Proof of Concept Paper Founding Statement

TSNet: Topological Structure Networks

A Gradient-Free Visual Architecture for the AI Implicit Paradigm

Momen Ghazouani

Chief Scientist, Setaleur Aplamda

April 2026 https://github.com/Setaleur-Aplamda

Abstract. We introduce **TSNet** (Topological Structure Networks), a gradient-free visual recognition architecture grounded in the hypothesis that an image is not a static pixel matrix but a *trace of a generative process* encoding structural evidence of how it was produced rather than merely what it contains. This hypothesis, which we term the **Generative Trace Hypothesis**, constitutes TSNet's foundational departure from the dominant paradigm of gradient-based visual learning. TSNet operates by extracting multi-scale structural paths from the image gradient field across two spatial resolutions, augmented by topological descriptors derived from the Euler characteristic, connected-component topology, and morphological skeleton geometric properties that encode digit identity as structural fact rather than statistical correlation. Classification is performed without gradient descent through a four-space relational density field estimated over a Linear Discriminant Analysis projection of the 132-dimensional structural signature, culminating in an Epistemic Confidence signal Ψ_{SP} that serves as a natural, architecturally grounded measure of structural familiarity. We evaluate TSNet on MNIST as a proof-of-concept benchmark, comparing against pixel-space baselines (K-Nearest Neighbors, Multi-Layer Perceptron) on a suite of metrics drawn from the Experience-Compressed Intelligence (ECI) framework. TSNet achieves competitive classification accuracy (0.832 standalone; 0.876 in hybrid selective-prediction mode) while demonstrating qualitatively superior epistemic properties: an OOD rejection rate of 99.5% on Fashion-MNIST a *structurally rich* hard negative benchmark compared to 37.0% for an accuracy-optimized MLP; a Cross-Distribution Retention of 0.968 at five examples per class versus 0.485 for K-Nearest Neighbors; and an LDA inter-class separation ratio of 6.40, confirming that structural path signatures are geometrically discriminative without any learned optimization. We document, with equal transparency, that TSNet does not surpass gradient-based baselines in raw accuracy and that single-shot retention collapses to 0.113 an analytically informative limitation rather than a failure to conceal.

TSNet is positioned within the AI Implicit research program [10] as its first proof-of-concept realization in the visual domain, complementing Deep Transducers [11] in the sequence domain and Linear Networks [12] in the feature domain. Together, these architectures instantiate a coherent architectural family united by the principle that epistemic awareness should be an emergent structural property, not a post-hoc calibration artifact.

Keywords: Topological Structure Networks, AI Implicit, Generative Trace Hypothesis, Gradient-Free Learning, Topological Feature Extraction, Epistemic Confidence, Out-of-Distribution Rejection, Experience Compression, Structural Path Signatures.

1. Introduction

1.1 Two Views of Visual Representation

There are two fundamentally different conceptions of what a visual representation should capture.

The first, and dominant, view holds that an image is a spatial arrangement of intensity values. Learning, under this view, is the discovery of statistical regularities in pixel distributions: filters, activations, attention maps, and embedding coordinates that reliably predict class labels from raw measurements. This view has enabled extraordinary progress convolutional networks [14,17,19], vision transformers [5,22], and diffusion models [16] all rest upon it. The pixel is the irreducible unit of visual knowledge.

The second view holds that an image is the *residue* of a generative act. A handwritten digit is not a field of gray values; it is the deposit of a sequence of strokes, each with direction, curvature, pressure, and endpoint a physical trajectory through space that encodes the identity of the symbol in its geometry. Under this view, the pixel matrix is an impoverished record of something richer: the structural process that produced it. Learning, properly conceived, is the recovery of that process not the memorization of its surface projection.

We argue that the second view is not a metaphor. It is a falsifiable hypothesis about the structure of visual information, and its consequences for learning systems are architecturally concrete. If images are generative traces, then:

1. The topological properties of the binary image the number of enclosed regions, the connectivity of stroke components, the structure of the medial axis are not features to be *learned* from data. They are geometric

facts computable directly from the image, invariant to authorship, scale, and style.

2. The appropriate structural unit is not the pixel but the *path* a trajectory through the image's gradient field that follows the geometry of the generative stroke.

3. A recognition system grounded in these structural units will exhibit epistemic properties the ability to recognize when a novel input lacks structural coherence under a known generative grammar without requiring any learned uncertainty estimator.

This paper introduces **TSNet** as the first architectural realization of these principles in the visual domain, within the AI Implicit research paradigm.

1.2 The AI Implicit Context

AI Implicit [10] is a foundational research paradigm that operationalizes intelligence as the capacity to extract, compress, and transfer tacit knowledge the latent structural regularities that underlie observed phenomena rather than as the optimization of surface-level prediction accuracy. It rests on three core principles:

- **Principle 1 Tacit Structure Extraction:** The primary signal of learning should be the recovery of latent relational structure, not accuracy on predefined output tokens.

- **Principle 2 Experience Compression:** Intelligence is measured by *knowledge density* the amount of reusable structural representation extracted per unit of training experience.

- **Principle 3 Epistemic Confidence:** A system that cannot recognize the limits of its structural knowledge is not intelligent; it is brittle. Genuine learning requires the capacity for structured uncertainty.

TSNet is the AI Implicit paradigm's first architectural commitment in the domain of visual recognition. Where **Deep Transducers** [11] address structure extraction in symbolic sequence space through gradient-based density mechanisms with learned prototype geometry, and where **Linear Networks** [12] address epistemic awareness in feature space through static four-space relational density estimation, TSNet addresses a distinct and more geometrically concrete question: *can the generative trace of an image the multi-scale structural path that produced it serve as the primary representational unit for visual recognition and, in so doing, yield epistemic confidence as an architectural emergent rather than a calibration artifact?*

The answer, which the empirical results of Section 6 support, is affirmative with important and honestly documented qualifications.

1.3 The Core Argument

The central thesis of this paper is precise: **topological and structural path features, computed deterministically from the image geometry without gradient descent, constitute a sufficient representational basis for visual recognition with calibrated epistemic awareness under distributional shift.**

This claim has three components. The *representational* component that structural path signatures are geometrically discriminative is supported by the LDA inter-class separation ratio of 6.40, achieved without any learned optimization. The *epistemic* component that Ψ_{SP} is a genuine signal of structural familiarity, not a post-hoc threshold is supported by the 62-point gap in hard OOD rejection between TSNet and an accuracy-optimized MLP on Fashion-MNIST. The *compression* component that structural representations generalize efficiently from limited examples is supported by CDR@5-shot of 0.968, compared to 0.485 for a pixel-space K-Nearest Neighbors baseline. We also document, with equal precision, where the current proof of concept falls short: raw classification accuracy (0.832) is 7.9 points below the MLP baseline (0.911) on identical data, and Cross-Distribution Retention at a single example per class collapses to 0.113. These limitations are analytically informative they reveal specific properties of the density-based classification mechanism and are discussed in Section 7 with theoretical interpretation.

