PsyDraw: A Multi-Agent Multimodal System for Mental Health Detection in Left-Behind Children

Anonymous submission

Abstract

Left-behind children (LBC) face severe mental health challenges due to parental migration for work. The House-Tree-Person (HTP) test, a psychological assessment method with higher child participation and cooperation, requires expert interpretation, limiting its application in resource-scarce areas. To address this, we propose **PsyDraw**, a multi-agent system based on Multimodal Large Language Models for automated analysis of HTP drawings and assessment of LBC's mental health status. PsyDraw comprises two main stages, i.e., feature analysis and report generation, accomplished by multiple collaborative agents. We evaluate the system on HTP drawings from 290 primary school students, with the generated mental health reports evaluated by psychological teachers. Results show that 71.03% of the analyses are rated as **Matching**, 26.21% as **Generally Matching**, and only 2.41% as **Not Matching**. These findings demonstrate the potential of PsyDraw in automating HTP test analysis, offering an innovative solution to the shortage of professional personnel in mental health assessment for LBC.

Demo — psydraw.zeabur.app

Code — anonymous.4open.science/r/psydraw-D65B

Introduction

Left-behind children (LBC), resulting from parental economic migration, face severe mental health challenges (Wang et al. 2020; Zhu, Wang, and Pan 2023; Fauk et al. 2024). Studies show LBC exhibit 50-80% higher incidence of suicidal thoughts compared to non-LBC peers (Fellmeth et al. 2018; Račaitė et al. 2024), along with increased rates of anxiety and depression. As of 2020, there are approximately 66.93 million LBC in China (Si Zihan 2023). Early detection of mental health issues is crucial for LBC. While standardized tests are common, projective tests like the House-Tree-Person (HTP) test offer a more engaging and less tedious approach, increasing children's willingness to participate (Palmer et al. 2000). Despite its benefits, the HTP test lacks objective indicators and requires expert interpretation, challenging its use in areas with limited mental health resources. The emergence of Multimodal Large Language Models (MLLMs) presents a new opportunity to address this challenge. MLLMs have been widely applied in fields such as Visual Question Answering (VQA) (Liu et al. 2023; Bai et al. 2023) and healthcare (Li et al. 2024), opening up new possibilities for their application in mental health assessment. Interpreting HTP drawings and generating analysis reports is a complex task that requires professional knowledge in psychology and image understanding.

To address this challenge, we propose PsyDraw, a multi-agent system based on MLLMs. PsyDraw aims to evaluate mental health status and identify positive and negative factors by analyzing HTP drawings. The system's workflow is divided into two main stages: feature analysis and report generation. Each stgae is completed by multiple collaborating agents, with each agent guided by prompts crafted with professional knowledge to accomplish its specific tasks. To validate the effectiveness of PsyDraw, we analyze HTP drawings from 290 Chinese primary school students, generating corresponding mental health reports. These reports are subsequently evaluated and annotated by the subjects' psychological teachers. The results illustrate that 71.03% of the analyses are rated as **Matching**, 26.21% as **Generally Matching**, and only 2.41% as **Not Matching**. These preliminary findings demonstrate PsyDraw's potential in automating HTP test analysis, offering an innovative solution to the shortage of mental health professionals in assessing the psychological well-being of LBC.

Method

The House-Tree-Person (HTP) test (Burns 1987; Palmer et al. 2000) is a projective psychological assessment tool applicable to both children and adults aged 3 and above. This test aims to provide insights into an individual's personality, emotions, and attitudes through the analysis of drawings. In the HTP test procedure, participants are instructed to draw a house, a tree, and a person. Mental health professionals analyze HTP drawings to evaluate cognitive, emotional, and social functioning, interpreting depicted elements as reflections of hidden emotions, desires, and internal conflicts not easily discerned through direct methods. Drawing upon professional HTP analysis manuals (Burns 1987), we divide this task into two stages, HTP image feature analysis (***Stage 1***) and report generation (***Stage 2***). Each stage employs a multi-agent collaborative approach, resulting in the development of PsyDraw, an automated system for HTP drawing analysis and report generation.

