ENREP: ENERGY-INDUCED REPRESENTATIONS FOR CLASSIFICATION WITHOUT BACKPROPAGATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Modern classification relies heavily on learned representations, raising the question of whether discriminative features must always emerge through data-driven learning. We introduce EnRep, a framework that constructs energy-induced representations from prescribed energy functionals and extracts spectral features. Unlike learned feature extractors, it does not use backpropagation or iterative learning for feature construction, and the resulting features are coupled with a simple closed-form classifier. Using a single prescribed energy functional, we evaluate EnRep on synthetic spatial pattern data and two image benchmarks spanning handwritten digits and colorectal histology, with comparisons against a raw-input and a trained neural baseline. Across the tested tasks, EnRep outperforms raw-input classification and, on the synthetic benchmarks, markedly exceeds the accuracy of the trained neural baseline, while its closed-form readout avoids iterative optimization once the representation is constructed. These results demonstrate that energy structure can provide informative and computationally efficient representations without learning the feature extractor from data. More broadly, these findings suggest that energy can serve not only as an objective to be optimized, but also as a source of representations for classification.

1 INTRODUCTION

A central strength of image classifiers is their ability to construct useful features from data. Convolutional neural networks (CNNs) and related architectures achieve this through trainable filters and nonlinear mappings, with learned representations playing a central role in their success (LeCun et al., 1998; Bengio et al., 2013). Obtaining such features, however, typically relies on backpropagation and iterative gradient-based optimization. For image data with exploitable mathematical structure, learning every useful feature directly from data may not always be necessary: some discriminative information may instead be accessible through quantities derived from a prescribed mathematical description of the input.

This possibility is consistent with a broader line of work exploring alternatives to fully learned feature extractors. Random feature methods use fixed random mappings followed by simple trainable predictors (Rahimi & Recht, 2007), while networks with largely frozen parameters can retain substantial expressive power (Frankle et al., 2021; Burkholz, 2024). Related approaches have employed randomly parameterized dynamical systems with linear readouts (Piatti et al., 2026) and physics-informed neural networks trained without gradient descent (Datar et al., 2026). Together, these studies demonstrate that useful prediction does not always require a fully learned representation. Nevertheless, the fixed structure in such methods is typically determined by random mappings, frozen network parameters, or prescribed dynamics, rather than by structural quantities derived from a prescribed mathematical functional.

A natural source of such structure is an energy functional. In spatially structured systems, energy functionals frequently combine complementary information about local bulk states and spatial variation across interfaces. Phase-field models provide a canonical example, where the Ginzburg–Landau (GL) energy contains both gradient and local potential contributions (Cahn & Hilliard, 1958). These components are closely related to pattern morphology and suggest a representation principle: instead of treating energy only as an objective to be minimized, one may use its internal structure to construct representations for classification.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

Prior work on Turing pattern classification provides useful context for examining this idea. Oh and Lee introduced GL-inspired nonlinear features that emphasize bulk and interfacial structures and showed that these features can improve classification with a single-layer neural network (Oh & Lee, 2020). More recently, Shin et al. studied trainable convolutional maps and showed that increasing the number of convolutional features can substantially improve Turing pattern classification (Shin et al., 2026). These studies represent two complementary directions: mathematically prescribed feature construction and learned convolutional feature extraction. The question considered here is whether the energy functional itself can serve as a source of fixed representations, and whether such representations remain informative beyond the reaction–diffusion systems in which the physical motivation is most direct.

To address this question, we introduce EnRep, a framework that constructs spectral representations from structural quantities induced by a prescribed energy functional and couples them with a closed-form classifier. Its representation map contains no trainable parameters and requires no backpropagation for feature construction; supervision enters only through the closed-form classifier. In this work, we instantiate the general framework with the GL energy and refer to the resulting method as GL-EnRep. We consider two complementary encodings: the GL energy-density mode, obtained by projecting local energy densities onto a spectral basis, and the GL energy-decomposition mode, obtained by decomposing the GL energy into mode-wise contributions.

