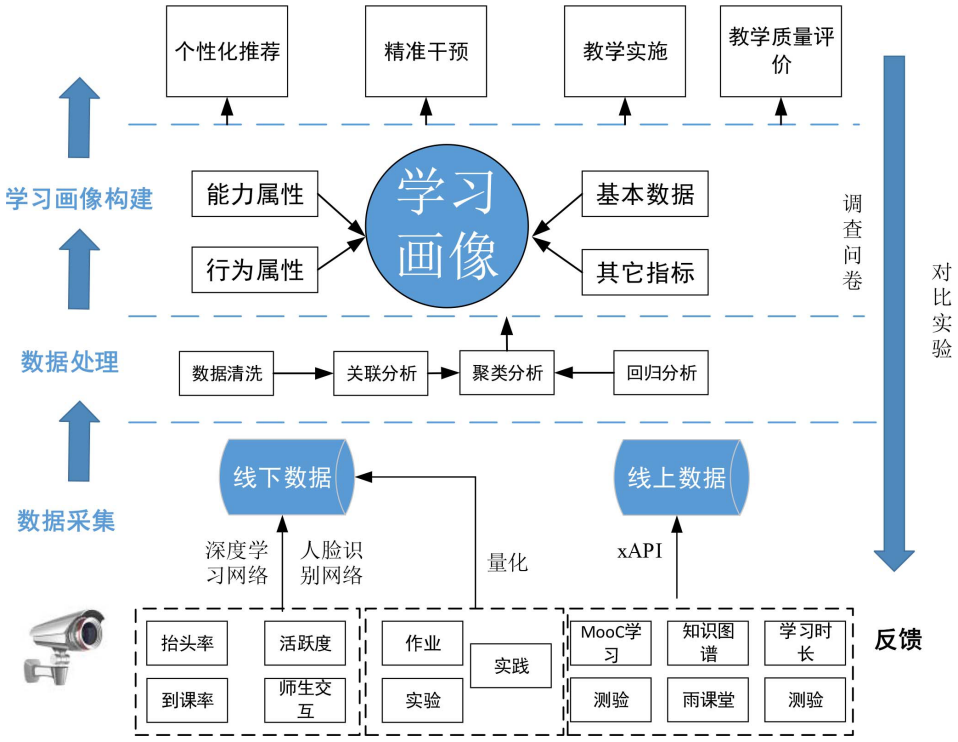


Figure 1. Roadmap for generating learning profiles in blended online and offline education
图 1. 线上线下混合式教育学习画像生成路线

(1) 混合式教育学习画像评价指标体系构建

为了全面、客观地刻画学生能力画像，指标兼顾客观行为记录与主观学习状态，同时**讨论单一指标的局限性**。本文通过线上、线下两类数据源共同支撑画像构建如表 1 所示。线上数据包括视频资源学习时长、在线测验成绩、作业提交情况、知识图谱学习记录；线下数据包括出勤率、课堂交互行为、课堂专注状态等。指标局限性说明：例如“抬头率”仅为课堂专注的间接代理指标，抬头不等于学生在进行深度思考，低头也不代表学生走神(如记笔记、演算)，不能将抬头率直接等同于学习效果，需要结合随堂测验、课堂发言等其他指标综合判断，避免单一指标带来误判。

Table 1. Multi-dimensional data collection
表 1. 多维数据采集

一级维度	二级维度	具体指标	数据来源	指标说明及局限提示
学习行为特征	线上行为	视频观看完成度、视频暂停/回放次数、课件点击数、登录频率、在线时长分布(高峰期/低谷期)、作业提交准时性、预习材料阅读时长	学习平台日志	反映线上参与，无法完全代表真实学习理解程度
	线下行为	出勤率、课堂互动(提问/回答)次数、小组讨论参与度、随堂测验正确率、抬头率(如 AI 课堂分析)	课堂观察、签到系统	专注状态识别为间接指标，需要结合课堂产出综合研判，不可单独作为评价依据

