ArtCognition：一种基于绘画视觉与运动学线索的情感状态感知多模态 AI 框架

Behrad Binaei-Haghighi、Nafiseh Sadat Sajadi、Mehrad Liviyan、Reyhane Akhavan Kharazi、Fatemeh Amirkhani 和 Behnam Bahrak

arXiv:2601.04297v1 [cs.LG] 7 Jan 2026

摘要—客观评估人类的情感与心理状态是一项重大挑战，尤其是通过非语言渠道进行评估时。本文提出将数字绘画作为一种丰富但尚未得到充分探索的情感感知模态。我们提出了一种名为 ArtCognition 的新型多模态框架，用于自动分析广泛使用的心理测量工具——房树人（HTP）测试。ArtCognition 独特地融合了两种不同的数据流：由计算机视觉模型从最终画作中提取的静态视觉特征，以及从绘画过程本身获取的动态行为运动学线索，例如笔画速度、停顿和平滑度。为弥合低层特征与高层心理学解读之间的差距，我们采用了检索增强生成（RAG）架构。该架构以既有心理学知识为分析提供依据，从而增强可解释性并降低模型产生幻觉的可能性。结果表明，视觉线索与行为运动学线索的融合比单独使用任一模态能够提供更细致的评估。我们发现，提取出的多模态特征与标准化心理指标之间存在显著相关性，验证了该框架作为一种可扩展工具来支持临床医生的潜力。本研究为非侵入式情感状态评估贡献了一种新方法，并为技术辅助心理健康护理开辟了新的途径。

索引词——目标检测、多模态学习、大语言模型、检索增强生成、数字绘画分析、心理评估

I. 引言

MENTAL 心理健康问题影响着全球相当大比例的人口，并造成显著的个人和社会成本。例如，2019年估计“全球每八人中就有一人，即9.7亿人，患有精神障碍” [1]。尽管患病率如此之高，准确诊断仍然面临挑战。当前的临床工具，如《精神障碍诊断与统计手册》（DSM）和《国际疾病分类》（ICD），存在诸多局限，包括标准重叠、二元分类，以及忽视情境或行为因素 [2]。这一诊断缺口在以下情境中尤为突出：心理健康资源有限，专科人员匮乏，导致检测和治疗延误 [3]。因此，亟需创新工具来补充传统诊断实践，提供易于获取、客观且可扩展的评估方法 [4]。

在各种临床评估技术中，投射绘画测验在洞察人类情绪和态度，以及揭示潜在心理动力学方面拥有悠久的历史 [5]。房树人绘画测验由 Buck 于 1948 年 首次提出，至今仍是应用最广泛的投射测量工具之一；在美国心理学会的一项调查中，它在常用心理评估中排名第八 [6], [7]。房树人测验具有多项优势，包括自发性、结构复杂性、非语言表达方式以及文化独立性，这些特点共同增强了其诊断效用 [8], [9]。

然而，传统的绘画评估面临着实践和方法论方面的双重挑战。人工评分耗时较长，而且往往无法捕捉重要的过程层面线索，例如绘画对象的绘画顺序、绘制一个对象所需的动作数量、停顿以及擦除行为。与此同时，心理学正经历快速数字化，数字工具生态系统不断扩展，为心理健康护理提供支持[10], , 推动投射测验及其分析实现人工智能驱动的转型。近期计算机视觉和深度学习领域的进展，使得大规模分析图画成为可能，并提升了分析的客观性和一致性。

基于这些发展，我们提出了一种框架，将数字绘画网络应用与人工智能辅助解读平台相结合，使投射性绘画评估更加易于使用。该系统会自动记录绘图操作，在无需持续专家监督的情况下捕获过程层面元数据。此外，计算机视觉模型能够检测并分类关键组成部分，并通过基于规则的精神分析指标将其归入临床相关类别。随后，描述生成器综合计算机视觉模型的结果以及输入中的过程层面元数据，生成对绘图的全面文本描述。该描述将作为检索增强生成模块的查询，由该模块基于经过组织的心理学知识生成解释。因此，通过结合行为日志记录、对象层面分析和知识引导的推理，使用相关数据片段有助于减少幻觉。ArtCognition 推进了

Behrad Binaei-Haghighi 和 Nafiseh Sadat Sajadi 对本研究作出了同等贡献；Nafiseh Sadat Sajadi 和 Behnam Bahrak 为共同通信作者。Behrad Binaei-Haghighi、Mehrad Liviyan 和 Reyhane Akhavan Kharazi 隶属于伊朗德黑兰的德黑兰大学电气与计算机工程系。NafisehSadatSajadiandBehnamBahrak 隶属于伊朗德黑兰的哈塔姆大学德黑兰高等研究院。Fatemeh Amirkhani 隶属于伊朗德黑兰的阿拉梅·塔巴塔巴伊大学心理学系。

房树人评估，迈向标准化、可扩展且用户友好的工作流程。

II. 相关工作

在解读房树人测验时，John Buck 结合结构化评分系统与评估者对绘画过程的观察来得出临床推断[6]，，尽管传统评分流程耗时较长，且严重依赖专家判断。这些局限性促使人们开发自动化框架，以支持可扩展的评估[11]。

先前研究表明，房树人分析的部分自动化主要聚焦于目标检测和基于规则的特征提取。单阶段检测器（如 YOLO 系列）能够可靠地预测边界框和实例，从而分离房树人元素 [12]。包括 EfficientNet、视觉变换器（ViT）、ConvNeXt 和 Swin Transformer 在内的近期架构，进一步提升了目标分析的准确率[13]–[16]。Lee 等人 [17] 通过计算目标比例和空间位置等特征，根据检测到的元素生成心理分析表，展现了实现更客观、高效评估的潜力。然而，这些方法依赖人工制定规则，并孤立地解读各个元素，最终报告仍需由人类专家生成。

后续研究探讨了利用儿童绘画进行情绪分类。Alshahrani [18] 使用 YOLOv8n-cls 将图画分为四种情绪状态（幸福、悲伤、焦虑和愤怒/攻击性），并凭借紧凑且适合移动设备的架构，实现了 94% 的 Top-1 准确率。然而，该方法仅限于对最终绘图进行分析，忽略了能够提供更丰富解读线索的时间特征。

近年来，大语言模型（LLMs）取得了显著进展，例如 GPT-4 在需要达到人类水平推理能力的任务中展现出强大能力，其中包括认知心理学方面的挑战。研究人员已在 CommonsenseQA、SuperGLUE、MATH 和 HANS 等成熟的认知数据集上评估 GPT-4，结果显示，其准确率高于此前的最先进模型。这些结果表明，大语言模型能够整合上下文信息并模拟人类认知过程的某些方面，凸显了其在支持自动化心理评估与解读方面的潜力 [19]。

自动化房树人分析仍存在两个关键缺口。首先，现有大多数系统仅关注最终绘图，忽略了包含有价值信息的行为元数据 [11]。其次，既有方法通常将识别与解读分开处理，几乎没有框架能够在统一且可解释的处理流程中整合对象级检测、行为分析和心理测量分类。要弥补这些缺口，需要端到端的多模态模型，结合视觉、时间和上下文特征，从而实现可扩展、数据驱动且具有临床意义的心理评估。