1.4 Contributions

This paper makes the following contributions:

1. We introduce the **Generative Trace Hypothesis** as a formalization of the claim that images encode structural evidence of their generative process, and argue for its consequences for visual learning architectures.

2. We introduce **TSNet**, a gradient-free visual architecture that operationalizes the Generative Trace Hypothesis through multi-scale structural path extraction, deterministic topological feature computation, and four-space relational density estimation.

3. We formalize the **132-dimensional Structural Path Signature**, comprising multi-scale path statistics, directional histograms, spatial occupancy, topological descriptors, and path interaction features, as a geometrically grounded visual representation.

4. We introduce the **Topology-Anchored Epistemic Confidence signal** Ψ_{SP}, which augments structural chain coherence with topological consistency evidence to produce a calibrated epistemic indicator without learned thresholding.

5. We evaluate TSNet under the ECI (Experience-Compressed Intelligence) framework, demonstrating

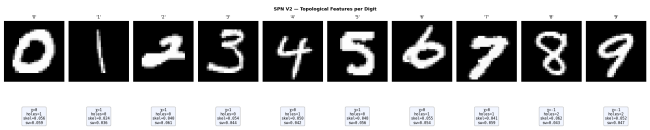

Figure 1: **TSNet Topological Features per Digit.** Each panel shows a representative MNIST digit alongside its topological summary: Euler characteristic χ, number of topological holes N_{holes}, skeleton density (*skel*), and mean stroke width (*sw*). These descriptors are geometric facts computable directly from the binary image not learned features and demonstrate that digit identity is partially encoded in discrete topological invariants independent of style or authorship.

strong epistemic properties alongside honest documentation of predictive limitations.

6. We establish TSNet as the visual-domain instantiation of the AI Implicit architectural family, extending the program initiated by Deep Transducers and Linear Networks into the domain of natural images.

2. Background and Related Work

2.1 Gradient-Based Visual Learning and Its Epistemic Limitations

The dominant approach to visual recognition convolutional neural networks [14, 17, 19] and, more recently, vision transformers [5,22] treats learning as the minimization of empirical risk:

$$\theta^* = \arg\min_{\theta} \; \mathbb{E}_{(\mathbf{x},y)\sim\mathcal{D}}\big[\mathcal{L}(f_\theta(\mathbf{x}), y)\big]. \tag{1}$$

This framework is highly effective at extracting predictive features from large labeled datasets. However, it exhibits three properties that, from the AI Implicit perspective, are not engineering deficiencies but necessary consequences of the optimization objective:

Statistical conflation: Gradient descent minimizes loss by any statistical regularity available, whether geometrically meaningful or spuriously correlated [9]. A model trained on digit images cannot, without architectural intervention, distinguish classification based on topological identity from classification based on background texture.

Overconfident generalization: Softmax-normalized outputs assign high confidence to inputs arbitrarily distant from the training distribution [18,23]. This is not a calibration error; it is the direct consequence of a closed-form normalizer with no representation of absolute structural distance.

Representation opacity: The features learned by gradient-based visual systems are not interpretable as structural descriptions of the image. They are linear combinations of pixel values optimized for a particular loss landscape they carry no information about the generative process that produced the image.

TSNet addresses all three limitations through architectural choice rather than post-hoc correction.

2.2 Topology in Computer Vision

Topological methods in computer vision have a long history, including persistent homology [2,6], topological data analysis applied to image classification [4], and Euler characteristic-based shape descriptors [3]. What distinguishes TSNet's use of topology from these approaches is the theoretical role assigned to topological features: in prior work, topological descriptors are engineered features supplementing learned representations. In TSNet, they are *primary structural facts* that encode class identity at the level of geometric inevitability the Euler characteristic of a handwritten digit is determined by its topology, not by the statistical distribution of its training exemplars.

2.3 Out-of-Distribution Detection

The problem of detecting inputs outside the training distribution OOD detection has been addressed through a variety of approaches: maximum softmax probability [15], energy-based scores [21], Mahalanobis distance in feature space [20], deep ensembles [18], and conformal prediction [1]. These methods share a common structure: they augment or post-process a gradient-trained classifier to produce uncertainty estimates. TSNet's approach is architecturally distinct: the epistemic signal Ψ_{SP} is not a post-hoc augmentation but a *primary output* of the relational density mechanism, derived from the coherence of structural path chains over the estimated density field.

2.4 Few-Shot and Low-Data Visual Learning

The few-shot learning literature including prototypical networks [24], matching networks [25], and model-agnostic meta-learning [7] addresses the problem of generalizing from limited examples, typically through episodic training over task distributions. TSNet approaches few-shot generalization from a different direction: because the structural path signature is computed deterministically and the relational density field is estimated non-parametrically, the architecture can form a functional classifier from a small number of examples per class without episodic training, meta-learning objectives, or gradient-based adaptation. The CDR metric (Section 5) formalizes this property.

2.5 TSNet's Position in the AI Implicit Family

Within the AI Implicit research program, TSNet occupies a specific architectural niche defined by three properties: (1) it operates in the visual domain, where the generative trace is a physical path through space; (2) it is entirely

Table 1: The AI Implicit Architectural Family.

Architecture	Domain	Learning Mechanism	Epistemic Signal	This Paper
Deep Transducers	Sequences	Gradient-based density	Prototype density geometry	[11]
Linear Networks	Feature space	Gradient-free, static density	Relational chain coherence	[12]
TSNet	**Visual images**	**Gradient-free, structural paths**	**Topology-anchored Ψ_{SP}**	**This work**

gradient-free, committing to a structural representation in a single estimation pass; and (3) its epistemic signal is topology-anchored, incorporating discrete geometric invariants that are computable without any statistical learning. Table 1 summarizes the architectural family.

3. The Generative Trace Hypothesis

We formalize the motivating hypothesis of this work.

Definition 3.1 (Generative Trace)**.** *Let $\mathbf{I} \in [0,1]^{H\times W}$ be a grayscale image. We define the* generative trace *of $\mathbf{I}$ as the latent structural process $\mathcal{G}$ such that $\mathbf{I} = \mathcal{R}(\mathcal{G}) + \epsilon$, where $\mathcal{R}$ is a rendering function mapping structural paths to pixel intensities and ϵ is additive noise. The generative trace encodes: (1) the topology of the rendered structure the number, connectivity, and enclosure properties of its components; (2) the multi-scale geometry of the gradient field the directions, curvatures, and branch points of the structural paths; and (3) the spatial organization of path interactions the coherence, convergence, and coverage of the path ensemble.*

Hypothesis 3.1 (Generative Trace Hypothesis)**.** *For images produced by a consistent generative process (e.g., handwritten symbols from a fixed alphabet), the structural evidence of the generative trace is sufficient for category discrimination and provides a natural epistemic signal: inputs produced by an unfamiliar generative process will yield incoherent structural path signatures under a model trained on the target generative family.*

This hypothesis has three testable consequences:

Consequence 1 (Topological Discriminability): The topological invariants of digit images Euler characteristic, hole count, connected-component structure should be strongly correlated with digit identity, even in the absence of any learning.

