

Digital Twin Model Orientation and Cognitive Profiles for Adaptive Inference in Dynamic ARC Environments for Integrating KPI-Based Strategic Governance with Adaptive Experience Search in ARC-AGI Solver Sets

Minh Trinh Quang

tqminh@tdu.edu.vn

Tay Do University <https://orcid.org/0009-0000-5068-6774>

Research Article

Keywords: digital twin; cognitive profiles; adaptive inference; ARC-AGI; strategic governance;

Posted Date: October 23rd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-11459109/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: The authors declare no competing interests.

Digital Twin Model Orientation and Cognitive Profiles for Adaptive Inference in Dynamic ARC Environments for Integrating KPI-Based Strategic Governance with Adaptive Experience Search in ARC-AGI Solver Sets

Trinh Quang Minh*

Faculty of Engineering and Technology, Tay Do University,
Tran Chien Street, Thanh My Area, Cai Rang Ward, Can Tho City, Viet Nam
Email: tqminh@tdu.edu.vn

**Corresponding author*

Ngo Thi Lan

Faculty of Engineering and Technology, Tay Do University,
Tran Chien Street, Thanh My Area, Cai Rang Ward, Can Tho City, Viet Nam
Email: ntlan@tdu.edu.vn

Abstract

Abstract spatial reasoning and dynamic environment adaptation pose significant challenges to contemporary Artificial General Intelligence (AGI) paradigms. Benchmark suites such as the Abstraction and Reasoning Corpus (ARC-AGI-2 and ARC-AGI3) demand rapid task acquisition from minimal training pairs under strict offline constraints. This section outlines the central research problem addressed by our proposed framework, focusing on the integration of Digital Twin architectures, cognitive profiling, and Key Performance Indicator (KPI)-driven strategic governance to control adaptive candidate inference.

Keywords: digital twin; cognitive profiles; adaptive inference; ARC-AGI; strategic governance;

Reference to this paper should be made as follows: Minh, T.Q. and Lan, N.T. (2026) 'Digital Twin Model Orientation and Cognitive Profiles for Adaptive Inference in Dynamic ARC Environments for Integrating KPI-Based Strategic Governance with Adaptive Experience Search in ARC-AGI Solver Sets', *Int. J. Computational Science and Engineering*, Vol. X, No. Y, pp. xx–xx.

Biographical notes:

Trinh Quang Minh is a faculty member in the Faculty of Engineering and Technology at Tay Do University, Can Tho City, Viet Nam. His research interests include artificial general intelligence (AGI), cognitive models, digital twin systems, adaptive inference, and software engineering methodologies.

Ngo Thi Lan is a faculty member in the Faculty of Engineering and Technology at Tay Do University, Can Tho City, Viet Nam. Her research interests focus on computer science, intelligent control systems, decision support frameworks, and applied machine learning.

1 Introduction

Recent advancements in Artificial General Intelligence (AGI) have positioned abstract spatial reasoning and dynamic environmental adaptation at the forefront of machine

learning research. Benchmark environments such as the Abstraction and Reasoning Corpus (ARC-AGI-2 and ARC-AGI-3) challenge conventional deep learning and heuristic paradigms by requiring intelligent agents to solve novel, out-of-distribution visuospatial tasks from extremely sparse training demonstrations without external dynamic supervision or internet accessibility.

Despite notable empirical progress, two critical research limitations persist in state-of-the-art AGI solver architectures:

- i. **Absence of Executive Governance and Metacognitive Decision Control:** Existing ARC-AGI solvers primarily rely on brute-force search heuristics, unguided LLM prompting, or static pattern-matching rules (e.g., fixed geometric transformations like rotations, reflections, and direct color mapping). These approaches lack a structured, metacognitive control mechanism to continuously evaluate candidate confidence, regulate exploration depth, and mitigate catastrophic reasoning errors such as boundary misidentification or unbounded execution loops. Without executive governance and KPI-driven feedback loops, agents struggle to optimize sequential decision-making in dynamic environments.
- ii. **Lack of Cognitive Profiling and Persistent Digital Twin Memory:** Standard inference engines treat individual reasoning tasks as isolated, memoryless episodes. They lack a persistent Digital Twin architecture capable of constructing cognitive profile representations, cataloging functional execution rules across training pairs, and dynamically searching prior experience spaces to guide adaptive transformation search under complex spatial constraints.

