Automated Psychological Profiling from Children's Drawings Using Deep Learning and Natural Language Processing

Dhananjay Bhagat
School of Computer Engineering and Technology
MIT World Peace University
Pune, India
dhananjay.bhagat@mitwpu.edu.in

Kartik Maski
School of Computer Engineering and Technology
MIT World Peace University
Pune, India
kartikmaski29032004@gmail.com

Swamil Randive
School of Computer Engineering and Technology
MIT World Peace University
Pune, India
swamilrandive47@gmail.com

Harsh Dhimmar
School of Computer Engineering and Technology
MIT World Peace University
Pune, India
1032220410@mitwpu.edu.in

Ramadevi Vitthal Salunkhe
Department of Artificial Intelligence
G H Raisoni College of Engineering, Nagpur, India
ramadevisalunkhe@gmail.com

Dr. Devendra Joshi
School of Computer Engineering and Technology
MIT World Peace University
Pune, India
devendra.joshi1@mitwpu.edu.in

***Abstract*—Understanding the inner psychological world of children through their drawings has long been a cornerstone of developmental and clinical psychology. Traditional assessment techniques, such as the House-Tree-Person (HTP) test, rely heavily on expert interpretation, which is often subjective, time-consuming, and inconsistent across evaluators. In this research, we propose a novel AI-powered framework that fuses deep learning and natural language processing to automate the psychological interpretation of children's HTP drawings. The system employs ResNet-34, a convolutional neural network (CNN), for advanced image analysis to detect and interpret nuanced visual patterns in the drawings. These patterns are then transformed into coherent, human-like psychological insights using a transformer-based language generation model, specifically BART. Unlike static, template-driven systems, our approach dynamically generates rich, context-aware psychological narratives that align with clinical reasoning. This interdisciplinary methodology presents a significant step forward in augmenting traditional psychological evaluation with scalable, interpretable, and data-driven intelligence.**

Keywords*—*House-Tree-Person (HTP) test, AI-powered framework, ResNet-34, BART, Psychological insights

I. Introduction

The psychological assessment of children plays a crucial role in early diagnosis, intervention, and support planning for emotional and behavioral issues. Children, especially at a young age, often lack the verbal ability to express their internal emotional states and experiences [1]. In such cases, projective techniques like drawing-based assessments offer a non-verbal medium to uncover their psychological condition. Among these, the **House-Tree-Person (HTP) test**, developed by John N. Buck in 1948 [2], stands out as a widely used projective tool for understanding a child's personality, emotional functioning, and cognitive development. Through the symbolic interpretation of drawings of a house, tree, and person, psychologists have long inferred key aspects of a child's psychological state, including signs of stress, anxiety, introversion, depression, and emotional trauma.

Traditionally, the interpretation of HTP drawings has relied heavily on the subjective analysis by clinical psychologists, who evaluate aspects such as the size, placement, and completeness of the drawings, as well as the emotional tone, pressure, and detail [3]. While effective, this method is inherently dependent on the clinician's expertise and may suffer from inter-rater variability and a lack of quantifiable metrics. Moreover, the manual evaluation process is time-consuming and may not scale efficiently in institutional or educational settings where large-scale psychological screening may be required.

With advancements in technology, researchers are increasingly turning to automated and data-driven methods to support psychological assessment. Artificial Intelligence (AI) and Machine Learning (ML) have opened new possibilities for analyzing children's drawings, offering tools that can assist in identifying emotional and behavioral patterns [4]. These systems are capable of learning from large sets of drawing samples and detecting visual cues that may be associated with psychological conditions. In particular, image analysis techniques can be used to identify common traits across drawings, such as indicators of stress or withdrawal. In parallel, language-based AI systems are being used to interpret and communicate these findings in a more accessible way [5]. This integration of visual and linguistic AI tools allows for consistent and scalable

979-8-3315-1302-3/25/$31.00 ©2025 IEEE

psychological evaluation, supporting educators, psychologists, and caregivers in better understanding children's mental and emotional well-being.