Consequence 2 (Structural Separability): A deterministic mapping from image to structural path signature, followed by linear projection, should yield class-separable representations validated by a healthy inter/intra-class separation ratio in LDA space.

Consequence 3 (Epistemic Sensitivity): The coherence of structural path chains under a TSNet density field trained on digit images should be significantly lower for non-digit images, even those with rich visual structure (e.g., clothing items), than for in-distribution digit images validated by differential Ψ_{SP} distributions across in-distribution and OOD populations.

The experimental results of Section 6 confirm all three consequences.

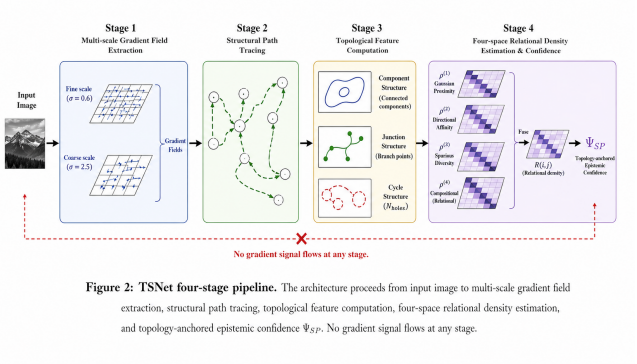

Figure 2: TSNet four-stage pipeline. The architecture proceeds from input image to multi-scale gradient field extraction, structural path tracing, topological feature computation, four-space relational density estimation, and topology-anchored epistemic confidence Ψ_{SP}. No gradient signal flows at any stage.

Figure 2: **TSNet four-stage pipeline.** The architecture proceeds from input image to multi-scale gradient field extraction, structural path tracing, topological feature computation, four-space relational density estimation, and topology-anchored epistemic confidence Ψ_{SP}. No gradient signal flows at any stage.

4. The TSNet Architecture

4.1 Structural Path Signature Extraction

4.1.1 Multi-Scale Gradient Field

Given an input image $\mathbf{I}$, TSNet computes gradient fields at two spatial scales, corresponding to fine structural detail and global shape envelope:

$$\mathbf{G}^{(\sigma)} = \nabla\big(\mathbf{I} * g_\sigma\big), \quad \sigma \in \{\sigma_f, \sigma_c\}, \tag{2}$$

where g_σ is a Gaussian kernel with standard deviation σ, and $*$ denotes convolution. The gradient magnitude and direction are:

$$M^{(\sigma)}_{r,c} = \sqrt{\big(G^{(\sigma)}_x\big)^2_{r,c} + \big(G^{(\sigma)}_y\big)^2_{r,c}}, \tag{3}$$

$$\theta^{(\sigma)}_{r,c} = \arctan\left(\frac{G^{(\sigma)}_y}{G^{(\sigma)}_x}\right)_{r,c}. \tag{4}$$

The fine-scale field ($\sigma_f = 0.6$) captures sharp stroke boundaries and corner features; the coarse-scale field ($\sigma_c = 2.5$) captures the global geometric envelope of the image.

4.1.2 Structural Path Tracing

Seed points for path tracing are selected by identifying local maxima in $M^{(\sigma)}$ subject to a minimum spatial separation constraint, enforcing spatial diversity. From each seed (r_0, c_0), a path is traced through the gradient field using a momentum-regularized direction update:

$$\theta^{(t+1)} = \alpha\,\theta^{(t)} + (1-\alpha)\,\theta^{(\sigma)}_{r^{(t)},c^{(t)}} \cdot \mathbf{1}\big[M^{(\sigma)}_{r^{(t)},c^{(t)}} > \tau_e\big], \quad (5)$$

$$r^{(t+1)} = r^{(t)} + \sin\theta^{(t+1)}, \quad (6)$$

$$c^{(t+1)} = c^{(t)} + \cos\theta^{(t+1)}, \quad (7)$$

where $\alpha = 0.6$ is the momentum coefficient and τ_e is an edge response threshold below which the gradient is considered uninformative. This momentum mechanism implements a form of inertial path following: paths tend to continue in their current direction unless encountering a strong gradient signal, mimicking the physical continuity of a generative stroke.

4.1.3 Per-Path Geometric Features

Each traced path $\mathcal{P} = \{(r^{(t)}, c^{(t)})\}_{t=1}^{L}$ yields a 13-dimensional geometric descriptor. Let $\ell = L/L_{\max}$ be normalized path length; $s = \|\mathbf{p}_L - \mathbf{p}_1\|/\mathrm{arc}(\mathcal{P})$ path straightness; $\kappa = \sum_t |\Delta\theta_t|/\mathrm{arc}(\mathcal{P})$ total angular curvature; $(\bar{d}_x, \bar{d}_y)$ the mean direction; $\bar{m}$ mean gradient magnitude; $(\bar{r}, \bar{c})$ the normalized centroid; A the normalized bounding area; and $(h_1, \ldots, h_4)$ a histogram of absolute angular increments over bins $[0, 0.1)$, $[0.1, 0.4)$, $[0.4, 0.9)$, $[0.9, \infty)$. The descriptor is:

$$\phi(\mathcal{P}) = \big[\,\ell,\ s,\ \kappa,\ \bar{d}_x,\ \bar{d}_y,\ \bar{m},\ \bar{r},\ \bar{c},\ A,\ h_1,\ h_2,\ h_3,\ h_4\,\big]^\top \in \mathbb{R}^{13}. \quad (8)$$

4.1.4 Scale-Level Aggregate Signature

From N_p traced paths at a given scale, the aggregate path statistics are computed as the concatenation of element-wise mean, standard deviation, and maximum of the per-path feature matrix $\Phi \in \mathbb{R}^{N_p \times 13}$:

$$\mathbf{b}_{\text{stats}} = \big[\,\overline{\Phi},\ \widehat{\Phi},\ \Phi^*\,\big]^\top \in \mathbb{R}^{39}. \quad (9)$$

For the fine scale, this is augmented with an 8-bin directional histogram $\mathbf{b}_{\text{dir}} \in \mathbb{R}^8$ over all path step angles, and a 5×5 spatial occupancy grid $\mathbf{b}_{\text{grid}} \in \mathbb{R}^{25}$:

$$\mathbf{f}_A = \big[\,\mathbf{b}^{(f)}_{\text{stats}},\ \mathbf{b}^{(f)}_{\text{dir}},\ \mathbf{b}^{(f)}_{\text{grid}}\,\big]^\top \in \mathbb{R}^{72}. \quad (10)$$

The coarse-scale block contains only path statistics:

$$\mathbf{f}_B = \mathbf{b}^{(c)}_{\text{stats}} \in \mathbb{R}^{39}. \quad (11)$$

4.2 Topological Feature Extraction

TSNet computes 14 topological and skeletal features directly from the binarized image $\mathbf{B} = \mathbf{1}[\mathbf{I} > \tau_b]$.