续表

认知投入特征	资源利用	拓展资料阅读量、论坛/讨论区发帖质量(如被点赞数)、回帖深度、搜索关键词记录	学习平台、论坛	发帖活跃度不等同认知水平，沉默学生也可能具备深度思考
	任务完成	作业/实验报告原创性、错题订正行为、项目式学习中的贡献度	作业系统、PBL 平台	需要教师人工复核，避免仅依靠系统量化打分
情感与交互特征	社交交互	师生互动频次、生生互评次数、小组协作中的活跃度、私信交流频率	讨论区、社交群组	交互频次不等于协作质量
	情感状态	论坛发言的情感倾向(积极/消极/困惑)、问卷调查中的学习焦虑度、自我效能感评分	NLP 情感分析、问卷	文本情感分析存在误识别风险，以学生问卷自我报告为重要补充
学业表现特征	过程性考核	单元测验得分、随堂练习得分、平时作业得分	测验系统	阶段性结果，需要跟踪动态变化
	总结性考核	期中/期末考试成绩、作品/项目答辩得分	考试系统	综合学习产出结果
学习风格特征	时间偏好	深夜学习 vs 清晨学习、均匀学习 vs 临时突击(规律性指数)	登录时间戳	仅描述行为模式，不做优劣判定
	媒体偏好	偏好视频学习 vs 文本阅读 vs 听播客	资源点击流数据	仅反映资源选择习惯

(2) 线上线下混合教育过程元数据采集

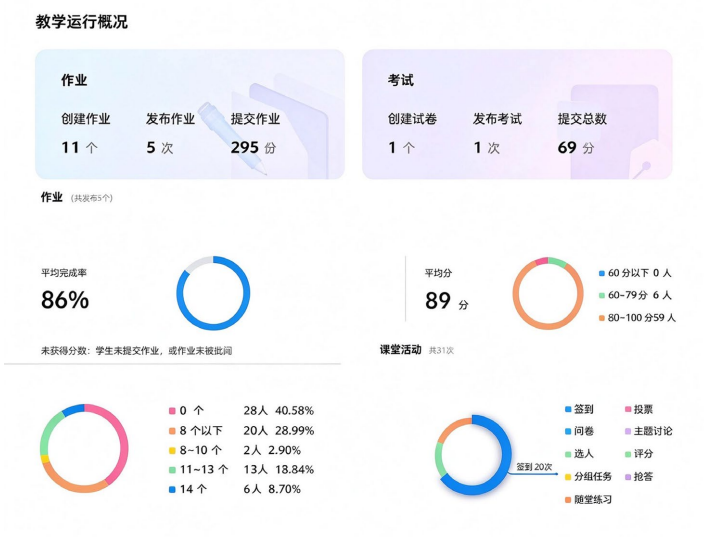
在**获取学生知情同意、满足伦理隐私规范前提下**，通过摄像头实时录制采集课堂数据，借助人脸识别，姿态动作识别技术，采集到课率、抬头率、学习行为状态等指标，如图 2 所示。系统基于 YOLOv7 深度学习的方法采集学生课堂状态(学习、睡觉和玩手机)行为状态，抬头率识别算法采用暨南大学 SCB-Dataset 数据集训练，并在真实课堂场景中进行测试。线上网络课程学习数据采用学习通平台采集视频学习时间、观看时长，测验、作业正确率等。

(3) 多维度学习画像构建

基于上述指标体系，我们需要抽象出具有可读性、概括性的**标签**。摒弃带有主观价值评判、易形成刻板印象的定性词汇，标签以**客观行为描述**为主；建立画像动态更新机制，按教学周滚动更新指标数据，体现学生学习状态的成长与变化，避免静态标签给学生贴上永久固化的“身份烙印”。标签分为：基础属性标签、行为描述标签、能力状态标签、动态预警标签。

a. 基础属性标签

- **自然属性：**年级、专业、性别、生源地(用于群体宏观统计分析，不用于个体评价)。
- **设备属性：**移动端高频学习者、PC 端高频学习者(描述学习访问场景)。