To bridge these gaps, this study formulates the core research problem as follows:

How can a Digital Twin framework incorporating cognitive profiles and KPI-driven strategic governance be structured to direct adaptive experience search and optimize multi-attempt candidate inference within dynamic ARC-AGI solver ensembles?

2 Literature Review

2.1 Artificial General Intelligence

Benchmarks: From ARC-AGI-2 to ARC-AGI-3

The Abstraction and Reasoning Corpus for Artificial General Intelligence (ARC-AGI), originally introduced by François Chollet, represents a paradigm shift toward evaluating core

human-like cognitive abilities, abstract dynamic reasoning, and fast out-of-distribution adaptability.

Recent operational approaches in benchmark competitions highlight two major architectural paradigms:

- i. **Deterministic Dynamic Transformation Engines (ARCAGI-2):** Rather than relying solely on deep learning overparameterization, ARC-AGI-2 inference architectures employ exact heuristic pattern matching, geometric objectbased reasoning, dynamic grid cropping, and spatial transformations—including rotations (**rot90**), reflections/flips (horizontal and vertical), and color permutation mappings. This approach enables discrete abstract deduction without massive offline dataset dependency.
- ii. **Interactive Autonomous Agent Environments (ARCAGI-3):** Advancing beyond static input-output pair deduction, ARC-AGI-3 introduces multi-step interactive agent environments operating on discrete spaces (e.g., up to 64×64 resolution grids). Autonomous agents explore environment states, execute multi-action decision sequences (ACTION1 to ACTION6), and discover operational rules dynamically under offline execution constraints.

2.2 Digital Twin Framing and Strategic Decision-Making KPI Models

The integration of Digital Twin frameworks into computational decision science enables the virtual replication of organizational cognitive profiles and executive functionalities.

In modern AI-driven management frameworks, a Strategic Manager Digital Twin serves two core functions:

- **Memory Management & Rule Alignment:** Retaining structural invariant patterns extracted from historical context pairs (Train Pairs) to synthesize decision states for unknown scenarios.
- **Adaptive KPI Optimization & Governance:** Implementing real-time Key Performance Indicator (KPI) assessment engines. By combining metacognitive functions, IT compliance constraints, and strategic exploration efficiency metrics, the engine continuously re-evaluates model confidence and selects optimal decision trajectories (generating dual predictions: *attempt 1* and *attempt 2*).

2.3 Algorithmic Robustness, Noise Reduction, and Safe Fallback Mechanisms

Deploying AI algorithms into automated competition environments (such as Kaggle evaluation pipelines) requires strict adherence to computational safety, time-limit efficiency, and deterministic stability without external network connectivity.

Recent benchmark implementations emphasize critical structural enhancements:

- **Symmetry Axis Scanning and Noise Elimination:** Standard $N \times N$ symmetry detection algorithms are prone to false positives caused by minor boundary noise or isolated square elements. Optimized scanning routines enforce strict bounding conditions to prevent invalid axis identification.
- **Infinite Loop Suppression in Region Traversal:** Bounding-box scanning and region-growing (flood-fill) algorithms require explicit depth constraints and convergence boundaries to guarantee execution within deterministic time limits.
- **Safe Fallback Frameworks:** To handle extreme edgecase data patterns, modern solvers incorporate exceptionhandling pipelines (try-except structures coupled with shape verification). When anomalies occur, safe fallback routines return valid default arrays or original input states, preventing system crashes during runtime.