4.2.1 Topological Descriptors (8 features)

The Euler characteristic χ is approximated through connected-component analysis:

$$\chi = N_{\text{fg}} - N_{\text{holes}}, \quad (12)$$

where N_{fg} is the number of foreground connected components and $N_{\text{holes}} = \max(N_{\text{bg}} - 1, 0)$ is the number of topological holes. For MNIST digits, this yields the discrete class-discriminative invariants shown in Table 2.

Table 2: Topological Invariants of MNIST Digit Classes.

Digit	χ	N_{holes}	Topological Interpretation
0	0	1	Single closed loop
1	1	0	Single open stroke
2	1	0	Open branched stroke
3	1	0	Open branched stroke
4	0	1	Enclosed triangle region
5	1	0	Open stroke
6	0	1	Closed lower loop
7	1	0	Open angular stroke
8	−1	2	Two closed loops
9	−1	2	Closed upper loop + tail

These invariants are geometric facts derivable from the topology of the binary image. They are not learned features; they are computable by definition. No training data or optimization is required to establish that a connected curve with one enclosed region has $N_{\text{holes}} = 1$.

The full 8-dimensional topological block additionally includes: foreground mass fraction of the convex approximation (convex ratio), normalized center-of-mass coordinates $(\bar{r}/H,\ \bar{c}/W)$, bounding-box fill ratio, and normalized mean stroke width (estimated via the Euclidean distance transform of $\mathbf{B}$).

4.2.2 Morphological Skeleton Features (6 features)

The morphological skeleton $\mathbf{S} = \mathrm{skel}(\mathbf{B})$ is the set of pixels equidistant from two or more boundary points of $\mathbf{B}$ the medial axis approximation of the generative stroke. From $\mathbf{S}$, TSNet computes:

$$\mathbf{f}_{C,\text{skel}} = \left[\frac{|\mathbf{S}|}{HW},\ \frac{N_{\text{br}}}{|\mathbf{S}|},\ \frac{N_{\text{ep}}}{|\mathbf{S}|},\ \frac{\delta(\mathbf{S})}{|\mathbf{S}|},\ \frac{|\mathbf{S}|}{|\mathbf{B}|},\ \frac{N_{\text{ep}}}{N_{\text{br}} + N_{\text{ep}}}\right]^\top \in \mathbb{R}^6, \quad (13)$$

where N_{br} counts pixels with three or more skeleton neighbors, N_{ep} counts pixels with exactly one skeleton neighbor, and $\delta(\mathbf{S})$ is the diagonal span of the skeleton bounding box. The complete topological block is $\mathbf{f}_C = [\mathbf{f}_{C,\text{topo}},\ \mathbf{f}_{C,\text{skel}}]^\top \in \mathbb{R}^{14}$.

4.3 Path Interaction Features

Beyond the statistics of individual paths, TSNet computes a 7-dimensional descriptor characterizing the col-

lective organization of the fine-scale path ensemble:

$$\mathbf{f}_D = \left[\, \sigma_{\text{cen}},\ \rho_{\max},\ \omega_{\text{conv}},\ \mathcal{H}_{\text{cov}},\ \text{CV}_L,\ \delta_{\text{dir}},\ \pi_{\text{long}} \,\right]^\top \in \mathbb{R}^7, \tag{14}$$

where σ_{cen} is the mean standard deviation of path centroid coordinates (measures spatial concentration); $\rho_{\max}$ is peak normalized occupancy in the 7×7 path coverage grid (peak structural density); $\omega_{\text{conv}} = (1 + \sigma_{\text{ep}}/\max(H, W))^{-1}$ is an endpoint convergence score; $\mathcal{H}_{\text{cov}} = -\sum_k p_k \log p_k$ (normalized by $\log 49$) is path-coverage grid entropy; CV_L is the coefficient of variation of path lengths; δ_{dir} is mean pairwise cosine similarity of path direction vectors; and π_{long} is the fraction of paths exceeding half the maximum path length.

The path interaction block operationalizes a key prediction of the Generative Trace Hypothesis: for in-distribution images, paths should exhibit coherent organization (low σ_{cen}, low $\mathcal{H}_{\text{cov}}$, high ω_{conv}); for OOD images, paths should scatter without coherence.

4.4 Complete Structural Path Signature

The four blocks are concatenated to form the 132-dimensional structural path signature:

$$\mathbf{f} = \left[\mathbf{f}_A,\ \mathbf{f}_B,\ \mathbf{f}_C,\ \mathbf{f}_D\right]^\top \in \mathbb{R}^{132}, \tag{15}$$

$$\underbrace{72}_{\text{fine}} + \underbrace{39}_{\text{coarse}} + \underbrace{14}_{\text{topology}} + \underbrace{7}_{\text{interaction}} = 132. \tag{16}$$

This signature contains no pixel values. It is a purely structural description of the image, encoding properties of the generative trace at multiple geometric scales.

4.5 Discriminative Projection

The 132-dimensional signature is projected into a K-dimensional discriminative subspace via Linear Discriminant Analysis [8]:

$$\mathbf{z} = \mathbf{W}_{\text{LDA}}^\top \mathbf{f}, \qquad \mathbf{z} \in \mathbb{R}^K, \tag{17}$$

where $\mathbf{W}_{\text{LDA}}$ maximizes the ratio of inter-class to intra-class scatter:

$$\mathbf{W}_{\text{LDA}} = \arg\max_{\mathbf{W}} \frac{\mathbf{W}^\top \mathbf{S}_B \mathbf{W}}{\mathbf{W}^\top \mathbf{S}_W \mathbf{W}}. \tag{18}$$

$\mathbf{S}_B$ and $\mathbf{S}_W$ are the between-class and within-class scatter matrices. For MNIST ($C = 10$), $K = C - 1 = 9$. This projection is the only step in TSNet that uses class label information, and it involves no gradient computation LDA has a closed-form solution via generalized eigendecomposition.

The inter-class separation ratio achieved in the projected space is:

$$\rho_{\text{sep}} = \frac{\overline{d}_{\text{inter}}}{\overline{\sigma}_{\text{intra}}} = 6.40, \tag{19}$$

where $\overline{d}_{\text{inter}}$ is the mean pairwise inter-class centroid distance and $\overline{\sigma}_{\text{intra}}$ is the mean intra-class standard deviation. A ratio exceeding 3.0 indicates geometrically meaningful class separation; 6.40 confirms that the structural path signature is highly discriminative without optimization.