III. 方法

本研究采用一种整合图像分析与时序行为数据分析的多阶段方法，弥补了以往研究的不足，从而实现可解释的房树人评估。首先，我们构建了一个数据集，其中包含高分辨率绘画及其对应的细粒度笔画序列，以捕捉静态视觉特征和行为信息。其次，我们采用两阶段目标检测模型，先定位房屋、树和人物等主要组成部分，再通过第二阶段检测识别这些对象中的特定组成部分。第三，我们利用分类模型处理检测到的组成部分，分析特定的心理标记，例如“扑克脸”或倾斜的房屋。第四，我们从输入元数据中提取行为数据，以追踪橡皮擦使用情况和笔画动作等行为模式。第五，描述生成器综合目标检测结果、分类结果和元数据，生成一份关于图画的综合文本总结。最后，这段文本作为检索增强生成模块的查询，由该模块生成以专家心理学知识为依据的最终解读。所提出的架构如图1所示。

A. 数据集

为构建数据集，我们使用数字绘画在整个绘画过程中捕获细粒度元数据。我们从志愿参与者中收集了146个样本，每位参与者均使用定制的基于网络的绘画应用程序，该应用程序旨在记录详细的用户交互。该应用程序以高时间和空间分辨率记录每个绘画动作，从而能够近乎精确地重建绘画过程。对于每个样本，元数据以JSON格式存储，并与PNG格式的最终图像配对 [20]。

每个数据集样本均由完成的绘图及其对应的元数据组成。元数据记录所有绘画动作，并包含绘画顺序、动作类型、颜色、不透明度、时间戳、线条宽度以及描述每条笔画轨迹的点序列等属性。图2展示了一个绘画动作的记录示例。

此外，21名参与者完成了一份补充性的房屋-树木-人物问卷（附录B）。该问卷由研究团队与领域专家合作设计，旨在提供补充的自我报告信息，并提高对图画的可解释性。

B. 目标检测

目标检测阶段采用两阶段层次化方法，涉及四个基于YOLOv11架构的专用模型。第一阶段由主要组件检测器定位房树人元素中的主要对象：房屋、树和人物。该检测器还会识别云、太阳和山脉等环境特征。

在第二阶段，三个独立的组成部分检测器同样采用YOLOv11，分别处理房屋、树和人物的裁剪图像。这些模型能够捕捉细致的解剖和结构细节，例如房屋中的窗户和门、树上的树干和枝条，以及人物的面部特征和四肢。精细检测确保所有相关子组件都被分离出来，以便进一步分析。这些检测结果直接为

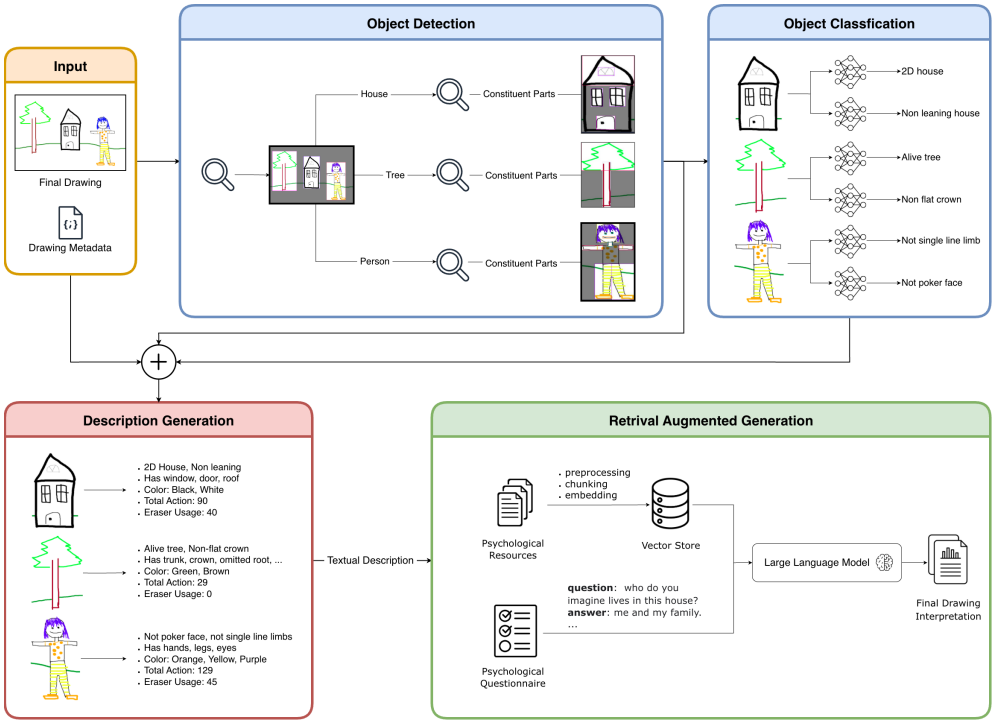

图1：ArtCognition 的工作流程



（a）绘制艺术作品


```
[{
    "order": 1    ,
    "action_type": "绘制线条",
    "color": "#000000",
    "opacity": 1,
    "线宽_": 5,
    "时间戳_开始": 1751293539626,
    "时间戳_结束": 1751293540253,
    "点":  [{
        "x":  333.95,
        "y": 102.76,
        "pointerType": "鼠标",
        "timestamp": 1751293539626
    }]
}, ...]
```


(b) 绘图元数据示例

图 2：数字绘画示例及其对应的元数据。

描述生成模块提供依据，确认特定元素是否存在或缺失。此外，生成的边界框能够进行精确的空间测量，从而揭示重要的心理学洞见。例如，房屋尺寸分为微小、正常或巨大，分别对应绘图面积比例低于1 9、介于1 9 和2 3之间，以及高于2 3 [21]。

C. 对象分类

检测到各个组件及其组成部分后，会通过六个专门的分类模型进行处理，以识别心理标记。这些模型选自高性能架构，包括 EfficientNet、ViT、ResNet50、MobileNetV2、ConvNeXt 和 Swin Transformer，以确保高准确率 [22], [23]。

房屋分类器 两个模型用于分析房屋的结构特征：

- 倾斜房屋分类器：识别房屋是否倾斜，这可能表明结构或情绪不稳定。
- 2二维/三维房屋分类器：确定绘图的视角，以评估空间复杂度。

树分类器 两个模型用于评估树的活力和形状：

- 枯树分类器：区分活树与被描绘为枯死或枯萎的树。
- 扁平化树冠分类器：检测树冠的变形，例如变平，这可作为临床指标 [11]。

人物分类器 最后两个模型分析人物的面部和身体表现：

- 扑克脸分类器：评估面部表情，以检测情绪中性或缺乏情感的状态 [11]。
- 单线条肢体分类器：识别四肢是否以简单线条绘制，这可能表明存在情绪或发展限制 [24]。

训练过程：所有模型均独立训练。首先使用 117 个带标签的样本训练 YOLOv11，以定位 HTP 绘画中的主要对象。随后裁剪检测到的对象，为后续分类器生成训练数据。训练完成后，将所有模型集成到处理流程中，用于端到端检测和绘图细节分析。

D. 元数据分析

如图 2 所示，我们的系统会捕获每个绘画动作的详细元数据，包括动作类型（例如，画线、使用橡皮擦、使用油漆桶），以及每个点的属性，例如以毫秒级分辨率记录的坐标和时间信息，从而能够精确重建绘图 [20]。更重要的是，这些数据支持高级分析，可提供量化的行为学洞见，而这些洞见通常难以或无法由人类或心理学家手动测量。