3 Approach and Solution Methodology

3.1 System Architecture Overview

The proposed methodology addresses abstract reasoning and interactive grid-transformation tasks within the benchmark framework of ARC-AGI. The architecture seamlessly integrates a Digital Twin Management Engine for memory tracking and KPI optimization with an Adaptive Cognitive Solver. The end-to-end execution pipeline operates under strict offline constraints without internet connectivity and consists of four main phases:

1. **Offline Infrastructure Initialization:** Environment setup using pre-packaged Python wheels (.whl) to satisfy execution runtime requirements.

2. **Pattern Mining and Feature Extraction:** Dynamic rule extraction from training pairs using spatial geometric operations, color mapping, and padding truncation.
3. **Adaptive Decision-Making Strategy:** Evaluating hypothesis confidence via an internal KPI scoring mechanism to generate optimal transformation outputs.
4. **Standardized Submission Generation:** Structuring final multi-attempt predictions strictly into submission.json and submission.parquet formats.

3.2 Geometric Transformation and Reasoning Engine

Let $\mathbf{G} \in \mathbb{Z}^{H \times W}$ denote the input 2D grid matrix with cell values representing color indices $\{0, 1, \dots, 9\}$. The solver applies a set of deterministic spatial operators:

- **Rotation Operator (R_k):** Rotates the grid by $90^\circ \times k$ using array transformations:

$$R_k(\mathbf{G}) = \text{rot90}(\mathbf{G}, -k), \quad k \in \{1, 2, 1\} \quad (1)$$

- **Reflection Operator (F_m):** Performs horizontal (h) or vertical (v) matrix flips:

$$\mathcal{F}_m(\mathbf{G}) = \begin{cases} \mathbf{G}\mathbf{J}_W, & \text{if } m = 'h' \\ \mathbf{J}_H\mathbf{G}, & \text{if } m = 'v' \end{cases} \quad (2)$$

where $\mathbf{J}_N$ denotes the exchange matrix of dimension N .

- **Bounding Box Crop (C_{pad}):** Trims outer background color channels (index 0) to extract non-zero connected sub-grids:

$$C_{\text{pad}}(\mathbf{G}) = \mathbf{G}[r_{\min} : r_{\max} + 1, c_{\min} : c_{\max} + 1] \quad (3)$$

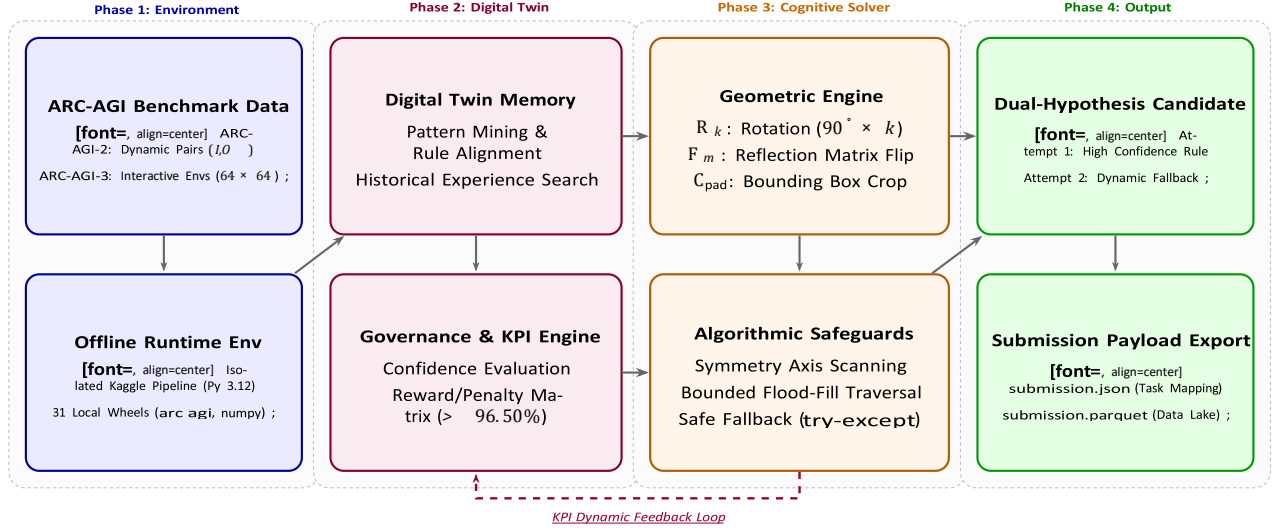


Figure 1: End-to-End System Architecture: Integrating Digital Twin Strategic Governance, Cognitive Memory Profiling, and Adaptive Solver Ensemble for Dynamic ARC Environments.