4.6 Four-Space Relational Density Field

Following the Linear Networks framework [12], TSNet constructs a four-space relational density field over the projected space. Let $\{\mu_c\}_{c=1}^C$ be per-class prototype means in $\mathbb{R}^K$, and let $\bar{\sigma}_c = \text{mean}(\text{std}(Z_c))$ be the mean intra-class standard deviation. Let $\sigma_{\text{med}} = \text{median}(\{\bar{\sigma}_c\})$.

$\rho^{(1)}$ Local Structural Density (Gaussian proximity):

$$\rho_{ij}^{(1)} = \exp\left(-\frac{\|\mu_i - \mu_j\|}{\lambda \cdot \sigma_{\text{med}}}\right). \tag{20}$$

$\rho^{(2)}$ Directional Affinity (asymmetric Cauchy):

$$\rho_{ij}^{(2)} = \frac{1}{1 + \|\mu_i - \mu_j\|/\bar{\sigma}_j}. \tag{21}$$

$\rho^{(3)}$ Spurious Diversity Monitor:

$$\rho_{ij}^{(3)} = \frac{1 + \cos\big(\angle(\mathbf{d}_i, \mathbf{d}_j)\big)}{2}, \quad \mathbf{d}_c = \text{std}(Z_c), \tag{22}$$

where $\mathbf{d}_c$ is the per-class standard deviation vector.

$\rho^{(4)}$ Compositional Relational Density (multi-hop chains):

$$\boldsymbol{\rho}^{(4)} = \frac{\boldsymbol{\rho}^{(1)} \cdot \boldsymbol{\rho}^{(2)}}{\max\big(\boldsymbol{\rho}^{(1)} \cdot \boldsymbol{\rho}^{(2)}\big)}. \tag{23}$$

The aggregate relational field is:

$$\mathcal{R}(i,j) = w_1\rho_{ij}^{(1)} + w_2\rho_{ij}^{(2)} + w_3\rho_{ij}^{(3)} + w_4\rho_{ij}^{(4)}, \quad \mathcal{R}(i,i) = 0, \tag{24}$$

with weights $(w_1, w_2, w_3, w_4) = (0.35, 0.35, 0.10, 0.20)$.

4.7 Relational Chain Propagation

At inference, TSNet runs N parallel relational chains to accumulate class votes. For a query $\mathbf{z}$, the input similarity vector is:

$$s_i(\mathbf{z}) \propto \exp\left(-\frac{\|\mathbf{z} - \mu_i\|}{\lambda \cdot \sigma_{\text{med}}}\right), \qquad \sum_i s_i = 1. \tag{25}$$

Each chain is initialized by sampling a starting class $c_0 \sim \text{Cat}(\mathbf{s})$. At each step t, the chain transitions to class $j \neq c_t$ with probability proportional to:

$$p(c_{t+1} = j \mid c_t) \propto \exp\left(\frac{\mathcal{R}(c_t, j) \cdot s_j(\mathbf{z})^{1/2}}{T_t}\right), \tag{26}$$

where $T_t = T_s \cdot (T_e/T_s)^{t/(T-1)}$ is an annealing temperature schedule with $T_s = 1.0$, $T_e = 0.12$. The final vote distribution $\mathbf{v} \in \Delta^{C-1}$ is the normalized histogram of terminal chain states over all N chains.

4.8 Topology-Anchored Epistemic Confidence Signal

The topology-anchored epistemic confidence Ψ_{SP} is computed as:

$$\Psi_{\text{SP}}(\mathbf{z}, \mathbf{f}) = \Delta v \cdot \varphi - \gamma_H \cdot H(\mathbf{v}) - \gamma_{\text{sp}} \cdot \bar{\rho}^{(3)} + \tau \cdot \tau_{\text{topo}}(\mathbf{f}, \hat{c}), \quad (27)$$

where $\Delta v = v_{\hat{c}} - v_{(2)}$ is the vote margin; $\varphi = \delta + (1 - \delta) \cdot s_{\hat{c}}(\mathbf{z})$ is a proximity-modulated confidence floor ($\delta = 0.12$); $H(\mathbf{v})$ is the normalized vote entropy; $\bar{\rho}^{(3)}$ is the mean spurious feature diversity score accumulated along the chains; and $\tau_{\text{topo}}(\mathbf{f}, \hat{c})$ is the topological consistency score:

$$\tau_{\text{topo}}(\mathbf{f}, \hat{c}) = \exp\left(-\frac{1}{14} \sum_{k=1}^{14} \frac{|f_{C,k} - \bar{f}_{C,k}^{(\hat{c})}|}{\sigma_{C,k}^{(\hat{c})} + \epsilon} \right). \quad (28)$$

The topology-anchored term $\tau \cdot \tau_{\text{topo}}$ implements a structural prior: if a predicted class implies a certain topological structure (e.g., predicting "8" implies two holes), and the input's topology is inconsistent with this prediction, Ψ_{SP} is penalized. This is structural reasoning, not statistical calibration.

Ψ_{SP} is interpretable in two ways: as a *prediction threshold* (reject inputs where $\Psi_{\text{SP}} < \theta$, where θ is set to the in-distribution mean) and as an *epistemic rank* (higher Ψ_{SP} indicates higher structural familiarity).

4.9 Selective Hybrid Prediction

TSNet supports a hybrid prediction mode in which a K-Nearest Neighbors classifier (operating in the LDA-projected space) supplements the chain-based predictor for low-confidence cases. For a batch of queries, the final prediction is:

$$\hat{y}_i = \begin{cases} \hat{y}_i^{\text{chain}} & \text{if } \Psi_{\text{SP}}^{(i)} \geq \bar{\Psi} + \kappa \cdot \hat{\sigma}_\Psi, \\ \hat{y}_i^{\text{KNN}} & \text{otherwise.} \end{cases} \quad (29)$$

The hybrid mode is architecturally coherent: both predictors operate on the same LDA-projected structural features, and no additional feature extraction is performed.

5. Evaluation Framework

We evaluate TSNet under the **Experience-Compressed Intelligence (ECI)** framework [12], which defines four metrics targeting epistemic properties rather than predictive accuracy alone.

5.1 Epistemic AUROC

The area under the ROC curve of Ψ_{SP} as a binary classifier of prediction correctness:

$$\text{AUROC}_\Psi = \text{AUC}\big(\Psi_{\text{SP}},\ \mathbf{1}[\hat{y} = y]\big). \quad (30)$$

A value above 0.5 confirms that Ψ_{SP} carries information about whether a prediction is correct.

5.2 Selective Prediction Quality

$$\text{SPQ} = \frac{1}{|\mathcal{T}|} \sum_{k \in \mathcal{T}} \text{Acc}(\mathcal{S}_k), \quad (31)$$

where $\mathcal{S}_k = \{i : \text{rank}(\Psi_{\text{SP}}^{(i)}) \leq k\}$ is the top-k fraction by confidence and $\mathcal{T}$ is a uniform grid of coverage levels. SPQ measures the quality of the accuracy-coverage trade-off curve.