In the present study, we propose an innovative framework that integrates computer vision and natural language processing to facilitate the psychological interpretation of children's drawings, with a particular emphasis on the House-Tree-Person (HTP) technique. The proposed approach utilizes deep learning methods to analyze visual elements within children's sketches, enabling the automated classification of psychological indicators such as stress, introversion, and extroversion etc. Complementing this, a natural language generation component is employed to produce structured, multi-sentence psychological assessments that correspond to the inferred emotional and cognitive states. By combining these technologies, the framework seeks to overcome the limitations of traditional manual interpretation, offering a scalable, consistent, and interpretable solution for psychological evaluation. The technical architecture, model design, and implementation details are elaborated in the subsequent sections.

II. Background and Preliminary Study

Psychological assessments have long utilized projective techniques to explore individuals' subconscious thoughts and emotional states, with drawing-based tests offering valuable insights, particularly in children. The House-Tree-Person (HTP) test [2], widely used in clinical psychology, involves drawing a house, tree, and person to reflect different psychological aspects, such as security, growth, and self-perception. However, traditional interpretations are often subjective and lack standardization. With advancements in Artificial Intelligence (AI) and Machine Learning (ML), the analysis of HTP drawings is being transformed. AI models, including convolutional neural networks (CNNs) and transformer-based Natural Language Processing (NLP), are now capable of automating emotional trait classification and generating consistent, objective psychological reports, enhancing the reliability and scalability of this tool in modern psychological practice.

A. The House-Tree-Person (HTP) Test

The House-Tree-Person (HTP) test is a projective drawing-based assessment widely used to gain insights into an individual’s psychological and emotional state. It involves instructing individuals—particularly children—to draw a house, a tree, and a person, each of which symbolically represents distinct aspects of the human psyche [3]. These elements are deliberately chosen for their universal familiarity and accessibility across different developmental stages.

Sample Drawing	Drawing Type	Predicted Trait
	House	Stress, Anxiety
	Tree	Introvert
	Person	Boredom, Lack of Motivation

Table 1: Samples of House, Tree, and Person drawings collected from the research dataset.

In the context of psychological interpretation, the **house** reflects perceptions of family life, safety, and internal comfort; features such as size, structure, and detailing offer cues about the individual's sense of security and attachment. The **tree** is associated with personal growth, stability, and self-identity, often analyzed through trunk strength, branches, and overall form. The **person** drawing serves as a projection of self-concept and interpersonal orientation, where characteristics like body proportion, posture, and expression provide insights into social confidence and self-image.

Children are typically asked to draw these figures freely, without specific constraints, allowing them to project subconscious feelings through their sketches. Following the drawings, descriptive responses about each figure are sometimes collected to further contextualize their emotional and cognitive associations. This format makes the HTP test particularly effective for use in automated visual analysis systems, as the structural attributes in the drawings provide consistent visual markers that can be learned and interpreted through machine learning models, as illustrated in **Table 1**, which presents sample House, Tree, and Person drawings from the dataset.

979-8-3315-1302-3/25/$31.00 ©2025 IEEE

B. Traditional Administration of the HTP Test and Its Limitations

The traditional administration of the HTP test has long been criticized for its reliance on subjective interpretation and lack of empirical grounding. One of the primary limitations stems from the examiner's personal biases, which can inadvertently influence the interpretation of a child's drawings. Without standardized scoring criteria, different psychologists may derive conflicting conclusions from identical drawings, leading to potential inconsistencies in psychological assessment. Furthermore, the variability in children's artistic abilities and cultural backgrounds adds another layer of complexity; a drawing interpreted as indicative of psychological distress in one context may simply reflect developmental or stylistic differences in another [6].

Additionally, traditional HTP assessments often fail to differentiate between symbolic and literal expressions in drawings, increasing the chances of over-interpretation or false positives. In clinical settings, especially in areas such as child abuse investigations, the misuse of HTP interpretations has raised serious concerns regarding the ethical and diagnostic reliability of such projective tools. Despite attempts to introduce structured scoring systems, studies continue to report limited validity and reliability in correlating specific drawing features with psychological conditions. These challenges underscore the need for more objective, reproducible, and data-driven approaches to enhance the credibility and utility of the HTP test in modern psychological evaluation [7].