We evaluate GL-EnRep on three groups of classification problems. First, we construct Turing pattern datasets from the Lengyel–Epstein model and consider both four broad morphology classes and a finer 36-class parameter-configuration task. Second, we study handwritten digit images over a wide range of training set sizes, with particular attention to the low-data regime. Finally, we consider colorectal histology images at 28×28 and 150×150 resolutions. Across all three benchmark groups, we use the same prescribed GL energy functional without training a task-specific feature extractor. Raw-pixel ridge classification provides a matched readout baseline, while a shallow CNN serves as a compact learned-feature reference for accuracy and computational cost.

2 Background and Related Work

2.1 Learned and Fixed Representations

Representation learning typically optimizes feature extractors jointly with downstream predictors (Bengio et al., 2013; Shi et al., 2022), whereas fixed-feature approaches move representation construction outside the training loop. Examples include random features (Rahimi & Recht, 2007), partially frozen networks (Frankle et al., 2021; Burkholz, 2024), prescribed random dynamics (Piatti et al., 2026), and gradient-free physics-informed models (Datar et al., 2026). Deterministic spectral constructions provide another reference point: wavelet scattering uses fixed multiscale filters (Bruna & Mallat, 2013), while recent ICLR work studies Fourier-based representations, learnable multigrid encodings, and geometry-adaptive harmonic representations (Koyama et al., 2024; Audia et al., 2025; Kadkhodaie et al., 2024). EnRep differs by deriving its fixed representation from a prescribed energy functional rather than from random, frozen, or learned spectral structure.

2.2 Energy Structure and Pattern Classification

The GL energy separates bulk and interfacial structure through its potential and gradient terms. GL-inspired quantities have previously been used as engineered features for Turing pattern classification (Oh & Lee, 2020), whereas EnRep constructs spectral representations from the energy terms and uses the same prescribed GL functional across datasets. As a learned reference, Shin et al. reported improved Turing pattern classification as the number of convolutional features increased (Shin et al., 2026). We use their study to motivate six convolution filters and the reference training schedule for our CNN baseline, which is re-trained under the present benchmark protocol.

2.3 Histopathology and Computational Efficiency

Histopathology provides a useful test of whether energy-derived features remain informative on complex real images. The colorectal histology benchmark introduced by Kather et al. contains heterogeneous tissue textures across multiple tissue categories (Kather et al., 2016). Computational

pathology often relies on large collections of patches, and the associated training and storage costs have motivated methods that reduce the amount of data used downstream (Cong et al., 2024). We therefore use colorectal histology to test whether a compact, fixed energy-based representation retains discriminative information without feature-extractor training.

3 Energy-Induced Representations

3.1 EnRep Framework

Let $u \in \mathcal{X}$ denote an input image or scalar field and let $\mathcal{E} : \mathcal{X} \to \mathbb{R}$ be an energy functional. EnRep does not use the scalar value $\mathcal{E}(u)$ alone as a feature. Instead, it treats the structure induced by $\mathcal{E}$ as a source from which a finite-dimensional representation can be constructed. The particular form of this structure depends on the chosen functional and may include local quantities, spatial interactions, variational terms, or other energy-associated descriptors.

We write the EnRep representation as

$$\Phi_{\mathcal{E}} = \mathcal{T} \circ \mathcal{S} \circ \mathcal{Q}_{\mathcal{E}}, \qquad z(u) = \Phi_{\mathcal{E}}(u) \in \mathbb{R}^d,$$

where $\mathcal{Q}_{\mathcal{E}}$ extracts structural quantities induced by the prescribed energy, $\mathcal{S}$ expresses these quantities in a chosen spectral coordinate system, and $\mathcal{T}$ retains or aggregates a finite set of coordinates. This formulation does not require every energy functional to admit the same analytic decomposition. Rather, it requires that the selected functional induce computable quantities that can be mapped deterministically into a fixed-dimensional representation.