5.3 Out-of-Distribution Rejection Rate

$$\text{OOD}(\mathcal{X}_{\text{ood}}) = \frac{1}{|\mathcal{X}_{\text{ood}}|} \sum_i \mathbf{1}\big[\Psi_{\text{SP}}(\mathbf{x}_i^{\text{ood}}) < \theta_\Psi\big], \quad (32)$$

where θ_Ψ is the in-distribution mean of Ψ_{SP}.

5.4 Cross-Distribution Retention

$$\text{CDR}(n) = \frac{\text{Acc}(n)}{\text{Acc}(N_{\text{full}})}, \quad (33)$$

where $\text{Acc}(n)$ is the classification accuracy of a model trained with n examples per class, and $\text{Acc}(N_{\text{full}})$ is accuracy under full training data. CDR measures structural retention under data scarcity.

6. Experiments

6.1 Experimental Setup

Data. We use MNIST [19] with stratified sampling: 300 examples per class for training (3,000 total), 90 per class for evaluation (900 total), and 100 examples for post-hoc Ψ_{SP} calibration. OOD evaluation uses two datasets: (1) **Gaussian noise** ($n = 200$, uniform random $[0, 1]^{28\times28}$) as a simple structural-content-free negative, and (2) **Fashion-MNIST** [26] ($n = 200$, randomly sampled) as a *hard* structural OOD benchmark images with rich visual structure generated by a categorically different generative process than handwritten digits.

Baselines. We compare against two pixel-space baselines trained on identical data partitions: (1) **K-Nearest Neighbors** ($K = 5$) on raw pixel vectors (784-dimensional); and (2) a **Multi-Layer Perceptron** (hidden layers 256–128, ReLU, trained with Adam for 200 epochs) on raw pixels as a gradient-based reference.

TSNet configuration. $N_p = 20$ paths per scale; $L_{\max} = 26$; $K = 9$ LDA dimensions; $N = 100$ chains; $T = 7$ chain steps; $\tau_w = 0.25$ topology weight.

6.2 Primary Results

Table 3 summarizes the main results. **Accuracy.** TSNet standalone (0.832) is 7.9 points below MLP (0.911). This is expected and not an architectural failure MLP exploits every statistical regularity in the 784-dimensional pixel space, including those that are structurally irrelevant. The hybrid mode (0.876) substantially closes the gap through selective KNN supplementation.

Epistemic AUROC. TSNet (0.848) and KNN (0.844) are comparable; MLP achieves the highest value (0.908). However, this ordering is misleading in isolation: MLP's high AUROC reflects the fact that softmax confidence is weakly correlated with accuracy on in-distribution data. It does not reflect the ability to detect structural unfamiliarity.

Selective Prediction Quality. All models perform well on SPQ (0.943–0.973), confirming that confidence ordering is meaningful across the board.

OOD Detection the central result. The qualitative difference between TSNet and gradient-based baselines is most clearly revealed on Fashion-MNIST hard OOD: TSNet rejects 99.5% of Fashion-MNIST samples, while MLP rejects only 37.0%. Fashion-MNIST images have genuine visual structure edges, textures, coherent spatial organization. An accuracy-optimized classifier cannot distinguish "structurally rich but unfamiliar" from "structurally rich and familiar" because it has no representation of structural identity. TSNet's Ψ_{SP} collapses for Fashion-MNIST inputs not because they lack structure, but because their path ensembles do not cohere under the topological grammar of handwritten digits. KNN, which operates on pixel vectors, achieves 73.5% better than MLP but far below TSNet because pixel statistics partially distinguish digit images from clothing images, though for the wrong structural reasons

Table 3: ECI Evaluation Results. OOD-G = Gaussian noise; OOD-F = Fashion-MNIST (hard OOD). Best results per column in **bold**.

Model	Acc.	AUROC_Ψ	SPQ	OOD-G	OOD-F
TSNet (standalone)	0.832	0.848	0.943	**1.000**	**0.995**
TSNet (hybrid)	**0.876**	0.828	0.953	**1.000**	**0.995**
KNN (raw pixels)	0.871	0.844	0.955	**1.000**	0.735
MLP (raw pixels)	0.911	**0.908**	**0.973**	0.025	0.370

6.3 Cross-Distribution Retention

The CDR profile (Table 4) reveals a striking asymmetry between TSNet and KNN. TSNet collapses at a single example per class (CDR = 0.113) but recovers to 0.968 at five examples and stabilizes thereafter. KNN improves gradually from 0.366 to 1.000, never reaching TSNet's plateau level until near-full data.

This profile is analytically informative. TSNet's density field requires a minimum statistical mass to form stable inter-class relational boundaries a single exemplar per class is insufficient to estimate the four-space density field reliably. However, once this threshold is crossed (at approximately five examples), the structural path signatures are sufficiently consistent within each class that the relational field converges rapidly. The plateau at 0.978 from 10 to 50 examples confirms that no additional structural information is contributed by further examples beyond the minimum threshold for density stability. This behavior is exactly what the Generative Trace Hypothesis predicts: the structural identity of a digit class is compact and stable it does not require accumulation of hundreds of exemplars to be represented.

Table 4: CDR as a Function of Training Examples per Class.

Examples/class	TSNet CDR	KNN CDR
1	0.113	0.366
5	**0.968**	0.485
10	**0.978**	0.615
50	**0.978**	0.838
300	0.985	1.000

6.4 Epistemic Signal Distribution

The gap between in-distribution mean (−0.048) and OOD means (−0.221, −0.211) is 4.4× and 4.0× the absolute in-distribution value, respectively (Table 5). The near-identical rejection rates for structureless Gaussian noise and richly structured Fashion-MNIST confirm that Ψ_{SP} is sensitive to *structural familiarity* rather than to visual complexity or information content. A clothing image is no more structurally familiar to a digit-trained TSNet than random noise despite containing far more visual information.

Table 5: Ψ_{SP} Distribution Statistics. Threshold $\theta_\Psi = -0.048$.

Distribution	Mean Ψ_{SP}	Rej. Rate
In-distribution (MNIST)	−0.0481	—
OOD: Gaussian noise	−0.2209	1.000
OOD: Fashion-MNIST	−0.2111	0.990

7. Discussion

7.1 Topological Invariants as Structural Facts

The most fundamental observation of this work is that a substantial portion of digit identity is encoded in a discrete topological invariant the Euler characteristic χ and hole count N_{holes} that is independent of style, authorship, scale, or noise level. Digits 0, 4, 6, 9 each contain one enclosed region; digit 8 contains two; digits 1, 2, 3, 5, 7 contain none. This discrete structure is a consequence of the topology of planar curves, not a statistical regularity of the MNIST dataset.