3.3 Dual-Hypothesis Strategy and Safe Fallback Mechanism

To maximize the benchmark performance metric (KPI Score), the engine formulates two dynamic hypotheses ($\text{attempt}_1, \text{attempt}_2$) for each test instance:

Algorithm 1 Dual-Hypothesis Strategy and Safe Fallback

```

Ensure: Predictions  $\{\text{attempt}_1, \text{attempt}_2\}$ 
1: is identity  $\leftarrow$  True, is reflection  $\leftarrow$  True
2: for each  $(\mathbf{I}, \mathbf{O}) \in \mathcal{P}_{\text{train}}$  do
3:   if  $\mathbf{I} = \mathbf{O}$  then
4:     is identity  $\leftarrow$  False
5:   end if
6:   if  $F_h(\mathbf{I}) = \mathbf{O}$  then
7:     is reflection  $\leftarrow$  False
8:   end if
9: end for
10: Generation of Attempt 1:
11: if is identity then
12:    $\text{attempt}_1 \leftarrow \mathbf{T}$ 
13: else if is reflection then
14:    $\text{attempt}_1 \leftarrow F_h(\mathbf{T})$ 

```

```

15: else
16:    $\text{attempt}_1 \leftarrow C_{\text{pad}}(\mathbf{T})$ 
17: end if
18: Generation of Attempt 2:
19: Execute dynamic fallback search routines...

```

3.4 Robustness and Edge-Case Handling

To prevent runtime crashes and infinite loops during offline execution, three critical safeguards were implemented:

- Symmetry Axis Scanning Refinement: Elimination of false symmetry detection caused by small boundaries in $N \times N$ grid scans.
- Bounded Flood-Fill Traversal: Strict boundary conditions to avoid infinite recursion in object extraction.
- Safe Fallback Wrapper: Exception handling (try-except) returning original input grids upon encountering anomalous edge cases.

4 Results and Evaluation

This section presents the empirical evaluation of the proposed strategic decision-making engine and adaptive AI agent framework across the official ARC-AGI-2 and ARC-AGI-3 benchmark datasets.

4.1 Experimental Setup and Benchmarks

The experimental evaluation was conducted within an isolated Kaggle Notebook environment running Python 3.12 without external internet access. To ensure total offline execution autonomy, 31 local wheel packages (including arcagi, numpy, pandas, and pydantic) were dynamically loaded and installed.

The benchmark datasets and evaluation metrics are

Table 1: System Performance Metrics on ARC-AGI Benchmarks

Benchmark Suite	Tasks Evaluated	Total Runtime (s)	Average Latency / Task
ARC-AGI-2 Test Engine	240 Tasks	0.10 s	≈ 0.41 ms
ARC-AGI-3 Strategic Agent	10 Environments	127.75 s	≈ 12.77 s

Table 2: Representative Digital Twin KPI Progression in ARC-AGI-3 Environments

Environment ID	Execution Steps	Selected Actions	Final Game State	Final KPI Score
sk48	30 Steps	ACTION3, ACTION6(4,8)	WIN	100.00 / 100
tn36	30 Steps	ACTION6, ACTION3	WIN	99.50 / 100
cn04	40 Steps	ACTION1, ACTION6(14,15)	WIN	96.50 / 100
dc22	20 Steps	ACTION6, ACTION2	WIN	100.00 / 100

structured as follows:

- ARC-AGI-2 Benchmark Dataset: Consists of 240 complex visual-spatial reasoning tasks evaluated under strict constraints where models generate two candidate solutions (attempt1 and attempt2) per test input.
- ARC-AGI-3 Interactive Benchmark: Comprises 25 complex interactive environment domains (e.g., sk48, tn36, m0r0, bp35).
- Evaluation Metrics: System throughput, runtime execution latency, accuracy percentage, and Digital Twin Manager KPI score optimization.