C. Integration of AI and Machine Learning in HTP Analysis

The integration of Artificial Intelligence (AI) and Machine Learning (ML) techniques into the analysis of House-Tree-Person (HTP) drawings represents a significant advancement in psychological assessment methodologies. Traditional interpretations of HTP tests are often subjective and labor-intensive, prompting researchers to explore automated approaches that can provide objective, efficient, and scalable analyses.

Deep Learning, particularly Convolutional Neural Networks (CNNs), has been instrumental in this transformation. CNNs are adept at processing visual data, making them suitable for interpreting the complex features present in children's drawings. For instance, a study employed CNNs to classify HTP drawings, focusing on reducing the subjectivity inherent in traditional evaluations. The researchers highlighted that CNNs could automatically learn and extract relevant features from the drawings, thereby enhancing the reliability of psychological assessments.

In addition to image analysis, Natural Language Processing (NLP) has been utilized to interpret textual data associated with psychological assessments. NLP techniques can analyze language patterns to detect indicators of mental health conditions. A comprehensive review of NLP applications in mental illness detection revealed an upward trend in research focusing on this intersection, emphasizing the potential of NLP to empower proactive mental healthcare and assist in early diagnosis.

Combining these AI methodologies, systems like **PsyDraw** have been developed to assist mental health professionals in analyzing HTP drawings [8]. PsyDraw employs a multi-agent framework based on Multimodal Large Language Models (MLLMs) to perform comprehensive feature analysis and generate professional psychological reports. Evaluation of HTP drawings from primary school students demonstrated that a significant portion of the analyses achieved high consistency with professional evaluations, indicating the system's effectiveness as a preliminary screening tool.

Despite these advancements, challenges remain. The subjective nature of drawing interpretation, cultural differences, and individual artistic abilities can influence outcomes. Moreover, while AI models can process visual and textual data efficiently, they may lack the nuanced understanding that human clinicians provide. Therefore, it is crucial to develop AI systems that complement human expertise, ensuring that automated analyses serve as supportive tools rather than replacements for professional judgment [9].

In summary, the integration of AI and ML techniques into the analysis of HTP drawings holds significant promise for enhancing the objectivity, efficiency, and scalability of psychological assessments. Ongoing research and development are essential to address existing challenges and to refine these technologies to better support mental health professionals in their evaluative processes.

III. Proposed Methodology

This study employs a curated dataset of children's house drawings to infer psychological traits based on the visual characteristics present in each image. The drawing samples were sourced from the **Google Quick, Draw! repository** [10] [11], a publicly available dataset consisting of millions of sketch-style drawings contributed by users worldwide. From this repository, we filtered and selected relevant house, tree, and person images and further refined them based on clarity and structural completeness. Each drawing is associated with one or more psychological outcomes—such as stress or anxiety—derived from established psychological interpretations. The dataset was manually annotated according to a predefined schema of visual indicators, such as excessive smoke from chimneys, intricate architectural detailing (e.g., inclusion of windows and doors), and shaded areas, which prior psychological studies have linked to specific emotional states. These labels were compiled and organized in structured CSV files corresponding to each drawing category. The annotation process was carried out by trained reviewers following consistent guidelines to ensure reliability and reduce subjectivity in label assignment. Since the original dataset is open-source and non-identifiable, ethical concerns related to the use of children's private data were not applicable;

however, the annotation framework was designed to simulate psychological interpretation of child-like drawings for research purposes only.

The classification framework relies on these visual indicators, which are learned by the model through exposure to labeled examples during training. As illustrated in Fig 1, the presence or absence of such features is used to infer the psychological state of the child. This feature-to-trait mapping forms the conceptual foundation for subsequent automated classification of the drawings.