(a)



(b)



(c)

Figure 2. Data collection (from the Chaoxing Platform website) (Left: Online data; Middle: Head-up detection training; Right: Real scene test)
图 2. 数据采集(超星学习通网站) (左图线上数据；中图抬头检测训练；右图真实场景测试)

b. 行为描述标签(仅客观描述行为模式，不做道德、能力优劣定性)

- **学习时间模式：**清晨高频学习、夜间高频学习、碎片化时段学习、集中截止日前集中完成任务。
- **资源使用模式：**高频观看视频资源、高频查阅文档资料、较少参与论坛互动、论坛高频发表观点。
- **交互行为模式：**课堂与线上均高频参与互动、较少公开互动、小组任务中高频输出观点。

c. 能力状态标签(基于多指标综合判定，设置置信度，允许状态迁移)

- **知识掌握综合层级：**综合表现优良、综合表现中等、综合表现有待提升(由测验、作业、课堂产出多维度综合，而非单一分数)。
- **学习投入模式：**深度探究投入、任务导向型完成、参与度偏低。
- **任务节律模式：**多数任务按时完成、常在截止节点完成任务、多次任务逾期。

d. 动态预测标签(风险提示标签，仅作为教师干预线索，非定性结论)

- **需要教学关注：**学业波动风险、情绪表达消极信号；
- **发展潜能线索：**论坛观点思辨性突出、小组任务中承担组织协调工作。
- **画像动态更新机制：**以教学周为最小更新周期，每周同步线上平台日志、课堂采集数据、测验作业结

果；当学生连续 2 到 3 周指标发生显著正向变化时，自动更新标签状态，允许学生画像状态向上迁移，避免静态标签固化学生印象。

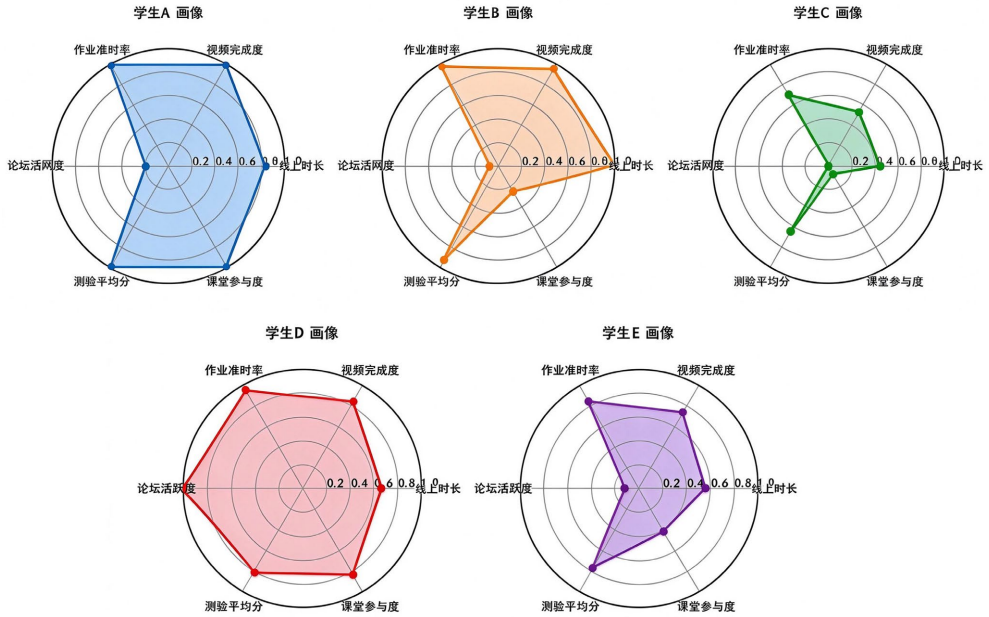


Figure 3. Learning profiles of different learners (radar chart)
图 3. 不同学习者学习画像(雷达图)