笔画速度： 通过计算连续点之间的欧几里得距离，得出笔画路径总长度。这种精细测量能够精确计算单个笔画以及每个检测到的对象（例如房屋、树、人物）的绘画速度。笔画总长度L通过对所有连续点对（x_i， y_i）和（x_{i+1}， y_{i+1}）之间的距离 d_i 求和得出。平均笔画速度 v 则通过笔画总长度除以笔画持续时间 T计算得出。速度的计量单位为像素每秒。

$$d_i = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}, \tag{1}$$

$$L = \sum_{i=1}^{n-1} d_i \tag{2}$$

$$v = \frac{L}{T} \tag{3}$$

绘画速度反映了控制运动、心理运动和自动动作的能力 [25]。此外，另一项研究得出结论，发育性协调障碍（DCD）儿童的速度和速度变化等运动学参数较弱 [26]。

笔画间隔： 笔画间隔是指一笔结束到下一笔开始之间的时间间隔，直接从带时间戳的绘画元数据中提取。我们计算停顿的总时长，以及绘图过程中的均值、方差和分布等统计量。笔画之间的这些停顿可能反映认知运动协调、加工迟缓、犹豫和注意力转移。此外，路易体痴呆（LBD）组的绘图速度明显较慢，而阿尔茨海默病组和路易体痴呆组的停顿次数及总时长均有所增加 [27]。ArtCognition 能够精准追踪用户绘图时的停顿和速度，从而为进一步研究特定群体的行为特征提供支持。图 3 展示了这些时间间隔的分布及其中位数。

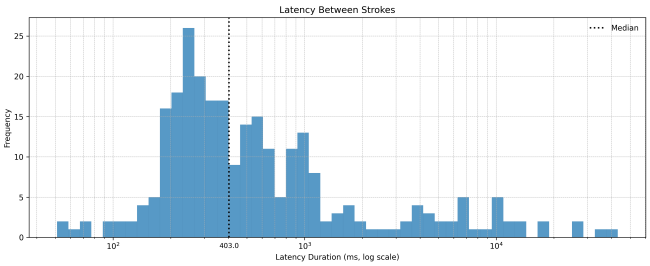

图 3：绘画过程中的笔画间隔。

笔画平滑度： 谱弧长（SPARC）是一种常用指标，通过计算线条归一化傅里叶幅度谱的弧长来量化线条的平滑度。计算时，对归一化频谱积分，直至自适应频率截止值（ω_c）；其中，$\mathbf{V}(\omega)$ 是线条一阶导数（压力变化率）的傅里叶幅度谱，而 $\hat{}\mathbf{V}(\omega)$ 是以其零频率处的值 $V(0)$ 归一化后的频谱。谱弧长（SAL）定义如下：

$$\begin{aligned} \mathrm{SAL} &\triangleq -\int_0^{\omega_c} \left[\left(\frac{1}{\omega_c} \right)^2 + \left(\frac{d\hat{V}(\omega)}{d\omega} \right)^2 \right]^{\frac{1}{2}} d\omega \\ \hat{V}(\omega) &= \frac{V(\omega)}{V(0)} \end{aligned} \tag{4}$$

这种方法中，数值越接近零表示平滑度越高；该方法已用于分析运动学信号，尤其用于量化吞咽研究中咽部高分辨率测压曲线的平滑度 [28]。

绘制对象的位置：神经心理学研究表明，位置模式会随年龄和空间加工差异而变化 [29]。例如，儿童往往会将图画画在中心略偏左的位置，而惯用右手的人表现出更明显的左偏差 [30], ，类似倾向也出现在不同文化中。例如，在流域绘图中，儿童经常将河流描绘成从左向右或向下流动 [31]。

为了量化位置，我们将画布划分为 3×3 网格，以捕捉整体绘图位置和各个对象的位置 [32]。每个笔画由多个点组成，我们计算这些点在各网格单元中的归一化分布：

$$P_{ij} = \frac{n_{ij}}{\sum_{k=1}^{3} \sum_{l=1}^{3} n_{kl}}, \quad i, j \in \{1, 2, 3\} \tag{5}$$

其中， n_{ij} 表示单元格（i， j）中的点数。由此得到概率分布 P_{ij}，可用于比较尺寸和密度各异的图画。

图4展示了九个区域内点的分布，为绘图位置模式提供了定量表示。

橡皮擦使用模式： 频繁擦除可能表明存在焦虑、自我怀疑和完美主义 [33]。另一方面，一项 2020年研究比较了实体与数字绘图环境中的擦除行为，发现数字环境中的擦除行为更频繁，这可能是因为“撤销”功能使用方便 [34]。这些差异并未显著影响最终图画的可解释性，这表明橡皮擦

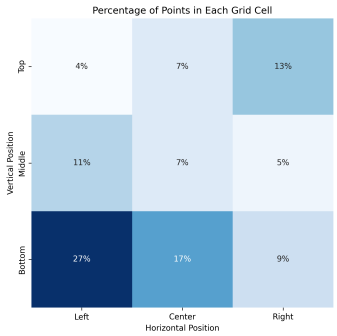

图 4：绘画点在 3×3 画布网格上的分布。

的使用反映的是补偿行为，而非诊断标记。

在 ArtCognition 中，橡皮擦行为会从多个维度进行分析。我们利用目标检测模型生成的边界框，统计每个对象的擦除事件次数，以此捕捉用户对特定元素的敏感性。我们还会计算总擦除时长和累计擦除面积，从而呈现绘画过程中修改策略和自我监控的行为概况。通常，重描的铅笔线条会引发自我批判和自我怀疑的感受 [35]。因此，我们的模块能够利用元数据深入了解用户的橡皮擦使用模式。此外，借助 YOLO 检测到的边界框，我们的模块可以确定用户使用橡皮擦修改每个对象的次数。

绘画顺序：在房树人测验中，受测者选择先画房屋、树还是人物，其顺序被视为一种行为指标，反映个人自发的注意力焦点和情绪优先事项 [36]。此外，先画人物通常与较强的自我关注、身体意象关注或人际敏感性有关，尤其是在伴有细致修改或犹豫时。先画房屋通常与对家庭关系、安全需求或家庭事务的关注有关。先画树有时被解读为反映了对活力、力量或内在资源的关注，尤其是在儿童中 [37]。利用 YOLO 检测到的边界框以及元数据中的动作时间戳，我们的模块可以确定每个绘制对象的绘画顺序和完成时间。

线条宽度：绘画界面支持调整画笔和橡皮擦的宽度。在投射绘画研究中，粗重的线条通常与紧张、攻击冲动、挫折感或强烈的情绪激活相关。相反，非常轻淡、浅淡的线条可能反映焦虑、不安全感或精力不足 [38]。

除了目标检测之外，我们还利用捕获的元数据重建完整的图画。该重建过程验证了记录数据的一致性，并支持对绘画过程进行可视化回放，从而开展能够捕捉草图时间演变的定量分析，而不仅仅关注其最终结果。图 5 展示了原始图画与其重建结果之间的对比。



(a) 原始绘图　　(b) 重建后的绘图

图 5：原始绘图及其基于元数据的重建图。

E. 描述生成

描述生成模块充当一个综合模块，将各种数据流转换为绘图的结构化文本表征。该模块整合了三类主要信息来源，包括目标检测结果、分类标记和行为元数据。

首先，该模块利用两阶段目标检测的结果，识别存在哪些元素或缺少哪些元素。对于每个检测到的对象，它结合六个分类模型的结果，描述具体的定性特征，例如倾斜的房屋或扑克脸。其次，对从绘图元数据中提取的数据进行进一步归纳，生成结构化文本描述。