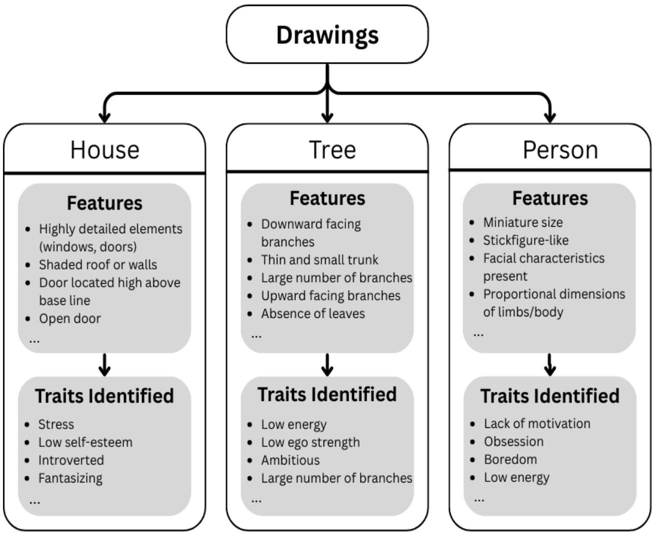

Fig 1: Analysis of image classification based on absence/presence of features

To systematically analyze children's drawings and generate interpretable psychological assessments, an integrated multi-stage framework is proposed, encompassing convolutional neural network-based image classification and transformer-based natural language processing (NLP) for report generation. The approach aims to map visual patterns present in house, tree, and person drawings to corresponding psychological traits, followed by the synthesis of descriptive textual feedback grounded in clinical interpretation.

The initial phase of the pipeline employs the ResNet-34 architecture, a deep convolutional neural network that leverages residual learning via skip connections to facilitate gradient flow and improve learning stability in deeper networks [12]. ResNet-34 is selected for its demonstrated efficacy in complex image recognition tasks, making it well-suited for interpreting variable and often abstract drawing inputs characteristic of children's illustrations. The model is trained independently on three distinct datasets—corresponding to house, tree, and person drawings—each annotated with psychological traits based on established feature-based psychological heuristics.

All input images are subjected to a standardized preprocessing protocol, including resizing and normalization, to ensure consistency in dimensionality and pixel distribution across the training set. This preprocessing is critical to reducing noise and enhancing the model's capacity to generalize.

To further enhance classification robustness, ensemble learning techniques are introduced. Multiple ResNet-34 models are trained on the same dataset using varying hyperparameters, including epoch counts, learning rates, batch sizes, and optimizer configurations. The resultant ensemble of models promotes diversity in learned representations, reducing the risk of overfitting to specific data patterns. Final predictions are derived via **majority voting**, allowing the ensemble to converge on a consensus prediction that reflects the most probable psychological classification per image. This ensemble strategy has been shown to improve classification stability and mitigate the limitations of relying on single-model predictions [13].