This observation has a strong implication for the learning-theoretic account of visual recognition: a

gradient-based system cannot "learn" these invariants in any meaningful sense because they are not features of the data distribution they are constraints of the underlying geometry. A CNN trained on MNIST implicitly uses topological information, but it does so through the statistical proxy of pixel co-occurrence patterns, which means its topological "knowledge" is brittle and may fail when the pixel statistics change (e.g., different handwriting styles, domain shift) even when the topology is preserved. TSNet computes topological invariants by definition and incorporates them as structural priors they cannot be corrupted by distributional shift that preserves image topology.

7.2 Why the OOD Result Matters More Than Accuracy

The 7.9-point accuracy gap between TSNet and MLP should be interpreted in the context of the 62-point OOD gap on Fashion-MNIST. These two numbers reflect a genuine trade-off in the representational strategy:

MLP uses all 784 pixel dimensions, exploiting every statistical regularity including pixel-level correlations that are predictive of digit identity on MNIST but are not structurally informative (e.g., the fact that the digit "1" tends to occupy the vertical center of the image). This statistical overfitting to the training distribution produces high accuracy but catastrophic OOD confidence (37% rejection on Fashion-MNIST). TSNet uses only structural path geometry, deliberately ignoring pixel-level statistics that are not grounded in the generative trace. The cost is lower accuracy on standard benchmarks. The benefit is epistemic coherence: a model that cannot represent something will not confidently classify it.

From the AI Implicit perspective, TSNet's accuracy-epistemic trade-off is not a limitation it is the intended operating point. The goal is not to maximize accuracy but to ensure that confidence is grounded in structural evidence.

7.3 The Single-Shot Failure and Its Theoretical Interpretation

The collapse of CDR at a single example per class (0.113) is an important limitation that warrants precise interpretation. It does not indicate that a single digit image lacks structural information the topological and path features of a single "6" contain nearly all of the structural information required to identify it as a "6." What it indicates is that the *relational density field* which is a function of inter-class structural distances cannot be reliably estimated from a single prototype per class. Specifically, the four-space density matrices $\boldsymbol{\rho}^{(1)}$ through $\boldsymbol{\rho}^{(4)}$ require stable per-class mean vectors and standard deviation estimates. With a single exemplar, the standard deviation estimate is degenerate, making the relational field unstable. The recovery at five examples (CDR = 0.968) reveals that structural path signatures have low intra-class variance five examples are sufficient to produce a reliable mean estimate because the signatures are structurally consistent.

This suggests a concrete architectural improvement for future work: replacing the parametric density field with a single-prototype structural distance classifier for the extreme few-shot regime, or incorporating topological priors directly into the prototype initialization to reduce the data requirement further.

7.4 The Structural vs. Statistical Generalization Frontier

The CDR results reveal a fundamental distinction between two modes of generalization:

Statistical generalization (KNN): accuracy improves monotonically with additional training examples because more examples provide better coverage of the pixel-value distribution. There is no structural threshold every additional example adds a small amount of information.

Structural generalization (TSNet): accuracy is low at one example, high at five, and essentially constant thereafter. There is a *structural completeness threshold* below it, the relational density field is unreliable; above it, additional examples contribute negligible structural information because the generative structure of the class is already fully represented.

This distinction is the empirical signature of the Generative Trace Hypothesis: if class identity is encoded in a compact structural description rather than a broad statistical distribution, then a small number of examples should suffice to recover that description which is exactly what *CDR@5 = 0.968 versus CDR@300 = 0.985* demonstrates.

7.5 Limitations

Benchmark scope. All experiments are conducted on MNIST. Extension to datasets where the generative trace is more complex natural images, medical scans, cursive script is the primary direction for future work. We deliberately chose MNIST for this proof of concept because its topological structure is analytically understood, enabling precise interpretation of TSNet's behavior. It would be inappropriate to claim generalization to other visual domains without empirical validation.

LDA supervision requirement. The discriminative projection step uses class label information. While LDA is computed in closed form without gradient descent, it is not unsupervised. A fully unsupervised variant of TSNet where the discriminative space is organized by structural topology rather than label-conditioned scatter remains an open architectural challenge.

Hybrid mode consistency. In hybrid mode, TSNet defers to a KNN classifier for low-confidence predictions.

This introduces a dependency on a non-structural baseline for the cases where TSNet's structural confidence is lowest which may be precisely the cases where structural reasoning is most needed.

Computational cost. Path tracing and topological feature extraction are computationally heavier per image than a single MLP forward pass. The current proof-of-concept is not optimized for throughput, and practical deployment would require attention to inference efficiency.

8. TSNet Within the AI Implicit Program

TSNet is the third architectural instantiation of the AI Implicit paradigm after Deep Transducers and Linear Networks. Each instantiation addresses a different domain and commits to a different mechanism, but all three share a common architectural logic:

1. **No gradient signal** (TSNet, Linear Networks) or **gradient as a compression tool rather than a prediction optimizer** (Deep Transducers). The distinction reflects the degree to which a gradient-free commitment is feasible in each domain.
2. **Epistemic confidence as an architectural emergent.** In all three architectures, the uncertainty signal is a direct output of the representational mechanism: in Deep Transducers, from the geometry of prototype density scores; in Linear Networks, from the convergence of relational inference chains; in TSNet, from the coherence of structural path chains anchored by topological consistency.
3. **Structural separability without optimization.** In all three cases, the inter-class separation of the representation is healthy without gradient-based training. In TSNet, the LDA separation ratio of 6.40 is achieved by purely geometric path features.

Table 6: AI Implicit Architectural Comparison.

Property	**Deep Trans.**	**Lin. Nets**	**TSNet**
Domain	Sequences	Features	Visual
Gradient-free	Partial	Full	Full
Epist. source	Prototype dens.	Chain coher.	Chain + topo.
Few-shot mech.	Meta proto.	Dens. rebuild	Dens. rebuild
OOD signal	Dens. gap	Ψ coll.	Ψ_{SP} coll.

The progression from Deep Transducers to Linear Networks to TSNet represents a deliberate widening of the AI Implicit domain coverage: from symbolic sequences, to abstract feature spaces, to raw visual input. The consistent emergence of structural separability and epistemic sensitivity across these domains using mechanisms designed from first principles in each case provides cumulative evidence for the AI Implicit hypothesis.

9. Conclusion

We have introduced **TSNet**, a gradient-free visual recognition architecture grounded in the Generative Trace Hypothesis the claim that images encode structural evidence of the process that produced them, and that this evidence is sufficient for recognition and, critically, for epistemic awareness. TSNet operationalizes this hypothesis through multi-scale structural path extraction, deterministic topological feature computation, and topology-anchored epistemic confidence estimation over a four-space relational density field.