这一结构化描述是原始视觉／行为数据与最终心理学解读之间的关键桥梁。通过结合绘制的对象与绘画过程，生成器为后续的检索增强生成阶段提供全面的查询。绘图描述的可视化示例见图 6。

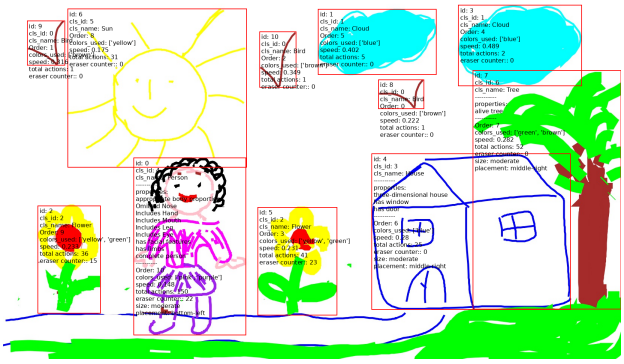

图 6：ArtCognition 标注对象并记录绘画行为的示例。

F. 基于 *RAG*架构的心理学解读

检索增强生成架构使系统能够将原始绘图描述转化为专家级的心理学洞见。通过将前序模块综合得出的数据与经过筛选的知识库相结合，系统能够生成有依据的解读，而非孤立的观察结果。这种方法确保最终输出既以证据为基础，又符合既定的临床框架。

1）数据预处理：检索增强生成知识库由权威心理学资料构建，包括《投射性绘画解读：自我心理学》

方法 [35]以及Guo等人的研究 [11]。通过去除非ASCII字符清理原始文本，随后进行停用词去除和词形还原。为提高检索精度，按照章节标题和小标题对语料库进行分类，从而得到针对下游任务优化的标准化数据。

2）数据分块：我们采用四种策略对语料库进行分段，以评估检索性能：a）字符级分块，生成固定长度的文本片段；b）递归字符切分，沿语义边界（段落、句子）递归划分，直至达到尺寸阈值；c）语义分块，根据句子嵌入的相似度对句子进行分组；以及 d）语义聚类，在句子嵌入上应用K均值，形成语义连贯的分块。

3）使用大语言模型生成解读：我们采用提示工程，将检索到的上下文提供给大语言模型。通过以特定领域的参考资料为大语言模型提供条件约束，使生成的解读以经过验证的心理学知识为依据，减少幻觉（这是临床人工智能领域的一项关键顾虑），并生成易于解读且具有上下文感知能力的报告。

IV. 结果

A. 目标检测

我们使用一系列绘图元素（包括房屋、树木、人物及其组成部分）对目标检测模型进行评估。表I报告了所有类别的性能。房屋、树木和人物等主要元素的精确率和召回率较高，而鸟类等较小或较少出现的元素，其检测准确率则较低。

表I：YOLOv11 在主要目标检测任务中的性能。

类别	#	P	R	mAP@50	mAP@50-95
All	103	0.898	0.897	0.949	0.833
Bird	4	1.000	0.710	0.768	0.421
云	15	0.963	0.867	0.956	0.895
花	3	0.567	1.000	0.995	0.908
房屋	21	0.977	1.000	0.995	0.941
人物	22	0.903	0.909	0.968	0.848
Sun	6	0.949	1.000	0.995	0.949
Tree	23	0.886	0.913	0.978	0.902
烟囱烟雾	9	0.942	0.778	0.939	0.802

组成部分的检测结果汇总于表 II–III-IV。对于房屋（表 II），窗户和屋顶表现出色，mAP@50-95值分别为0.92和0.84；门的准确率也很高（0.87）。烟囱的表现相对较弱（mAP@50-95 $= 0.61$），这可能是由于其尺寸较小且样本有限。对于人（表 IV），头部的检测最为可靠（mAP@50-95 $= 0.73$），其次是眼睛、腿和嘴巴；而鼻子和颈部等更精细的组成部分表现较弱，这反映出它们空间范围较小且变化性较大。对于树木（表 III），树冠展现出稳健的检测性能（mAP@50-95 $= 0.95$），树干和树根的准确率中等，而枝条和果实得分较低，这与它们形状多样且标注有限的情况相符。

表II：YOLOv11 在房屋部件检测上的性能。

类别	#	P	R	mAP@50	mAP@50-95
All	83	0.965	0.836	0.940	0.808
烟囱	11	0.911	0.636	0.830	0.610
Door	20	0.993	0.850	0.968	0.866
Roof	16	0.987	0.938	0.983	0.837
窗户	36	0.971	0.920	0.980	0.917

表 III：YOLOv11 在树木部件检测上的性能。

类别	#	P	R	mAP@50	mAP@50-95
All	63	0.725	0.931	0.875	0.668
枝条	8	0.495	0.875	0.745	0.494
树冠	23	0.968	1.000	0.995	0.950
果实	3	0.316	1.000	0.806	0.672
Root	1	0.928	1.000	0.995	0.597
树干	28	0.916	0.782	0.833	0.628

总体而言，目标检测模型在主要元素上表现稳健，并能令人满意地检测组成部分，为从HTP绘画中提取下游特征提供了可靠基础。

B. 分类模型

在心理学解读中，研究人员针对六项语义分类任务（例如，倾斜的房屋与非倾斜的房屋）对房屋、树木和人物进行了标注。研究评估了六种模型：ConvNeXt-Base、ViT-B/16、Swin Transformer、EfficientNet-B0、ResNet50 和 MobileNetV2，采用准确率、精确率、召回率和 F1分数作为评估指标（表VII–VIII）。每项任务中表现最佳的模型均被整合到处理流程中。

C. 心理学解读

为确保我们的框架具有临床相关性，我们评估了基于绘图生成的描述的精确率，以及检索增强生成处理流程的两个核心组成部分：心理背景的检索和最终评估报告的生成。

表IV：YOLOv11 人体部位检测性能。

类别	#	P	R	mAP@50	mAP@50-95
All	182	0.822	0.762	0.825	0.601
Eye	39	0.802	0.949	0.923	0.708
Hand	41	0.794	0.656	0.783	0.649
Head	21	1.000	0.995	0.995	0.731
Leg	39	0.892	0.795	0.884	0.686
嘴	19	0.813	0.842	0.908	0.746
Neck	14	0.906	0.688	0.827	0.333
Nose	9	0.547	0.406	0.456	0.352

表 V：倾斜房屋分类器的性能。

模型	准确率	精确率	召回率	F1分数
ConvNeXt-Base	**0.952**	**0.975**	0.750	**0.821**
ViT-B/16	0.905	0.452	0.500	0.475
Swin Transformer	0.857	0.639	0.697	0.659
EfficientNet-B0	0.810	0.447	0.447	0.447
ResNet50	0.810	0.667	**0.895**	0.691
MobileNetV2	0.857	0.450	0.474	0.462

表 VI：二维房屋分类器的性能。

模型	准确率	精确率	召回率	F1分数
ConvNeXt-Base	0.619	0.683	0.653	0.611
ViT-B/16	0.571	0.607	0.597	0.568
Swin Transformer	**0.810**	**0.846**	**0.833**	**0.809**
ResNet50	0.476	0.486	0.486	0.476
MobileNetV2	0.714	0.708	0.708	0.708
EfficientNet-B0	0.524	0.543	0.542	0.523

表 VII：枯树分类器的性能。

模型	准确率	精确率	召回率	F1分数
ConvNeXt-Base	**0.970**	**0.978**	**0.955**	**0.965**
ViT-B/16	0.788	0.879	0.682	0.698
Swin Transformer	0.848	0.907	0.773	0.802
EfficientNet-B0	0.303	0.273	0.250	0.260
ResNet50	0.909	0.907	0.886	0.895
MobileNetV2	0.788	0.879	0.682	0.698