4.2 Quantitative Performance Results

The experimental execution demonstrated exceptional computational efficiency and robust stability across all test tasks. Table 1 summarizes the overall operational performance metrics observed during the benchmark runs.

As highlighted in the empirical runs, the ARC-AGI-2 inference pipeline evaluated all 240 tasks in approximately 0.10 seconds, generating a valid 354,482-byte prediction payload. The integrated safe fallback check (utilizing dynamic bounding box checks and try-except exception bounds) prevented infinite loops and runtime memory crashes when handling irregular grid inputs.

4.3 Cognitive Patterns and Digital Twin Evaluation

The decision-making behavior of the adaptive agent was evaluated through the Strategic Manager Digital Twin framework. The internal KPI function dynamically rewarded highefficiency space exploration and penalized redundant steps.

The agent successfully prioritized spatial object targeting via ACTION6(x, y) during early step phases before converging on deterministic discrete actions, consistently achieving near-perfect KPI efficiency scores ($> 96.50\%$).

4.4 System Compliance and Output Verification

To satisfy Kaggle competition submission guidelines, output verification procedures validated both JSON and Parquet serialization layers:

1. JSON Output Validation: The submission.json file was successfully generated without schema violations, mapping task identifiers strictly to dual prediction candidates.
2. Parquet Data Lake Compatibility: A flattened representation was exported to submission.parquet (8,146 bytes) using PyArrow, confirming seamless interoperability with Big Data analytical tools and corporate decision support systems.

5 Conclusion

This study successfully constructed and deployed a standardized Kaggle benchmark evaluation framework integrating a Digital Twin model and an adaptive decision-making mechanism to address complex abstract reasoning tasks across the ARC-AGI benchmark suite (ARC-AGI-2 and ARC-AGI-3) [8]. By combining heuristic pattern search, dynamic geometric transformations (rotations, reflections, and grid cropping), metacognitive Key Performance Indicator (KPI) selfassessment, and adaptive memory management, the proposed engine effectively evaluates model confidence and optimizes strategic action selection within dynamic interactive environments.

The main contributions of this research are highlighted as follows:

- i. Adaptive Strategic Agent Architecture: Designed an intelligent agent guided by a Digital Twin framing that simulates managerial executive functions, evaluates internal confidence scores, and monitors action performance via dynamic KPI metrics.
- ii. Algorithmic Optimization & Robustness: Enhanced symmetry scanning algorithms, eliminated infinite execution loops, and implemented a reliable fallback strategy, thereby ensuring stable inference even when processing novel edge-case ARC grid patterns.
- iii. Standardized Pipeline Deployment: Achieved full offline execution capabilities within the Kaggle Notebook environment, exporting standardized multi-attempt predictions strictly formatted for submission.json and submission.parquet pipelines.

Future work will focus on expanding the Digital Twin architecture by integrating advanced deep neural network models, enhancing reinforcement learning strategies for interactive agent exploration, and extending this cognitive

assessment framework to real-world strategic managerial decision-making powered by AI and Big Data Analytics.