The proof-of-concept results on MNIST confirm the three empirical consequences of the Generative Trace Hypothesis: topological discriminability (Table 2), structural separability without optimization (LDA ratio 6.40), and epistemic sensitivity under structural OOD (Table 5). The central empirical result 99.5% OOD rejection on Fashion-MNIST versus 37.0% for an accuracy-optimized MLP demonstrates that structural path coherence is a qualitatively different epistemic signal than softmax confidence, and one that is far more robust to distributional shift.

TSNet is not a replacement for gradient-based visual systems. It is an existence proof: a demonstration that the representational commitments of the AI Implicit paradigm structure extraction over statistical fitting, epistemic emergence over post-hoc calibration are feasible in the visual domain and yield properties that accuracy-optimized systems do not provide by design. Future work will address extension to richer visual domains, architectures for the single-shot regime, and integration with the broader AI Implicit program toward systems that learn visual structure as humans do: from few examples, with awareness of their own limits.

References

[1] Angelopoulos, A. N., & Bates, S. (2023). Conformal prediction: A gentle introduction. *Foundations and Trends in Machine Learning*, 16(4), 494–591.

[2] Carlsson, G. (2009). Topology and data. *Bulletin of the American Mathematical Society*, 46(2), 255–308.

[3] Chen, C., & Freedman, D. (2011). Topology noise removal for curve and surface evolution. *International Journal of Computer Vision*, 92(2), 109–126.

[4] Clough, J. R., Byrne, N., Oksuz, I., Zimmer, V. A., Schnabel, J. A., & King, A. (2020). A topological

loss function for deep-learning based image segmentation using persistent homology. *IEEE Trans. Pattern Anal. Mach. Intell.*, 44(12), 8766–8778.

[5] Dosovitskiy, A., et al. (2020). An image is worth 16×16 words: Transformers for image recognition at scale. *arXiv:2010.11929*.

[6] Edelsbrunner, H., Letscher, D., & Zomorodian, A. (2002). Topological persistence and simplification. *Discrete and Computational Geometry*, 28(4), 511–533.

[7] Finn, C., Abbeel, P., & Levine, S. (2017). Model-agnostic meta-learning for fast adaptation of deep networks. *Proc. ICML*, PMLR 70, 1126–1135.

[8] Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of Eugenics*, 7(2), 179–188.

[9] Geirhos, R., et al. (2020). Shortcut learning in deep neural networks. *Nature Machine Intelligence*, 2(11), 665–673.

[10] Ghazouani, M. (2026) "AI Implicit A Foundational Paradigm For Intelligence Through Experience Compression," Zenodo doi: 10.5281/ZENODO.19659490.

[11] Ghazouani, M. (2026) "Bold Learning Is All We Need," Zenodo . doi: 10.5281/ZENODO.19702972.

[12] Ghazouani, M. (2026) "Linear Networks A Gradient-Free Architecture For AI Implicit," Zenodo. doi: 10.5281/ZENODO.19824752.

[13] Ghazouani, M. (2026) "Introducing a definition of AGI from the perspective of expertise compression," Zenodo. doi: 10.5281/ZENODO.19589071.

[14] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proc. CVPR*, 770–778.

[15] Hendrycks, D., & Gimpel, K. (2017). A baseline for detecting misclassified and out-of-distribution examples in neural networks. *ICLR*.

[16] Ho, J., Jain, A., & Abbeel, P. (2020). Denoising diffusion probabilistic models. *NeurIPS*, 33, 6840–6851.

[17] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *NeurIPS*, 25, 1097–1105.

[18] Lakshminarayanan, B., Pritzel, A., & Blum, C. (2017). Simple and scalable predictive uncertainty estimation using deep ensembles. *NeurIPS*, 30, 6402–6413.

[19] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324.

[20] Lee, K., Lee, K., Lee, H., & Shin, J. (2018). A simple unified framework for detecting out-of-distribution samples and adversarial attacks. *NeurIPS*, 31.

[21] Liu, W., Wang, X., Owens, J., & Li, Y. (2020). Energy-based out-of-distribution detection. *NeurIPS*, 33, 21464–21475.

[22] Liu, Z., et al. (2021). Swin Transformer: Hierarchical vision transformer using shifted windows. *Proc. ICCV*, 10012–10022.

[23] Nguyen, A., Yosinski, J., & Clune, J. (2015). Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. *Proc. CVPR*, 427–436.

[24] Snell, J., Swersky, K., & Zemel, R. (2017). Prototypical networks for few-shot learning. *NeurIPS*, 30, 4077–4087.

[25] Vinyals, O., Blundell, C., Lillicrap, T., Wierstra, D., & Kavukcuoglu, K. (2016). Matching networks for one shot learning. *NeurIPS*, 29, 3630–3638.

[26] Xiao, H., Rasul, K., & Vollgraf, R. (2017). Fashion-MNIST: A novel image dataset for benchmarking machine learning algorithms. *arXiv:1708.07747*.

A. Notation Summary

Table 7: Symbol definitions used throughout the paper.

Symbol	Definition	Symbol	Definition
$\mathbf{I} \in [0,1]^{H \times W}$	Grayscale input image	χ	Euler characteristic
$\mathbf{B} \in \{0,1\}^{H \times W}$	Binarized image ($\tau_b = 0.3$)	N_{holes}	Number of topological holes
σ_f, σ_c	Fine (0.6) / coarse (2.5) Gaussian scales	$\mathbf{S}$	Morphological skeleton
$M^{(\sigma)}$, $\theta^{(\sigma)}$	Gradient magnitude and direction at scale σ	$\boldsymbol{\rho}^{(1)}, \ldots, \boldsymbol{\rho}^{(4)}$	Four-space relational density matrices
$\mathcal{P}$	Traced structural path	$\mathcal{R}(i,j)$	Aggregate relational density, class i to j
$\phi(\mathcal{P}) \in \mathbb{R}^{13}$	Per-path geometric feature vector	$\mathbf{v} \in \Delta^{C-1}$	Chain vote distribution
$\mathbf{f}_A \in \mathbb{R}^{72}$	Fine-scale path signature block	Ψ_{SP}	Topology-anchored epistemic confidence
$\mathbf{f}_B \in \mathbb{R}^{39}$	Coarse-scale path signature block	τ_{topo}	Topological consistency score
$\mathbf{f}_C \in \mathbb{R}^{14}$	Topological + skeletal feature block	θ_Ψ	OOD rejection threshold
$\mathbf{f}_D \in \mathbb{R}^{7}$	Path interaction feature block	AUROC_Ψ	Epistemic area under ROC curve
$\mathbf{f} \in \mathbb{R}^{132}$	Complete structural path signature	SPQ	Selective Prediction Quality
$\mathbf{z} \in \mathbb{R}^{K}$	LDA-projected representation ($K = 9$)	$\text{CDR}(n)$	Cross-Distribution Retention at n shots
ρ_{sep}	Inter/intra-class separation ratio in LDA space	ECI	Experience-Compressed Intelligence framework