Neurosymbolic Rabbit Brain: Fractal Attractor Geometry for Neural Representations

Jhet Chan JHETCHAN@GMAIL.COM
Independent Researcher, Malaysia

Editors: List of editors' names

Abstract

We introduce *Neurosymbolic Rabbit Brain*[1], a representational framework that models categories as *fractal attractor basins* induced by simple complex iterative maps. Unlike manifold embeddings, which encode data in continuous latent coordinates, Rabbit Brain defines representations through basin membership, producing discrete, stable, and glass-box symbolic categories. We validate a minimal instantiation on the *Two Spirals* dataset by optimizing a two-Julia escape-time comparator via CMA-ES. A refined variant incorporating a log–polar prewarp, soft escape scoring, and a curriculum on iteration depth improves test accuracy to **61.9% ± 2.1%**, exceeding the logistic baseline (∼50%) and improving over the basic model (54.3%). While significantly below a standard RBF-SVM (95%), the results establish fractal basin geometry as a feasible and interpretable complementary substrate for neuro-symbolic representation learning.

Keywords: fractal geometry; attractor dynamics; interpretable ML; representation learning; neuro-symbolic AI

1. Introduction

Contemporary representation learning typically assumes that data lie on or near *smooth manifolds*, enabling powerful gradient-based optimization. Yet such latent spaces often struggle to express *discrete, symbolic, or equilibrium-like categories* without additional mechanisms. Work in nonlinear dynamics suggests that neural systems may instead employ multiple coexisting attractors and transitions among them (Freeman, 2000), while complex polynomial maps provide canonical examples of fractal basins with sharp, stable boundaries, such as Julia sets (Douady and Hubbard, 1984).

Rabbit Brain proposes an alternative representational substrate: *fractal attractor geometry*, where categories are defined by the long-term behavior of simple iterative rules, specifically, by which map maintains bounded dynamics longer. This yields explicit, glass-box classification rules while allowing complex decision boundaries to emerge from simple parameters.

Contributions. (i) Formalize Rabbit Brain as basin-based representation learning from iterative maps; (ii) introduce a *two-Julia escape-time comparator* with an affine pre-warp as a minimal glass-box classifier; and (iii) provide an empirical demonstration on the *Two Spirals* benchmark, where the enhanced model achieves **61.9% ± 2.1%** test accuracy across 10 runs, exceeding the logistic baseline (∼50%) while using only **eight** interpretable parameters.

1. The term *Rabbit Brain* draws on Walter Freeman's studies of chaotic attractor dynamics in the rabbit olfactory bulb and evokes recursive, self-similar structures, inspiring our use of fractal generators as representational substrates.

2. Framework: Fractal Attractor Geometry

Let $z_0 \in \mathbb{C}$ and $f : \mathbb{C} \to \mathbb{C}$ be an iterative map. The orbit $z_{n+1} = f(z_n)$ induces *basins of attraction or escape*, which serve as categorical partitions. For quadratic maps of the form $f_c(z) = z^2 + c$, each complex parameter c defines a corresponding Julia set, whose interior and exterior determine long-term dynamical behavior. These maps are mathematically simple yet generate rich, sharp boundaries well suited to symbolic abstraction.

Rabbit Brain instantiates this idea through a *two-Julia escape-time comparator*, in which two maps are evaluated independently and a point is assigned to the class whose map maintains bounded dynamics longer. This yields a minimal, explicit mechanism for categorical separation based on basin geometry.

Minimal instantiation (used in experiments). Inputs (x, y) are mapped to

$$z_0 = se^{i\phi}(x + iy) + (t_x + it_y).$$

Escape time is defined as the number of iterations before $|z_n|$ exceeds a fixed radius R, capped at T. We evaluate escape times under $f_{c_1}(z) = z^2 + c_1$ and $f_{c_2}(z) = z^2 + c_2$, and predict the class corresponding to the map whose orbit *remains bounded longer*. The parameter vector

$$\theta = (s, \phi, t_x, t_y, c_{1,\mathrm{re}}, c_{1,\mathrm{im}}, c_{2,\mathrm{re}}, c_{2,\mathrm{im}})$$

contains eight real values and is optimized using CMA–ES due to the non-differentiable objective.