描述生成器性能： 我们首先评估描述生成器的准确率。该生成器将基于视觉的检测结果和绘画动态转化为结构化语义描述，并将其作为检索模块的输入。对于每幅绘图，生成器都会生成一组对象级描述，分别对应房树人的核心元素（房屋、树和人物）。每条描述都包括空间定位（边界框和位置）、类别属性（例如对象类型、尺寸和维度）、视觉属性（例如是否存在窗户或门、色彩运用），以及从绘画过程中提取的行为特征（例如绘画顺序、笔画速度和动作数量）。

为定量评估描述质量，标注人员人工审阅了生成的描述，并统计每个对象中错误检测或错误归属的特征数量。如果某个特征被错误检测（例如，在不存在窗户时报告有窗户）、虽有明确视觉证据却被遗漏，或描述不准确（例如，位置或尺寸类别错误），则视为错误特征。根据这些标注，我们计算测试数据集上的平均精度，定义为正确检测的特征数占预测特征总数的比例。

在整个评估集中，描述生成器的平均精度达到 97.57%，表明其具有很高的

表 VIII：扁平化树冠分类器的性能。

模型	准确率	精确率	召回率	F1分数
ConvNeXt-Base	0.818	0.474	0.429	0.450
ViT-B/16	0.909	0.476	0.476	0.476
Swin Transformer	0.864	0.475	0.452	0.463
EfficientNet-B0	**0.955**	**0.477**	**0.500**	**0.488**
ResNet50	0.909	0.476	0.476	0.476
MobileNetV2	0.955	0.477	0.500	0.488

表IX：扑克脸分类器的性能。

模型	准确率	精确率	召回率	F1分数
ConvNeXt-Base	0.783	0.450	0.429	0.439
ViT-B/16	**0.870**	**0.455**	**0.476**	**0.465**
Swin Transformer	0.783	0.450	0.429	0.439
ResNet50	0.783	0.450	0.429	0.439
MobileNetV2	0.565	0.433	0.310	0.361
EfficientNet-B0	0.826	0.452	0.452	0.452

表 X：单线肢体分类器的性能。

模型	准确率	精确率	召回率	F1分数
ConvNeXt-Base	**0.864**	0.771	**0.819**	**0.790**
ViT-B/16	0.864	**0.781**	0.722	0.745
Swin Transformer	0.864	0.771	0.819	0.790
EfficientNet-B0	0.864	0.771	0.819	0.790
ResNet50	0.818	0.675	0.597	0.614
MobileNetV2	0.727	0.542	0.542	0.542

一致性，即生成的描述与人工标注高度吻合。错误主要出现在细粒度视觉属性上，例如维度模糊不清（二维与三维表示难以区分）以及细微的结构元素；而核心的对象识别、空间位置和绘图顺序特征则始终能够以很高的准确率被检测出来。这些结果表明，生成的描述能够对图画提供可靠且基于语义的表示，为后续检索和心理学解读奠定稳固的地基。

检索性能：其次，我们评估了检索器为特定视觉描述检索相关心理学解读的准确率。我们为多个查询集建立了真实标注，以衡量检索片段的语义对齐情况。如表XI所示，我们比较了多种分块策略。语义分块的余弦相似度最高（0.991），其次是K均值语义聚类（0.989）。这些方法的表现显著优于基于字符的切分，表明语义连贯的分段对于在检索过程中保持心理学概念的完整性至关重要。

表XI：不同分块策略下检索增强生成的余弦相似度。

分块策略	余弦相似度
字符级分块	0.978
递归字符文本切分	0.961
语义分块	**0.991**
使用K均值进行语义聚类	0.989

生成质量保证和临床有效性： 我们通过与两个基线进行受控比较，评估生成的解读结果的质量和可靠性：(1) 遵循Guo等人 [11], 提出的方法的基于规则的报告系统；(2) 在不使用检索增强的情况下运行的标准gemini-2-flash模型。所有输出均经过匿名化处理，并由持证临床心理学家在盲法环境下进行审阅，以评估结构准确性、理论一致性和实际可解释性。

从计算机视觉的角度来看，所提出的框架在结构化提取视觉信息方面展现出明显优势。该模型明确将每幅HTP绘画分解为具有语义意义的组成部分（房屋、树和人物），并利用特定的视觉属性对其进行分析，包括空间配置（例如左侧—中心—右侧的位置）、几何维度（二维与三维表示）、颜色分布以及特征完整性。这种结构化分解通过以可量化指标取代印象式观察，回应了投射测验长期以来关于可复现性的批评。

至关重要的是，将绘画特征与检索增强生成机制相结合，显著提高了输出的可靠性。标准的gemini-2-flash基线模型经常会产生幻觉，例如解读实际不存在的特征，或赋予文献中未提及的含义；而我们的框架通过将分析锚定于经过验证的心理学证据和参与者自己的口头描述，有效缓解了这一问题。为对此进行定量评估，我们测量了幻觉率，其定义为涉及不受支持的视觉或行为特征的解释性主张所占的比例。标准的gemini-2-flash基线模型的幻觉率约为45.72%。通过利用RAG架构将视觉检测结果与既有文献进行交叉比对，我们的框架将错误率降至零，因为它严格依赖检索到的参考数据；任何残余错误均仅源于目标检测器或分类器。

除了结构准确性之外，生成的解读还体现了对高层次理论构念的扎实运用。视觉指标并非只是被描述，而是始终与心理状态相联系；例如，空间朝向和特征缺失会对应到自我发展、情感收缩和感知的内在效能等构念。

此外，该模型将图像数据与参与者自我报告的“言语归因”（例如对树的年龄的感知或房屋的情感意义）相结合，从而提供临床价值。系统会明确识别这两种模态之间的一致与分歧之处。例如，自我报告的“平静”感与抑郁色调的视觉标记之间若存在差异，系统会将其指出，但不会视为错误，而是视为可能存在防御过程或情绪觉察有限的临床相关指标。

最后，这些解读与标准化心理测量指标表现出有意义的对应关系。对颜色变化有限、四肢几何形态呈线性以及人物简化等视觉指标的解读，与贝克焦虑量表（BAI）和贝克抑郁量表（BDI）的得分升高相一致。这种一致性表明，基于 AI 的分析并非随意而为，而是与成熟测量工具具有聚合效度，可反映情绪困扰。

总体而言，结果表明，与缺乏依据的大语言模型相比，ArtCognition生成的解读更有条理、可复现，且有证据支撑。该框架虽不能替代临床专家的判断，但通过将投射分析建立在可观察的运动学和视觉证据之上，提高了透明性。

五、讨论

本研究提出了一套可靠且新颖的自动化HTP分析处理流程，整合了目标检测、分类、运动学元数据和RAG架构。我们探讨了该方法的有效性、数据集与方法的局限性，以及人工智能辅助心理健康评估中的伦理考量，并指出了未来研究方向，以及将视觉模型与行为元数据相结合以实现可扩展、可解释评估的潜力。

A. 临床适用性与可信度

投射绘画测验广泛应用于心理学领域，尤其适用于评估难以用语言表达情绪的儿童 [17], [39]。以往的自动化HTP解释方法往往依赖非标准化评分系统和人工制定规则，导致输出结果不一致，进而可能削弱临床医生的信任 [9], [40]–[43]。