Figure 1 illustrates the workflow of PsyDraw. In ***Stage***

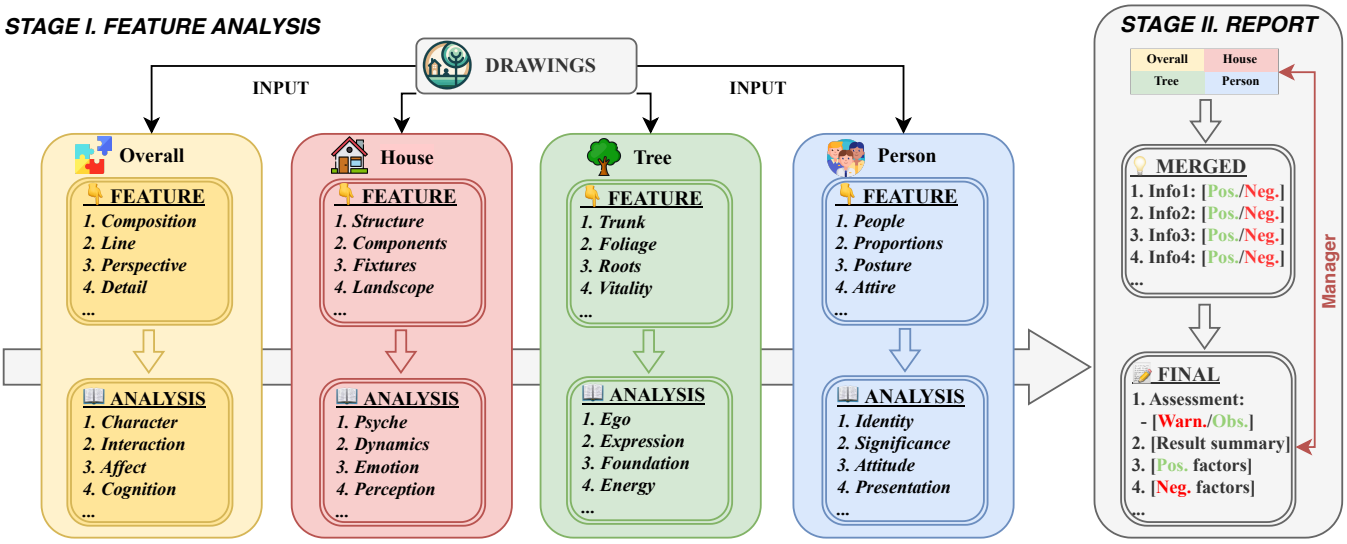

Figure 1: The workflow of PsyDraw.

1, we design four agents, including **Overall**, **House**, **Tree**, and **Person**, each focusing on a different aspect of HTP image interpretation. These agents follow a consistent strategy: given a drawing, they first extract features using specialized prompts and then analyze these features with professional prompts to generate preliminary analysis reports. The **Overall** Agent extracts features such as depth, size, and position, analyzing the subject's cognition, attitude, and emotional response to the environment. The **House** Agent extracts components like roofs and windows, analyzing family background and attitudes towards kinship relationships. The **Tree** Agent analyzes tree features, assessing subconscious self-image and psychological maturity. The **Person** Agent analyzes human figures, evaluating difficulties in self-image establishment and psychological defence mechanisms.

In **Stage II**, we focus on integrating and finalizing the analysis. Our agents consolidate the preliminary reports from the four feature analysis agents, summarize the initial results, and label each feature analysis with a tendency (positive, negative, neutral). Subsequently, they generate a comprehensive report that includes an assessment result (**Warning** or **Observation**), an analysis summary, and analyses of positive and negative factors. Throughout this process, the **Manager Agent** monitors for abnormal situations in HTP drawings. If it detects severe negative content (e.g., suicide, violence), it issues a warning and recommends professional medical assistance.

Experimental

This study is conducted in a Chinese primary school for left-behind children, accessed through a non-profit organization, collecting a total of **290** HTP drawings. The sample comprises 145 third-grade students and 145 fifth-grade students. During the drawing process, each participant is instructed on the basic HTP test rules: using a provided pen and paper to create a drawing containing a house, tree, and person, with no time limit. In our experiments, the PsyDraw system utilizes Claude-3.5-sonnet[1] as the text-only model and GPT-4o[2] as the multimodal model. We set the temperature parameter to 0.2 and Top P to 0.75 for all models while keeping other parameters at their default values.