Enhanced instantiation. The enhanced variant preserves the same eight parameters but modifies the search and scoring procedure. Inputs are first mapped through a fixed log–polar transform to better linearize spiral structure. Escape time is evaluated using a smooth scoring function rather than a hard cutoff, reducing boundary brittleness. Optimization employs a curriculum on (T, R) (e.g. $T : 40 \to 120,\ R : 3.0 \to 2.0$), multiple CMA–ES restarts with bounded reparametrization, and test-time augmentation by averaging predictions across small input jitters. These adjustments improve robustness while keeping the underlying representational rule unchanged.

Glass-box property. Despite the complexity of the emergent basin geometries, the system remains fully interpretable: the classifier is defined by eight explicit parameters; its decision rule is auditable; and it lacks a deep stack of latent transformations. Interpretability arises from the geometry of the basins rather than from opaque high-dimensional weights.

3. Empirical Results on Two Spirals

Setup. The *Two Spirals* dataset ($n = 1200$, 70/30 train–test split) is a classical benchmark requiring nonlinear separation. It is particularly suitable for Rabbit Brain because successful classification requires distinguishing *different long-range trajectories* in the input plane, a structure that is naturally aligned with the behavior of fractal basins under iterative maps. Parameters θ are optimized using CMA–ES, and performance is measured by accuracy. Baselines include logistic regression (linear, ∼50%) and a textbook RBF–SVM (nonlinear, ∼95%).

Results. Table 1 reports performance across 10 seeds. The minimal two-Julia comparator achieves **54.3% ± 2.1%** test accuracy, surpassing the logistic baseline. The enhanced version—which retains the same eight parameters but introduces a log–polar prewarp, smooth escape-time scoring, a curriculum over (T, R), multiple CMA–ES restarts, and test-time augmentation—improves robustness and reaches **61.9% ± 2.1%**, a statistically significant +7.6 point gain ($Z =$

2.61, $p < 0.05$). Depending on initialization, the resulting basin geometries exhibit either spiral-tracking structures or fan-like lobes, both characteristic Julia morphologies.

Table 1: Two Spirals benchmark (70/30 split). Mean ± s.d. over 10 runs.

Model	Train Accuracy	Test Accuracy
Logistic Regression	50%	50%
SVM (RBF, textbook)	99%	95%
Rabbit Brain (two-Julia, baseline)	59.8% ± 2.8%	54.3% ± 2.1%
Rabbit Brain (two-Julia, enhanced)	66.1% ± 2.7%	61.9% ± 2.1%