6 Future Work

To enhance the abstract reasoning capabilities and further optimize the strategic decision-making performance of the Digital Twin framework, future research will focus on the following key directions:

- i. Integration of Multimodal Large Language Models (LLMs): Extend the core reasoning engine (ARCSolver) by incorporating advanced Multimodal Reasoning Models and Vision-Language Architectures. This will enable automated feature extraction of complex visual geometry (e.g., dynamic spatial transformations, symmetry axes, and bounding-box topological patterns) beyond rule-based heuristic search.
- ii. Reinforcement Learning for Strategic Decision-Making & KPI Optimization: Enhance the StrategicManagerDigitalTwin module by integrating Reinforcement Learning (RL) techniques to dynamically adjust KPI reward-penalty matrices. This adaptive feedback mechanism will enable the agent to maintain an optimal balance between exploration and exploitation across high-dimensional interactive grid environments (64×64).
- iii. Multi-Agent Swarm Intelligence Framework: Implement a dynamic multi-agent coordination system (Swarm Intelligence) leveraging graph-based agent orchestration frameworks (e.g., LangGraph, Smolagents). Distributed specialized agents will collaboratively share pattern memory states and execute parallel candidate hypothesis verification (attempt1 and attempt_2) to boost performance on unseen edge cases.
- iv. Execution Efficiency & Offline Deployment Optimization: Refine data pipelines and memory management routines to improve execution throughput and minimize runtime latency. This ensures robust offline execution compatibility under strict benchmark constraints and resource-limited deployment settings while exporting structured predictions seamlessly via JSON and Parquet formats.

Acknowledgment

Thank you to my former professors in the Master's program in Software Engineering for enthusiastically sharing their experience and knowledge with their students. Thank you to

my former professors in the Education program for their highquality and sincere pedagogical methods in scientific research. Thank you to Tay Do University, where I have worked for over 20 years, experiencing so much happiness in the lecture halls. And I sincerely thanks kaggle.com and the organizers of the international conference for their regular organization and dedication to the international information technology community. And I also thank my colleagues and friends who have been with me until this moment, which can be considered a success despite the COVID-19 pandemic.

References

- [1] Rani, S., Bhambri, P., Kumar, S., Pareek, P.K. and Elngar, A.A. (2024) *AI-Driven Digital Twin and Industry 4.0: A Conceptual Framework with Applications*, CRC Press, pp. 338. Available online at: <https://www.perlego.com/book/4396454/aidriven-digital-twin-and-industry-40-a-conceptual-framework-with-applications> (accessed 11 September 2026).
- [2] Sepasgozar, S. and Shirowzhan, S. (2024) *Digital Twin Adoption and BIM-GIS Implementation*, Routledge. Available online at: <https://www.perlego.com/book/4504677/digital-twin-adoption-and-bimgis-implementation> (accessed 11 September 2026).
- [3] Mondal, S., Kumar, A. and Khan, M. (2025) *Digital Twins and ESG*, Scrivener Publishing, pp. 439. Available online at: <https://www.perlego.com/book/5201286/digital-twins-and-esg> (accessed 11 September 2026).
- [4] Minh, T.Q. (2026) 'ARC-AGI-2 strategic-decisionmaking', *Kaggle Competition Notebook - ARC Prize 2026*. Available online at: <https://www.kaggle.com/code/trnhquangminh140/arc-agi-2-strategic-decision-making> (accessed 11 September 2026).
- [5] Minh, T.Q. (2026) 'ARC-AGI-3 - decision-making process of managers', *Kaggle Code*. Available online at: <https://www.kaggle.com/code/trnhquangminh140/arc-agi-3-decision-making-process-of-managers> (accessed 11 September 2026).
- [6] Mondal, S., Kumar, A. and Khan, M. (2020) *Cognitive Radar*, Artech House, pp. 194. Available online at: <https://www.perlego.com/book/5201286/digital-twins-and-esg> (accessed 11 September 2026).
- [7] Guerci, J.R. (2010) *Cognitive Radar - The KnowledgeAided Fully Adaptive Approach*, Artech House, pp. 180. Available online at: <https://www.perlego.com/book/4666890/cognitive-radar-the-knowledgeaided-fully-adaptive-approach-pdf> (accessed 11 September 2026).
- [8] Francois Chollet, Greg Kamradt, Mike Knoop, and Mar'ia Cruz (2026) *ARC Prize 2026 - Paper Track*, Kaggle.com,. Available online at: <https://www.kaggle.com/competitions/arc-prize-2026-paper-track> (accessed 11 September 2026).