	Total (%)	Warn. (%)	Obs. (%)
Matching	71.03	58.89	76.50
Generally Matching	26.21	26.21	22.00
Not Matching	2.41	4.44	1.50

Table 1: Matching rates of results with teacher feedback.

After analyzing each drawing using the PsyDraw system, psychological teachers review their students' comprehensive analysis reports, evaluating how well they align with their observations and knowledge. The evaluations are categorized as **Matching**, **Generally Matching**, or **Not Matching**. Table 1 presents the result of the teacher's feedback. Overall, the system achieves a 71.03% match rate with teacher feedback, indicating high effectiveness in assessing the psychological states of LBC. In the Warning group (Warn.), the match rate is 58.89%, with 35.56% generally matching and 4.44% not matching. The Observation group (Obs.) show a higher match rate of 76.50%, with 22.00% generally matching and 1.50% not matching.

Given that this system is designed for early psychological detection of LBC, the high match rate in the Observation group and the relatively lower match rate in the Warning group do not significantly impact its practical application. This result distribution aligns with the conservative principle of early screening. Nevertheless, these findings provide clear directions for further system optimization.

[1]https://www.anthropic.com/news/claude-3-5-sonnet
[2]https://openai.com/index/hello-gpt-4o/

References

Bai, J.; Bai, S.; Yang, S.; Wang, S.; Tan, S.; Wang, P.; Lin, J.; Zhou, C.; and Zhou, J. 2023. Qwen-VL: A Versatile Vision-Language Model for Understanding, Localization, Text Reading, and Beyond. arXiv:2308.12966.

Burns, R. C. 1987. *Kinetic-house-tree-person drawings (KHTP): An interpretative manual.* Brunner/Mazel.

Fauk, N. K.; Seran, A. L.; Aylward, P.; Mwanri, L.; and Ward, P. R. 2024. Parental Migration and the Social and Mental Well-Being Challenges among Indonesian Left-Behind Children: A Qualitative Study. *International Journal of Environmental Research and Public Health*, 21(6).

Fellmeth, G.; Rose-Clarke, K.; Zhao, C.; Busert, L. K.; Zheng, Y.; Massazza, A.; Sonmez, H.; Eder, B.; Blewitt, A.; Lertgrai, W.; Orcutt, M.; Ricci, K.; Mohamed-Ahmed, O.; Burns, R.; Knipe, D.; Hargreaves, S.; Hesketh, T.; Opondo, C.; and Devakumar, D. 2018. Health Impacts of Parental Migration on Left-behind Children and Adolescents: A Systematic Review and Meta-Analysis. *Lancet*, 392(10164): 2567–2582.

Li, C.; Wong, C.; Zhang, S.; Usuyama, N.; Liu, H.; Yang, J.; Naumann, T.; Poon, H.; and Gao, J. 2024. Llava-med: Training a large language-and-vision assistant for biomedicine in one day. *Advances in Neural Information Processing Systems*, 36.

Liu, H.; Li, C.; Wu, Q.; and Lee, Y. J. 2023. Visual Instruction Tuning. arXiv:2304.08485.

Palmer, L.; Farrar, A. R.; Valle, M.; Ghahary, N.; Panella, M.; and DeGraw, D. 2000. An Investigation of the Clinical Use of the House-Tree-Person Projective Drawings in the Psychological Evaluation of Child Sexual Abuse. *Child Maltreatment*, 5(2): 169–175.

Račaitė, J.; Antia, K.; Winkler, V.; Lesinskienė, S.; Sketerskienė, R.; Maceinaitė, R.; Tracevskytė, I.; Dambrauskaitė, E.; and Šurkienė, G. 2024. Emotional and Behavioural Problems of Left behind Children in Lithuania: A Comparative Analysis of Youth Self-Reports and Parent/Caregiver Reports Using ASEBA. *Child and Adolescent Psychiatry and Mental Health*, 18(1): 33.

Si Zihan, S. W. 2023. The Status of China's Migrant Children – China Development Brief. Technical report, New Citizen Program.

Wang, F.; Lin, L.; Lu, J.; Cai, J.; Xu, J.; and Zhou, X. 2020. Mental health and substance use in urban left-behind children in China: A growing problem. *Children and Youth Services Review*, 116: 105135.

Zhu, Z.; Wang, Y.; and Pan, X. 2023. Health problems faced by left-behind children in low/middle-income countries. *BMJ Global Health*, 8(8).