The three components play distinct roles. The energy functional determines which structural properties of the input are exposed, the spectral representation determines how those properties are organized across modes, and the truncation or aggregation rule determines the final representation dimension. These choices may reflect prior knowledge about the input domain, including its geometry and boundary conditions.

Once $\mathcal{E}$, $\mathcal{Q}_{\mathcal{E}}$, $\mathcal{S}$, and $\mathcal{T}$ are fixed, $\Phi_{\mathcal{E}}$ contains no trainable parameters and is independent of the class labels. Supervision enters only through the downstream readout fitted to $z(u)$. EnRep therefore separates a structural representation from classifier optimization. The following subsections instantiate this general framework using the Ginzburg–Landau functional.

3.2 GL-EnRep

We instantiate EnRep using the Ginzburg–Landau energy. For exposition, consider

$$\mathcal{E}_\epsilon(u) = \int_\Omega \left[\frac{\epsilon^2}{2} |\nabla u|^2 + \frac{1}{4}(u^2 - 1)^2 \right] d\xi. \tag{1}$$

In computation, we use discrete constructions based on this functional on the normalized image lattice with unit pixel spacing. Throughout the experiments, $\epsilon = 1$. No dataset-specific coefficient in the energy functional is learned or tuned. Figure 1 summarizes the GL-EnRep pipeline and the two representations detailed below.

3.3 GL Energy-Density Mode

The two local energy-density terms are

$$e_g(\xi) = \frac{\epsilon^2}{2} |\nabla u(\xi)|^2, \qquad e_p(\xi) = \frac{1}{4}\big(u(\xi)^2 - 1\big)^2. \tag{2}$$

Let $\{\phi_k\}_{k \geq 0}$ denote an orthonormal Laplacian eigenbasis,

$$\langle \phi_j, \phi_k \rangle = \delta_{jk}, \qquad -\Delta \phi_k = \lambda_k \phi_k.$$

For the derivation below, we consider a real orthonormal Laplacian eigenbasis; the corresponding complex Fourier form is given in Appendix A.2. Projecting the two energy densities gives

$$a_k = \langle e_g, \phi_k \rangle, \qquad b_k = \langle e_p, \phi_k \rangle. \tag{3}$$

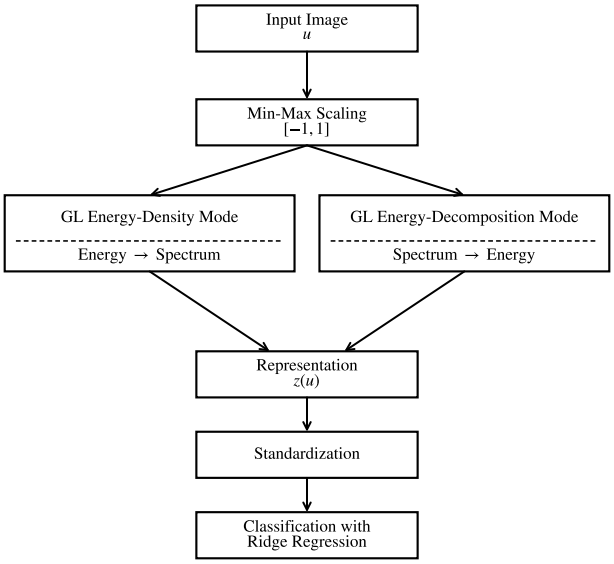

Figure 1: GL-EnRep pipeline. Normalized inputs are mapped to one of two energy-based spectral representations and classified by closed-form ridge. The representation is fixed and uses no back-propagation.

These coefficients summarize how interfacial and bulk energy are distributed across spectral coordinates. For non-periodic image benchmarks, the discrete implementation uses an orthonormal two-dimensional DCT-II basis (Ahmed et al., 1974). For the periodic Turing pattern data, it uses the two-dimensional Fourier basis and aggregates spectral quantities over radial shells.