雷达图各维度代表可量化的行为与学业指标，雷达图仅作为教师的辅助研判工具，不单独作为评价学生的唯一依据(见图 3)。

- **学生 A (均衡高投入模式)：**雷达图各维度得分整体较高，线上线下学习参与与学业产出均表现较好；
- **学生 B (线上参与充分、线下公开互动较少)：**线上时长、视频完成度指标较高，课堂公开互动、论坛发帖指标偏低；
- **学生 C (多维度参与度偏低)：**多数指标处于较低区间，是教师需要重点关注干预的对象；
- **学生 D (协作交互突出)：**论坛、小组交互维度得分显著偏高，其余指标处于中等区间。
- **学生 E (碎片化参与模式)：**各指标处于中等偏低区间，缺少深度学习行为记录。

画像输出仅作为教学辅助线索，教师结合面对面沟通，实施差异化教学策略：

- **对多维度参与度偏低的学生开展沟通，安排助教针对性辅导；**
- **对线上充分、线下公开互动较少的学生，鼓励小组内部小规模发言，逐步提升课堂参与；**
- **对协作交互突出的学生，可发挥其协作优势，承担小组协作任务；**
- **对均衡高投入的学生，推送高阶拓展学习资源；**
- **对碎片化参与模式的学生，推送时间管理学习建议与轻量化微课程资源。**

4. 实证研究

4.1. 实证研究设计

研究背景与对象：选取高校混合式专业课程《电工电子技术》开展教学实验；参与学生共 124 人，将班级划分为干预组 62 人、对照组 62 人；时间跨度为完整 16 周教学学期。

- **干预组**：教师定期使用本研究构建的线上线下学习画像开展教学研判，针对画像提示的线索开展分层干预(学习提醒、答疑辅导、小组角色调整、资源推送)；
- **对照组**：教师沿用传统教学方式，不使用学习画像系统，采用常规作业、测验开展教学。

数据采集：两组均保留统一的线上学习平台数据；干预组额外启用课堂视觉采集模块(全部学生签署知情同意书后方才采集课堂图像数据)。**评价指标**：期末课程考试成绩、课程综合得分(过程 + 期末)、学习投入问卷(包含学习专注度、学习自我效能感量表)。**实验假设**：使用学习画像开展教学干预的干预组，在学业成绩、主观学习投入上会显著优于对照组。

4.2. 实验结果

学期结束后对两组数据做独立样本 t 检验，主要结果如下：

1. 期末考试成绩：干预组均值 78.62 分，对照组均值 74.37 分，干预组成绩显著高于对照组($p < 0.05$)；

2. 课程综合得分：干预组综合得分均值 80.11，对照组 75.83，存在统计学差异；

3. 学习投入问卷得分：干预组学习投入整体水平高于对照组，其中“获得教师针对性指导”维度提升最为明显。

同时实验也发现局限性：视觉识别模块依然存在少量识别误差；画像标签仅为线索，教师面对面沟通必不可少，完全依赖画像系统会造成误判。实证说明学习画像可以作为教师教学辅助工具，但不能替代教师本身的教学判断。

5. 伦理与隐私保护

将课堂人脸图像、线上学习行为数据用于学习画像，会涉及学生个人隐私、算法标签偏见等风险，本研究从知情同意、数据安全存储、数据访问权限、规避标签固化风险四个层面建立规范。

5.1. 知情同意机制

课堂视觉采集、学习行为数据用于画像分析前，必须向学生充分告知：采集哪些数据、数据用途、数据存储周期；学生拥有选择权，可拒绝课堂摄像头图像采集，对拒绝采集的学生，仅使用线上平台、测验作业等非视觉数据构建画像，不得强制采集人脸信息。所有知情同意以书面形式留存。