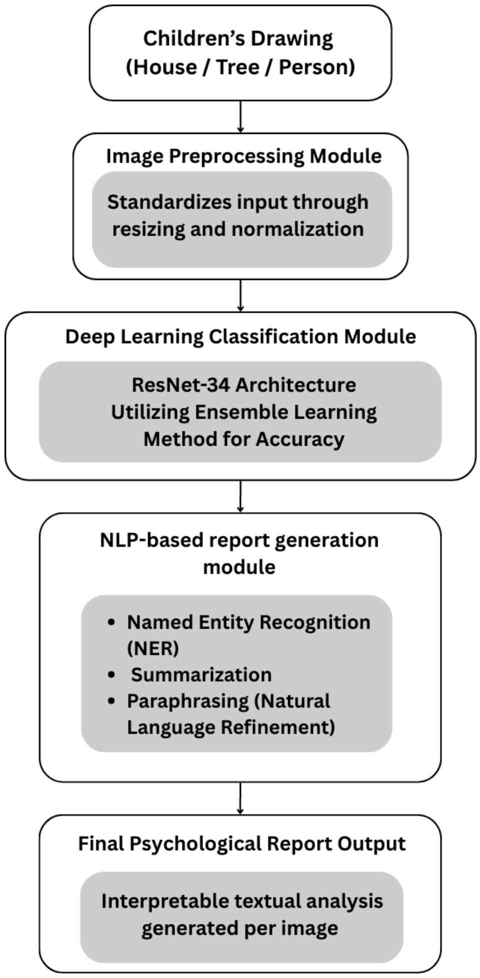

Fig. 2: End-to-End Flow of Classification and Report Generation

Upon obtaining the classification outputs, the system transitions into a natural language generation phase, wherein the structured predictions produced by the image classification models are transformed into coherent, human-readable psychological reports. This transformation is achieved through a sequenced integration of transformer-based natural language processing (NLP) models. The first stage employs Named Entity Recognition (NER) to extract key psychological constructs, traits, and emotional indicators from the classification labels. Subsequently, an abstractive summarization model, specifically BART, is utilized to synthesize these extracted elements into concise, clinically relevant summaries that convey the underlying psychological implications. To enhance the readability and interpretability of the output, a paraphrasing module based on T5-small is employed to rephrase the summarized content into fluent and contextually accurate language. These NLP modules work in unison to simulate a multi-stage cognitive synthesis pipeline, in which raw classification results are semantically enriched and linguistically refined. The final output is a structured and interpretable psychological report that contextualizes each child's drawing within a clinical framework, effectively bridging model inference with human-centered communication.

The complete flow—from visual feature extraction and classification to natural language generation—is visualized in **Fig 2**, which illustrates the modular architecture and the sequential relationship between deep learning and NLP components. This end-to-end system enables automated yet interpretable psychological analysis, facilitating its application in clinical, educational, and developmental psychology contexts.

IV. Results and Analysis

To evaluate the effectiveness of the proposed classification system, multiple ResNet-34 models were independently trained on each of the three drawing categories—House, Tree, and Person. For each category, three independent instances of the ResNet-34 model were trained with different epoch configurations. The predictive results of these individual models are reported and compared against the ensemble learning approach, which integrates the outputs through majority voting. Performance is primarily assessed using classification accuracy, and the results are summarized in Table 2.

Models	**Number of Epochs**	**Accuracy**
House	10	94.6 %
	12	97.3 %
	15	98.1 %
Tree	10	93.5 %
	12	95.8 %
	15	97.0 %
Person	10	94.3 %
	12	98.4 %
	15	99.1 %

Table 2: Accuracy of Individual ResNet-34 Models Trained on Different Epochs for House, Tree, and Person Drawing Classifications

Prior to adopting ResNet-34, we also experimented with a basic Convolutional Neural Network (CNN) architecture. However, the CNN model demonstrated suboptimal performance and inefficiency. It required approximately **70 epochs** to reach a maximum accuracy of only **93.4%**. This was notably less efficient and less accurate than the ResNet-34 models, which achieved higher performance in significantly fewer epochs. Due to this disparity, the CNN approach was not selected for deployment in this study but is documented here for comparison purposes.

In addition to classification accuracy, the report generation module was assessed based on its ability to translate prediction outputs into coherent psychological narratives. The NLP pipeline—comprising Named Entity Recognition (NER), BART-based summarization, and T5-based paraphrasing—produced consistent and interpretable reports aligned with classification results.

Fig. 3 demonstrates a complete case example of the system in action. It showcases the classification and report generation for a Tree drawing, detailing both the predicted psychological trait and the corresponding natural language interpretation. This visual example affirms the system's end-to-end functionality and highlights the accuracy of both classification and linguistic generation components in capturing the psychological indicators present in children's drawings.

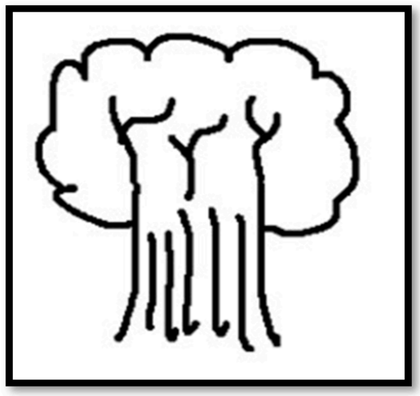

Fig. 3. Illustration of Drawing Classification and Psychological Report Generation

V. Conclusion and Future Scope

This study presents a unified, multi-stage framework that integrates deep learning-based image classification with transformer-driven natural language processing for the psychological interpretation of children's drawings. Utilizing a ResNet-34 architecture, the system classifies house, tree, and person sketches into psychological traits such as stress, anxiety, and introversion, with enhanced accuracy achieved through ensemble learning strategies. These classification outcomes are then processed by a sequence of NLP modules—Named Entity Recognition, summarization, and paraphrasing—to generate clinically meaningful and human-readable psychological reports. The methodological pipeline not only improves prediction reliability but also bridges the gap between automated inference and psychological interpretability, making it a valuable tool for early behavioral assessment.