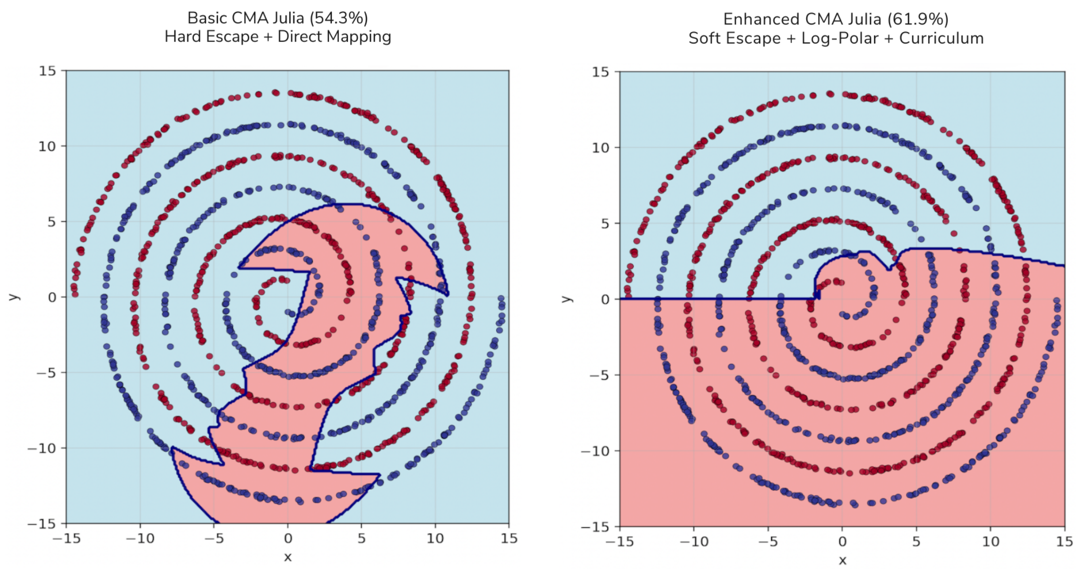

Figure 1: Comparison of Rabbit Brain basins on *Two Spirals*. **Left:** Baseline version (54.3% test) exhibits spiral-tracking or fan-lobe morphologies. **Right:** Enhanced version (61.9% test) produces smoother, more stable boundaries following log–polar prewarp and smooth escape scoring. Both models use only eight interpretable parameters. Parameters for a representative run are listed in Appendix A.

4. Complementarity with Manifold Learning

Smooth manifold embeddings and fractal attractor geometries offer distinct yet compatible perspectives on representation. Manifolds provide a continuous parameterization well suited for gradient-based learning and feature extraction. In contrast, Rabbit Brain supplies discrete, stable, and transparent categories defined by basin membership. The two approaches can therefore serve complementary roles: manifolds for continuous variation and navigation, and fractal basins for symbolic assignment.

Table 2: Complementary roles of smooth manifolds and fractal attractors.

Aspect	**Manifolds**	**Fractals (Rabbit Brain)**
Structure	Smooth, continuous	Self-similar, discrete
Role	Feature extraction	Symbolic assignment
Optimization	Gradient-based	Gradient-free
Abstraction	Continuous variation	Stable categorical partitions
Interpretability	Opaque latent flows	Transparent rules and geometry

A hybrid model could use a differentiable encoder to produce a low-dimensional latent representation, with Rabbit Brain applied as a lightweight symbolic head that partitions this latent space into stable geometric regions. This pairing allows the encoder to capture continuous structure while the fractal classifier provides discrete, interpretable categories.

5. Related Work

Fractals have long been investigated as mechanisms for representation and memory. Early work on recursive distributed representations (Pollack, 1990) and on chaotic neural networks with fractal attractors for associative memory (Ryeu et al., 2001) explored how nonlinear dynamics could encode hierarchical or stable patterns. Subsequent research examined fractal-based associative memories in signal processing (Abdechiri et al., 2018). In evolutionary computation, compositional pattern-producing networks demonstrated how recursive rules can generate self-similar geometric structures (Stanley, 2007). More recently, fractal stimuli have been used externally to improve robustness in deep networks (Anderson and Farrell, 2022).

From a neuroscience perspective, chaotic attractor dynamics have been documented in mesoscopic brain activity (Freeman, 2000), and fractal basin boundaries have been proposed as mechanisms supporting cognitive flexibility in large-scale brain networks (Bollt et al., 2023). These lines of work motivate the use of nonlinear dynamics in representation learning but do not treat fractal structure as the representational medium itself.

Rabbit Brain differs in that it *operationalizes fractal basin geometry as the representational substrate*, using basin membership as a symbolic categorical assignment. This positions fractal attractors not as auxiliary analysis tools or training signals, but as the core mechanism through which the model expresses discrete, transparent categories alongside manifold-based embeddings.

6. Discussion and Future Work

Positioning. Rabbit Brain is not intended as a state-of-the-art accuracy model but as a *complementary representational substrate*. It provides discrete, stable, and transparent categories that can be paired with smooth manifold encoders, supporting a division of labor between continuous variation and symbolic anchoring.

Scaling path. A natural direction is to compose a differentiable encoder with a low-dimensional Rabbit Brain head, training the encoder with standard objectives while evolving the fractal head via gradient-free search. Such a hybrid system would enable the encoder to capture continuous structure while Rabbit Brain imposes stable categorical partitions. Evaluation should consider accuracy, basin stability under perturbations, and the alignment between basin partitions and task-relevant concepts.