ArtCognition通过整合行为元数据与视觉特征来增强可解释性，从而应对这些局限性。它借助检索增强生成生成以知识为依据的报告，为临床医生提供透明且可验证的分析依据，同时避免黑箱预测 [44], [45]。通过自动化特征提取与解读，该系统减少了人工评分所需的时间和精力 [46]，使临床医生能够专注于治疗规划和干预。这种高效性有助于扩大心理健康服务的覆盖范围，而系统的适应性和透明的输出则能增强临床医生信任，推动其更广泛地应用。此外，该设计遵循可信医疗人工智能的关键原则，包括可解释性和隐私保护 [47]。

B. 局限性

尽管取得了令人鼓舞的结果，本研究仍面临若干关键局限，包括数据集规模较小、缺少行为模态、文化背景单一，以及自动化解读方面的限制。这些挑战指明了未来研究的重要方向。

数据集规模： 数据集相对较小，限制了模型性能，尤其是对罕见绘画模式的处理能力。数据稀缺也降低了检测和分类任务的泛化能力。

数字笔压力： 未采集笔压和倾斜角度数据，因此无法获得与神经运动行为相关的洞见，也限制了心理学解读的丰富性。

文化与人口统计特征的同质性： 所有参与者均来自同一文化背景，这可能引入偏差，并限制模型对更广泛人群的普适性。

解读方面的局限： 尽管可以精确量化图画的高分辨率特征，但支持对许多特征进行可靠解读的临床证据有限。此外，某些心理细微差异仍难以通过计算方式捕捉，这凸显了持续需要专家监督。

C. 伦理考量

自动化心理评估引发了有关隐私和负责任使用的重要担忧。解决这些问题对于安全、合乎伦理地部署数字心理健康应用至关重要。

VI. 结论

我们提出了一种先进的人工智能流程，用于自动分析数字化 HTP 绘画，整合了目标检测、分类、行为元数据，以及用于心理学解读的 RAG架构。通过结合视觉特征、运动学数据和外部心理学知识，

该系统提供了超越传统人工评分的可解释且可扩展的评估。实验结果表明，该系统具有较高的检测和分类准确率；此外，元数据分析还能捕捉细微的行为指标，例如绘画速度、线条平滑度和绘画顺序。

主要局限性包括规模较小且文化背景同质化的数据集、缺少笔压数据，以及量化复杂心理状态所固有的挑战。未来的工作将扩展数据集，纳入其他神经运动模态，并在不同人群中验证该方法。总体而言，本研究展示了AI辅助数字绘画分析作为心理评估补充工具的潜力，可提供高效性、可解释性和基于证据的洞见。

致谢

作者谨此感谢Shayan Talebiani对数据标注的贡献，以及其实现了用于计算与每个对象相关的动作数量的基于元数据的模块。

附录A 数据集

本附录呈现了全面的数据集统计信息，详细说明对象类别及其组成部分的出现频率。这些指标旨在促进可复现性，并为正文中讨论的实验结果和分析提供必要的背景信息。表XII呈现了数据集中主要对象类别的出现频率，展示每个高层级对象（例如房屋、树、人物）出现的次数。该表概述了数据集的整体构图，并有助于说明类别平衡。

表XII：数据集中主要对象类别的分布。

类别	实例数
Bird	66
云	112
房屋	147
Tree	181
人物	156
花	46
山	14
Sun	63
烟囱烟雾	45

表 XIII：数据集中房屋、树和人物对象类别的分布。

（a）房屋

类别	# 实例
Door	135
窗户	249
Roof	107
烟囱	58

(b) 树

类别	# 实例
枝条	40
树冠	158
果实	144
Root	21
树干	180

(c) 人物

类别	# 实例
Eye	213
Hand	266
Head	134
Leg	258
嘴	105
Neck	74
Nose	37

附录B

房屋-树木-人物问卷

这份问卷补充了有关参与者对其绘图解读的背景信息。它由临床专家设计，旨在了解参与者对房屋、树和人物图像的主观感知。

A. 参与者信息

姓名：年龄：性别：

B. 问题

1) 你认为谁住在这栋房屋里？2) 这栋房屋带给你什么感受？3) 你认为这棵树有多大年纪？4) 这棵树是活着还是已经枯死？5) 你认为现在是一年中的哪个季节？6) 这幅图像让你想起某个人吗？7) 你认为这个人物有多大年纪？8) 你认为这个人物是做什么的？9) 这个人物可能在想什么？10) 你认为这个人物有什么感受？

附录 C

使用大语言模型进行绘图解读的示例

图 7 展示了用于示例分析的绘图，作为主要输入。原始模型仅查看静态图像，而 ArtCognition 则整合多个数据层，生成更全面的评估。以下各节将比较基础解读与 ArtCognition 提供的进阶输出。



图 7：用于比较原始模型与 ArtCognition 解读方法的绘图示例。

A. 原始解读

1) 总体印象： 这幅绘图描绘了一个由房屋、树和人物组成的简单多彩场景，表明其关注生活和环境中的基本元素。明亮色彩和微笑的面容体现出总体上积极的情感。构图似乎较为分散地铺展在页面上，可能表明需要更多的结构或界限。*2)* 房屋：

- 颜色：房屋的粉色可能代表温暖、亲昵，或与女性气质相关联。
- 形状：房屋采用简单、传统的形式绘制，表明其遵循规范并渴望稳定性。
- 窗户和门：窗户和门的存在表明开放性，体现出社交互动的意愿，以及

与外部世界建立联系的愿望。窗户宽大且轮廓清晰，进一步强调了这种开放性。

• 地基：房屋似乎由一条起伏的绿色线条支撑，这可能代表环境以及与自然相连的感受。

3）树：

• 尺寸和形状：这棵树很大，树冠饱满圆润，暗示着强烈的成长感、活力和情感表达。

• 树干：树干由两条线条描绘，可能象征着力量和稳定性。

• 位置：这棵树位于房屋和人物之间，可能象征着家庭环境与个人自我表达之间的桥梁。

4）人物：

• 面部表情：微笑的脸庞清楚地表明了幸福和积极的心态。

• 头发：卷发可能代表精力、创造力，或渴望受到关注。

• 手臂和腿：手臂向外伸展，表明愿意与外界互动。腿部线条简单，体现出一种脚踏实地感。

• 服装：连衣裙可能代表女性气质，或体现自我呈现。

•鞋子：粉色鞋子增添了一丝奇趣和个性。

5）色彩运用：使用粉色、绿色和红色等明亮色彩，暗示着乐观、精力充沛，以及对刺激的渴望。

6) 页面上的位置：绘图位于页面的下半部分，这可能表明一种安全感或对支持的需求。

7) 局限性：必须谨记，这只是完全基于绘图得出的基础解读。全面分析还需要有关创作者背景、经历和情绪状态的额外信息。此解读应被视为进一步探索的起点，而非确定性的诊断。

B. ArtCognition 的解读

1) 总体印象：这幅绘图描绘了房屋、树和人物，位于页面中央，表明绘画者具有安全感，并且能够与当下建立联系。然而，恰好位于中心的位置也可能引发对平衡和僵化的需求。绘图的简化特征，加上缺乏额外装饰，可能表明心理动力低下，或存在回避和退缩的倾向。正常的着色表明绘画者能够表达广泛的主观体验。