3.4 GL Energy-Decomposition Mode

A complementary encoding is obtained by decomposing the GL energy itself. With

$$u = \sum_{k \geq 0} c_k \phi_k, \qquad c_k = \langle u, \phi_k \rangle, \qquad r_k = \langle u^3, \phi_k \rangle,$$

and using orthogonality and the Laplacian eigenvalue relation,

$$\int_\Omega |\nabla u|^2 \, d\xi = \sum_{k \geq 0} \lambda_k c_k^2, \qquad \int_\Omega u^2 \, d\xi = \sum_{k \geq 0} c_k^2.$$

Moreover,

$$\int_\Omega u^4 \, d\xi = \langle u, u^3 \rangle = \sum_{k \geq 0} c_k r_k.$$

Substituting these identities into equation 1 gives

$$\mathcal{E}_\epsilon(u) = \frac{|\Omega|}{4} + \sum_{k \geq 0} \left[\frac{1}{2}(\epsilon^2 \lambda_k - 1) c_k^2 + \frac{1}{4} c_k r_k \right]. \tag{4}$$

where Ω denotes the spatial domain and $|\Omega|$ its measure. Therefore, we separate the mode-wise gradient and potential contributions as

$$g_k = \frac{\epsilon^2}{2} \lambda_k c_k^2, \qquad p_k = -\frac{1}{2} c_k^2 + \frac{1}{4} c_k r_k. \tag{5}$$

Unlike the energy-density construction, which first forms local energy maps, the energy-decomposition construction assigns the GL gradient and potential contributions directly to spectral modes.

3.5 Discrete Construction and Representation Dimension

For handwritten digit and colorectal histology images, we retain an $m \times m$ low-frequency DCT block from each of the two GL components and append the image mean. Both GL-EnRep variants therefore have

$$d_{\mathrm{DCT}} = 1 + 2m^2. \tag{6}$$

We use $m = 10$ for digit images, giving $d = 201$, and a common $m = 14$ for both histology resolutions, giving $d = 393$. The latter choice keeps the histology representation dimension fixed while the raw input dimension changes from 784 to 22,500. Mode-count sensitivity results are reported in Appendix C.

For the periodic Turing pattern data, we retain the first 25 radial shells. Each shell contributes a gradient-energy summary, a potential-energy summary, and a directional anisotropy statistic. Together with the image mean,

$$d_{\mathrm{Turing}} = 1 + 3 \times 25 = 76. \tag{7}$$

The radial pooling and anisotropy construction are given in Appendix A.2.

3.6 Closed-Form Ridge Readout

Following ridge regression (Hoerl & Kennard, 1970), let $Z \in \mathbb{R}^{n \times d}$ denote the training features after standardizing each coordinate by its training-set mean and standard deviation, and let $Y \in \mathbb{R}^{n \times C}$ be the one-hot class matrix. The closed-form ridge solution is written as

$$W = (Z^\top Z + \alpha I)^{-1} Z^\top Y, \qquad \alpha = 1. \tag{8}$$

For a standardized test feature vector z, prediction is

$$\widehat{y} = \arg\max_c (zW)_c.$$

The raw-pixel baseline uses the same preprocessing and ridge readout, isolating the effect of the input representation.

4 Experiments

4.1 Common Experimental Protocol

Across all datasets, accuracies are reported over 10 stratified 80:20 train/test splits as mean $\pm$ standard deviation. Each image is independently min-max scaled to $[-1, 1]$. This range aligns the normalized image values with the two preferred states of the GL double-well potential; the same normalization is applied to all baselines. The main GL-EnRep experiments use the ridge readout in equation 8. All reported timings were obtained on an NVIDIA RTX A6000 GPU using CUDA with `torch.float32` including CPU-to-GPU transfer of the already normalized images and labels, feature extraction, standardization, fitting, prediction, and accuracy computation. CPU-side data loading, normalization, split construction, one-time transform-cache initialization and warm-up, plotting, and file I/O are excluded. CNN timing includes all 10,000 training epochs and final test prediction.