5.2. 数据采集、存储与匿名化

1. 课堂视频原始录像**不作为长期存储数据**，系统只保存经过算法提取后的统计指标(出勤率、专注占比等聚合统计结果)，原始视频定期自动删除；

2. 个体画像数据做匿名化处理，学生真实身份与画像数据建立加密映射，禁止对外泄露可识别个人身份的原始数据；

3. 设置数据存储有效期，课程结束之后在规定周期内清理个体层面原始数据，只保留脱敏后的群体统计数据用于教学研究。

5.3. 严格的数据访问与使用规范

1. 访问权限最小化：学生个体画像仅授课教师、课程助教可访问，管理层面只能查看群体汇总统计，不得调取单个学生详细画像；

2. 使用边界：画像**仅限教学干预、课程教学改进，禁止用于评优、评奖、处分等直接评价依据，不得对外输出学生画像数据**。

5.4. 规避标签固化与算法偏见

1. 画像标签为动态线索，不是定性结论，标签会随学生行为数据周度更新，允许学生状态发生正向迁移；

2. 教师使用画像必须结合线下访谈、作业产出，不能仅凭系统标签给学生定性；

3. 设置人工复核机制，对于系统输出的“需要教学关注”类预警标签，教师必须人工核实，避免算法误判对学生造成负面标签化伤害。

6. 总结

学生画像是基于大数据标签聚类系统，聚类不同群体的行为特征，并提供多维度的数据分析，实现对学生群体特点的多视角研判，有利于教育管理者开展差异化教学管理。本文针对线上线下混合教育的过程和特点，优化完善学习画像指标体系，对单一指标局限性进行讨论，采用中性客观的行为描述标签，建立画像动态更新机制；通过真实课堂开展完整的对照实验，证实画像系统对教学干预具有辅助效果；同时专门建立伦理隐私保护框架，规范人脸与学习数据采集使用。本研究对混合式教学下学生分类指导、精准教学有一定的借鉴作用。后续研究将进一步优化算法识别精度，完善多场景下画像应用边界。

基金项目

山东省教育科学“十四五”规划一般课题“线上线下混合式教育学习画像构建方法研究”阶段性成果，课题编号：2023YB089。

参考文献

[1] 冯永刚, 吕鑫源. 数据驱动精准教学的实然困境与应然进路[J]. 现代远距离教育, 2023(3): 30-38.

[2] 赵春, 李欣. 网络学习空间中学习画像的标签模型构建研究[J]. 计算机技术与发展, 2023, 33(10): 176-182.

[3] 肖君, 乔惠, 李雪娇. 基于 xAPI 的在线学习者画像的构建与实证研究[J]. 中国电化教育, 2019(1): 123-129.

[4] 孔德川, 郑芳莹, 叶晖, 等. 基于学习者画像的个性化学习资源推荐算法研究[J]. 河南科技学院学报(自然科学版), 2025, 53(4): 64-72.

[5] 肖丽萍, 张慧, 胡翰. 基于用户画像技术的高职院校学生在线学习精准干预研究与实践[J]. 中国教育技术装备, 2025(17): 93-95+107.

[6] 马婧, 吕纹奇, 周倩. 学习者画像赋能大规模因材施教: 为何与何为[J]. 教育发展研究, 2025, 45(Z2): 28-36.

[7] 赵娅, 周洪鑫, 单可欣, 等. 自主学习导向的虚拟仿真实验学生画像构建研究[J]. 实验室研究与探索, 2025, 44(6): 87-93.

[8] 赵建伟. 基于机器学习的学生画像系统设计与研究[J]. 智能物联技术, 2025, 57(6): 91-95.

[9] 余明华, 张治, 祝智庭. 基于学生画像的项目式学习评价指标体系研究[J]. 电化教育研究, 2021, 42(3): 89-95.

[10] 张可佳, 田枫, 王博, 等. 基于混合智能的学习状态实时采集与动态分析方法[J]. 现代教育技术, 2023, 33(7): 108-117.

[11] 赖英旭, 张雨薇, 庄俊玺. 面向学习者个性化画像的SOR标签体系构建与生成[J]. 计算机科学, 2026, 53(8): 9-19.