2) 房屋：

• 总体：立体房屋象征着一种表现深度的尝试。绘画者将房屋视为自己的童年之家，并将其与幸福和没有悲伤联系起来，这表明她可能对自己的过去持积极看法。

• 门：门的存在表明房屋具有一定的可进入性，但没有门把手则传达出一种不确定或矛盾心理，反映出对是否让他人进入的犹豫。

• 窗户：窗户的存在表明绘画者有能力与周围环境建立关联并进行互动。

•颜色：提供的文本中并未直接提及房屋的粉色，因此难以对此作出解读。

3) 树：

•总体：树是有生命的，也描绘了夏天，因此象征着活力与鲜活感。绘画者认为这棵树有20岁，而她29岁，这表明她感受到的成熟度有所减弱。

• 颜色：绿色唤起鲜活感和开阔感。*4)* 人物：

• 总体：人物形象完整，身体比例恰当，表明自我整合感完好。人物的尺寸适中，说明其胜任感处于最佳水平。这个形象让绘画者想起小时候的自己，强化了她与童年的联系。

• 面部特征：面部特征的出现表明个体关注自己的外表以及自己在他人面前的表现。然而，鼻子的缺失暗示个体在如何呈现自己、如何接收情感刺激方面存在严重障碍。

• 四肢：画出四肢（腿和手）表明个体具备机动能力，并能够与环境建立联系。

• 颜色：人物身上使用粉色并无直接含义，但可以作出一般性观察：彩色绘图能够带来更丰富的表达空间。

5) 颜色：

• 绿色：使用绿色暗示鲜活感和开阔感。

• 白色：白色背景暗示自我意识和心理防御的缺失。

• 粉色：所提供的文本中未提及颜色。*6)* 位置：页面中央位置表明不——
性。

7) 总结： 这幅绘图表明，绘画者与童年有所关联，并将其与幸福联系起来。绘图中展现出活力和渴望建立联系的意愿，但也透露出对可接近性的矛盾心理，以及一种倾向于退缩的表现，这可能是一种防御机制。省略鼻子并采用简化的风格，表明绘画者可能存在一些与自我感知及与外界互动相关的潜在焦虑。人物居中且尺寸适中，表明其具有基本的安全感和胜任感，不过对平衡与控制的需求或许也是一个影响因素。

C. 准确度分数计算

为量化解释性幻觉，我们基于专家验证衡量准确度分数。对于每项生成的解读，我们将其中的各项主张拆解为原子解释单元。根据领域专家（心理学家）的反馈，每个单元会被标记为正确 或错误 。

并依据既定的房树人方法规则进行判断。除非遗漏内容被错误推断取代，否则不予扣分。例如，在一个具有代表性的案例中，共识别出 13 个解释单元，其中 5 个被判定为正确，8 个因缺乏依据的象征性或情感推断而被判定为错误。最终准确度分数约为 38。5%；该数值越低，表示原始解读中出现幻觉的倾向越高。

附录D

房树人绘画特征

表XIV总结了房树人测验中所分析的绘画特征及其相关心理学解读的详细信息。我们使用此表将我们的方法与Guo等人的研究进行比较， [11], 重点介绍可观察到的绘画特征及其对应的推断心理构念。

表XIV：绘画特征及其相关心理学解读。

绘画特征	指示含义
物品之间过度分离；遗漏 房屋、树木或人物；没有门或窗户； 面部特征缺失／扑克脸；完整 或部分肢体缺失；人物绘制不完整； 页面左侧位置／左上角位 置；颜色：白色	自我意识丧失与心理防 御机制
倾斜的房屋；枯树；扁平树冠； 不恰当的身体比例；拳头；高位或 页面右侧位置；颜色：紫色、棕色	心理冲突与.....感 不真实感
冒烟的烟囱；根部；颜色：黄色， 紫色；页面顶部边缘位置	紧张、敏感性和易怒—— 易怒
无额外装饰；简化的绘画—— 绘图；绘画尺寸较小；强调直线 线条；非常小的房屋 / 树 / 人物；二维 房屋；单线条四肢；缺乏 颜色；页面位置偏低；颜色： 白色；浅淡线条；页面左侧位置	心理动机低下、回避， 和退缩
页面左侧位置；页面位置偏低； 颜色：棕色、　白色；左上角 位置	退行
页面中央位置；颜色：橙色， 绿色、蓝色	正常性
页面位置偏低；绘画尺寸较小； 线条非常淡；颜色：白色	抑郁、空虚
页面位置偏低；绘画尺寸较小； 颜色：棕色；页面左上边缘位置	不安全感
大幅绘画；粗重线条	攻击性
页面左侧位置；绘画尺寸较小； 线条非常淡；页面位置偏低；颜色： 绿色	自尊，幼稚
纸张底边	需要外部支持，依赖—— 密度
纸张侧边；大幅绘画；过度 使用颜色	环境限制、压力
大幅绘画；颜色：红色、橙色；页面上部 位置	活力增强、精力、躁狂 状态

参考文献

[1] 世界卫生组织，《精神障碍情况说明》，https://www.who.int/news-room/fact-sheets/detail/mental-disorders，2019年，访问日期：2025年9月11日。

[2] A. M. Kilbourne、K. Beck、B. Spaeth-Rublee、P. Ramanuj、R. W. O'Brien、N. Tomoyasu 和 H. A. Pincus，《衡量并改善心理健康照护质量：全球视角》，《世界精神病学》，第17卷，第1期，第30–38页，2018年。[3] F. Hanna、C. Barbui、T. Dua、A. Lora、M. van Regteren Altena 和 S. Saxena，《全球心理健康：我们做得如何？》，《世界精神病学》，第17卷，第3期，第367页，2018年。[4] L. Balcombe 和 D. De Leo，《数字心理健康面临的挑战与解决方案的未来》，《JMIR心理健康》，第8卷，第3期，e26811，2021年。[5] G. C. Gleser，《投射法》，《心理学年度评论》，第14卷，第1期，第391–422页，1963年。[6] J. N. Buck，《房树人技术：定性与定量评分手册》，《临床心理学杂志》，1948年。[7] W. J. Camara、J. S. Nathan 和 A. E. Puente，《心理测验的使用：对专业心理学的启示》，《专业心理学：研究与实践》，第31卷，第2期，第141页，2000年。[8] H. Smeijsters 和 G. Cleven，《在司法精神病学中运用艺术治疗攻击性：定性研究结果》，《心理治疗艺术》，第33卷，第1期，第37–58页，2006年。[9] L. Sheng、G. Yang、Q. Pan、C. Xia 和 L. Zhao，《合成房树人绘画测验：一种筛查癌症患者焦虑的新方法》，《肿瘤学杂志》，第2019卷，第1期，5062394，2019年。[10] T. Ostermann、J. P. R¨oer 和 M. J. Tomasik，《心理学数字化：一点挑战，一字节成功》，《Patterns》，第2卷，10，2021年，第 .[11] H. Guo、B. Feng、Y. Ma、X. Zhang、H. Fan、Z. Dong、T. Chen 和 Q. Gong，《房树人测验筛查和预测精神障碍特征的分析：系统综述与荟萃分析》，《精神病学前沿》，第13卷，1041770，2023年，第 .[12] J. Redmon、S. Divvala、R. Girshick 和 A. Farhadi，《YOLO：统一的实时目标检测》，载于《IEEE计算机视觉与模式识别会议论文集》，2016年，第779–788页。[13] M. Tan 和 Q. Le，《EfficientNet：重新思考卷积神经网络的模型缩放》，载于《国际机器学习会议》。PMLR，2019年，第6105–6114页。[14] A. Dosovitskiy、L. Beyer、A. Kolesnikov、D. Weissenborn、X. Zhai、T. Unterthiner、M. Dehghani、M. Minderer、G. Heigold、S. Gelly 等，《一张图像胜过16×16 个词：用于大规模图像识别的Transformer》，arXiv预印本 arXiv:2010.11929，2020年。[15] Z. Liu、Y. Lin、Y. Cao、H. Hu、Y. Wei、Z. Zhang、S. Lin 和 B. Guo，《Swin Transformer：采用移位窗口的层次化视觉Transformer》，载于《IEEE/CVF国际计算机视觉会议论文集》，2021年，第10 012–10 022页。[16] J. Feng、H. Tan、W. Li 和 M. Xie，《Conv2Next：重新审视用于图像识别的 ConvNext 网络设计》，载于《2022年计算机与人工智能技术国际会议（CAIT）》。IEEE，2022年，第53–60页。[17] M. Lee、Y. Kim 和 Y.-K. Kim，《运用深度学习为儿童绘画生成心理分析表》，《数据与知识工程》，第149卷，102266，2024年。[18] A. Alshahrani、M. M. Almatrafi、J. I. Mustafa、L. S. Albaqami 和 R. A. Aljabri，《一种基于计算机视觉和深度学习、通过绘画分析检测儿童心理与心理健康状况的模型》，《工程、技术与应用科学研究》，第14卷，第4期，第15 533–15 540页，2024年。[19] S. Dhingra、M. Singh、N. Malviya、S. S. Gill 等，《思维邂逅机器：揭示 GPT-4 的认知心理学》，《BenchCouncil基准、标准与评估汇刊》，第3卷，第3期，100139，2023年。[20]《Htp painter（v1.2）》，https://github.com/behradbina/Paint，2025, [在线；访问日期：19-Aug-2025]。[21] M. Takahashi，《绘画测验导论：房树人测验》，韩国色彩心理分析研究所，第29–34页，2009年。[22] M. Sandler、A. G. Howard、M. Zhu、A. Zhmoginov 和 L.-C. Chen，《MobileNetV2：倒残差与线性瓶颈》，《2018年IEEE/CVF计算机视觉与模式识别会议》，第4510–4520页，2018年。 [在线]。可访问：https://api.semanticscholar. org/CorpusID:45552077