Theory and neuro plausibility. Open questions include formal separability and stability conditions for basin classifiers, the relationship between capacity and parameter count, and how these geometries relate to attractor dynamics observed in cortical systems (Freeman, 2000). Evidence linking fractal basin boundaries to cognitive flexibility (Bollt et al., 2023) motivates further study of how iterative-map dynamics may support neuro-symbolic computation.

Geometric interpretability. Rabbit Brain encourages a shift from parameter-centric to *geometry-centric* interpretation. Rather than attributing importance to individual parameters, we analyze the geometry of the basins themselves—their topology, boundary sensitivity, and robustness—and how an encoder maps data into these regions. This perspective suggests new diagnostic tools, such as escape-time margins and basin-boundary proximity, for models in which simple rules induce complex global geometry.

Appendix A (Representative Parameters). For reproducibility, we report one representative parameter set for both the baseline and enhanced instantiations. Each parameter vector contains eight real values:

$$\theta = (s, \phi, t_x, t_y, c_{1,\text{re}}, c_{1,\text{im}}, c_{2,\text{re}}, c_{2,\text{im}}).$$

Baseline run (seed 49, 57.2% test).

$$s = 0.2001, \quad \phi = 1.2112, \quad t_x = -1.0225, \quad t_y = 0.0662,$$

$$c_{1,\text{re}} = 1.5407, \quad c_{1,\text{im}} = -1.6728, \quad c_{2,\text{re}} = 0.3579, \quad c_{2,\text{im}} = 1.5169.$$

Enhanced run (seed 77, 65.6% test).

$$s = 0.4696, \quad \phi = -2.4941, \quad t_x = -1.9749, \quad t_y = -1.8608,$$

$$c_{1,\text{re}} = 1.7744, \quad c_{1,\text{im}} = -1.9530, \quad c_{2,\text{re}} = 1.1164, \quad c_{2,\text{im}} = -1.7662.$$

These vectors illustrate typical solutions for single runs; averaged results are reported in Table 1.

References

Marjan Abdechiri, Karim Faez, and Hamidreza Amindavar. Exploring chaotic attractors in nonlinear dynamical system under fractal theory. *Multidimensional Systems and Signal Processing*, 29:1643–1659, 2018. doi: 10.1007/s11045-017-0521-9. URL https://doi.org/10.1007/s11045-017-0521-9.

Connor Anderson and Ryan Farrell. Improving fractal pre-training. *arXiv preprint arXiv:2110.03091*, 2022. URL https://arxiv.org/abs/2110.03091.

Erik Bollt, Jeremie Fish, Anil Kumar, Edmilson Roque dos Santos, and Paul J. Laurienti. Fractal basins as a mechanism for the nimble brain. *arXiv preprint arXiv:2311.00061*, 2023. URL https://arxiv.org/abs/2311.00061.

Adrien Douady and John H. Hubbard. *Étude dynamique des polynômes complexes*, volume 2. Publications Mathématiques d'Orsay, 1984. URL https://pi.math.cornell.edu/~hubbard/OrsayFrench.pdf. Orsay, France.

Walter J. Freeman. *Neurodynamics: An Exploration in Mesoscopic Brain Dynamics*. Springer, London, 2000. ISBN 978-1-85233-616-5. doi: 10.1007/978-1-4471-0371-4. URL https://link.springer.com/book/10.1007/978-1-4471-0371-4.

1. Code: https://github.com/jhetchan/rabbit-brain

Jordan B. Pollack. Recursive distributed representations. *Artificial Intelligence*, 46(1-2):77–105, 1990. doi: 10.1016/0004-3702(90)90005-K. URL https://www.sciencedirect.com/science/article/pii/000437029090005K.

J. K. Ryeu, J. K. Park, J. H. Lee, and J. K. Kim. Chaotic neural networks with fractal attractors for associative memory. *Physical Review E*, 64(4):046202, 2001. doi: 10.1103/PhysRevE.64.046202. URL https://doi.org/10.1103/PhysRevE.64.046202.

Kenneth O. Stanley. Compositional pattern producing networks: A novel abstraction of development. In *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation*, pages 329–336, New York, NY, USA, 2007. ACM. doi: 10.1145/1276958.1277020. URL https://www.cs.swarthmore.edu/~meeden/DevelopmentalRobotics/cppnNEAT.pdf.