The significance of this work lies in its potential application in educational and clinical contexts, where early detection of emotional distress can lead to timely interventions. Future research will focus on expanding the dataset with more diverse samples to improve generalization, incorporating attention-based visualization techniques to enhance model explainability, and domain-specific fine-tuning of NLP components to increase report precision. The proposed system establishes a foundation for scalable, interpretable psychological screening, and opens new avenues for the use of AI in child mental health evaluation, particularly in environments with limited access to professional expertise.

VI. References

[1] Andronikof, Anne & Fontan, Patrick. (2014). Psychological Assessment of Children: Guidelines and Ethics. Neuropsychiatrie de l Enfance et de l Adolescence. 62. 403-407.

[2] Gordon, R.M., Rudd-Barnard, A. (2011). House-Tree-Person Test. In: Kreutzer, J.S., DeLuca, J., Caplan, B. (eds) Encyclopedia of Clinical Neuropsychology. Springer, New York, NY. https://doi.org/10.1007/978-0-387-79948-3_2029

[3] Guo H, Feng B, Ma Y, Zhang X, Fan H, Dong Z, Chen T, Gong Q. Analysis of the screening and predicting characteristics of the house-tree-person drawing test for mental disorders: A systematic review and meta-analysis. Front Psychiatry. 2023 Jan 4;13:1041770. doi: 10.3389/fpsyt.2022.1041770. PMID: 36683989; PMCID: PMC9848786.

[4] Moonyoung Lee, Youngho Kim, Young-Kuk Kim, Generating psychological analysis tables for children's drawings using deep learning, Data & Knowledge Engineering, Volume 149, 2024, 102266, ISSN 0169-023X, https://doi.org/10.1016/j.datak.2023.102266.

[5] Alshahrani, A., Almatrafi, M. M., Mustafa, J. I., Albaqami, L. S., & Aljabri, R. A. (2024). A Children's Psychological and Mental Health Detection Model by Drawing Analysis based on Computer Vision and Deep Learning. *Engineering Technology & Applied Science Research*, *14*(4), 15533–15540. https://doi.org/10.48084/etasr.7812

[6] Y. Lin and N. Zhang, "The House-Tree-Person test is not valid for the prediction of mental health: An empirical study using deep neural networks," *Acta Psychologica*, vol. 230, p. 103734, Oct. 2022. [Online]. Available: https://doi.org/10.1016/j.actpsy.2022.103734

[7] Eyal, C., & Lindgren, H. C. (1977). The House-Tree-Person Test as a Measure of Intelligence and Creativity. Perceptual and Motor Skills, 44(2), 359-362. https://doi.org/10.2466/pms.1977.44.2.359 (Original work published 1977)

[8] Y. Zhang et al., "PsyDraw: A Multi-Agent Multimodal System for Mental Health Screening in Left-Behind Children," *arXiv preprint arXiv:2412.14769*, Dec. 2024. [Online]. Available: https://arxiv.org/abs/2412.14769.

[9] Zhang J, Yu Y, Barra V, Ruan X, Chen Y, Cai B. Feasibility study on using house-tree-person drawings for automatic analysis of depression. Comput Methods Biomech Biomed Engin. 2024 Jul;27(9):1129-1140. doi: 10.1080/10255842.2023.2231113. Epub 2023 Jul 7. PMID: 37417817.

[10] Guo, K., Ma, J., & Xu, E. (2018). *Quick, Draw! Doodle Recognition*. Stanford University. Available: https://cs229.stanford.edu/proj2018/report/98.pdf

[11] van Doorn, F. (2016). Children as co-researchers in design: Enabling users to gather, share and enrich contextual data. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:f16db80d-9f1a-4064-91be-decf8c805898

[12] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

[13] Ammar Mohammed, Rania Kora, A comprehensive review on ensemble deep learning: Opportunities and challenges, Journal of King Saud University - Computer and Information Sciences, Volume 35, Issue 2, 2023, Pages 757-774, ISSN 1319-1578, https://doi.org/10.1016/j.jksuci.2023.01.014.

979-8-3315-1302-3/25/$31.00 ©2025 IEEE