[23] K. He、X. Zhang、S. Ren 和 J. Sun，《用于图像识别的深度残差学习》，《2016年IEEE计算机视觉与模式识别会议（CVPR）》，第770–778页，2015年。 [在线]。可访问：https://api.semanticscholar.org/CorpusID:206594692

[24] M. C. Meehan，《儿童人物绘画的心理评估》，《美国医学会杂志》，第205卷，第3期，第190–190页，1968年。

[25] S. Rueckriegel、F. Blankenburg、R. Burghardt、S. Ehrlich、G. Henze、R. Mergl 和 P. Driever，《年龄与动作复杂度对童年和青春期手部运动学参数的影响》，《国际发展神经科学杂志：国际发展神经科学学会官方期刊》，第26卷，第655–63页，2008年09月。[26] A. Abu-Ata、D. Green、R. Sopher、S. Portnoy 和 N. Ratzon，《发育性协调障碍儿童与非患儿的上肢运动学》，Sensors，第22卷，第9224页，2022年11月。[27] Y. Yamada、M. Kobayashi、M. Nemoto、M. Ota、K. Nemoto 和 T. Arai，《绘画过程特征可区分阿尔茨海默病与路易体痴呆》，《阿尔茨海默病杂志》，第90卷，第693–704页，2022年09月。[28] A. Scholp、M. Hoffman、S. Rosen、M. Abdelhalim、C. Jones、J. Jiang 和 T. Mcculloch，《以谱弧长量化咽部高分辨率测压曲线平滑度的方法》，Neurogastroenterology & Motility，第33卷，2021年04月。[29] A. Barrett 和 C. Craver-Lemley，《是你看到的内容，还是你表达的方式？年轻和老年受试者的空间偏向》，《国际神经心理学会杂志：JINS》，第14卷，第562–70页，2008年08月。[30] D. Picard 和 B. Zarhbouch，《儿童绘画位置中的向左空间偏向：半球激活与方向性假说》，Laterality: Asymmetries of Body, Brain and Cognition，第19卷，第1期，第96–112页，2014年。[31] E. Apostolopoulou 和 A. Klonari，《学生在二维和三维地图上对河流的表征》，Procedia-Social and Behavioral Sciences，第19卷，第443–449页，2011年。[32] J. N. Buck，《房-树-人技术：修订版手册》。Western Psychological Services，1966年。[33] J. Buck 和 S. Sloan，《房-树-人技术修订版手册》。Ishi Press International，2019年。 [在线]。可访问：https: //books.google.com/books?id=zlZ0yAEACAAJ [34] J. Christie、M. Reichertz、B. Maycock 和 R. M. Klein，《擦除与否并非问题所在：在模拟或数字环境中进行观察绘画》，《高等教育中的艺术、设计与传播》，第19卷，第2期，第203–220页，2020年。[35] M. Leibowitz，《投射性绘画解读：自体心理学取向》。Routledge，2013年。[36] G. Groth-marnat，《心理评估手册》，第3版，2009年01月。[37] L. Handler 和 A. Thomas，《评估与心理治疗中的绘画：研究与应用》。Taylor & Francis，2013年。 [在线]。可访问：https://books.google.com.br/books?id=OtwkAgAAQBAJ [38] M. Cooper，《线描解读》，丛书：经济研究、不平等、社会排斥与福祉。Springer London，2010年。 [在线]。可访问：https://books.google.nl/books?id=pmAQxCUg12UC [39] N. Ali、A. Abd-Alrazaq、Z. Shah、M. Alajlani、T. Alam 和 M. Househ，《通过绘画感知儿童情绪的人工智能移动应用》，载于《医疗保健中的信息学、管理与技术进展》。IOS Press，2022年，第118–121页。[40] W. CAI、Y.-L. TANG、S. WU 和 Z.-Z. CHEN，《投射测验中的树》，《心理科学进展》，第20卷，第5期，第782页，2012年。[41] G. Chen 和 W. Yan，《罗夏墨迹测验在临床心理诊断中的效用》，China J Health Psychol，第30卷，80，2022年475pp. – 。[42] H. Zhou，《初中生反刍思维与房树人绘画特征关系的研究》，锦州：渤海大学，2021年。[43] J. Xiang、M. Liao 和 M. Zhu，《通过房-树-人测验评估小学低年级学生的抑郁倾向》，China J Health Psychol，第28卷，第1057–61页，2020年。[44] D. W. Joyce、A. Kormilitzin、K. A. Smith 和 A. Cipriani，《通过透明性和可解释性实现可理解的心理健康可解释人工智能》，npj Digital Medicine，第6卷，第1期，第6页，2023年。[45] S. Tutun、M. E. Johnson、A. Ahmed、A. Albizri、S. Irgil、I. Yesilkaya、E. N. Ucar、T. Sengun 和 A. Harfouche，《用于预测心理健康障碍的人工智能决策支持系统》，Information Systems Frontiers，第25卷，第3期，第1261–1276页，2023年。[46] M. O. Zeeshan、I. Siddiqi 和 M. Moetesum，《用于儿童绘画多标签分类的两阶段微调卷积神经网络》，载于《国际文档分析与识别会议》。Springer，2021年，第321–334页。[47] M. Kim、H. Sohn、S. Choi 和 S. Kim，《可信人工智能的要求及其在医疗保健中的应用》，Healthcare Informatics Research，第29卷，第315–322页，2023年10月。