4.2 Lengyel–Epstein Model

Turing patterns serve as the physics-aligned benchmark because their morphology naturally reflects the bulk/interface structure emphasized by the GL functional. We generate Turing patterns (Turing, 1952) using the Lengyel–Epstein reaction–diffusion model (Lengyel & Epstein, 1991):

$$\frac{\partial u}{\partial t} = D_u \Delta u + k_1 \left(v - \frac{uv}{1+v^2} \right), \qquad \frac{\partial v}{\partial t} = D_v \Delta v + k_2 - v - \frac{4uv}{1+v^2}. \tag{9}$$

We set $D_u = 1$, $k_2 = 11$, $D_v \in \{0.01, 0.02, 0.03, 0.04, 0.05, 0.06\}$, and $k_1 \in \{1, 3, 5, 7, 9, 11\}$. Simulations use $\Omega = (0, 10)^2$, periodic boundaries, a 64×64 grid, a spectral Laplacian, explicit Euler with $\Delta t = 2^{-9}$, and $T = 512$. Initial conditions are

$$(\bar{u}, \bar{v}) = \left(1 + \frac{k_2^2}{25}, \frac{k_2}{5} \right), \qquad u_0 = \bar{u} + 0.2\xi_u, \quad v_0 = \bar{v} + 0.2\xi_v, \quad \xi_u, \xi_v \sim \mathcal{U}[-1, 1],$$

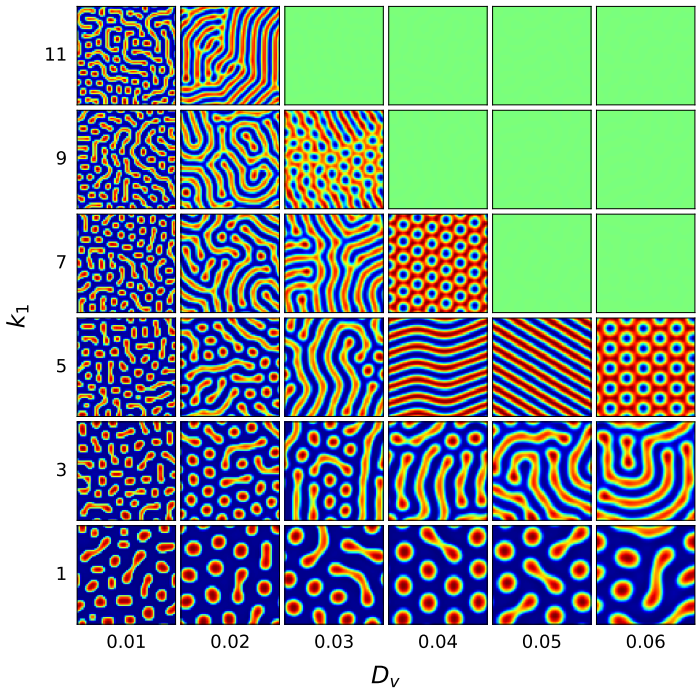

Figure 2: Final v fields over the 36 parameter combinations. Columns vary D_v, rows vary k_1.

and the final v field is used for classification.

Each of the 36 parameter configurations is simulated from 20 random perturbations, yielding 720 images. The 36-class task labels parameter configurations directly. For the 4-class task, each configuration is assigned a representative morphology (small spot, large spot, stripe, inverted spot, stable); stable configurations are removed, leaving 540 images. The full mapping is given in Appendix E.

4.3 Image Benchmarks

Handwritten Digits. We use the MNIST handwritten digits (LeCun et al., 1998) as a nonphysical canonical image benchmark and construct disjoint class-balanced subsets for each run. Unlike the Turing setting, MNIST provides no physical mechanism that directly motivates the GL functional, allowing us to test a prescribed representation after its physical motivation is removed. In the main-text experiments, train/test sizes range from $200/50$ to $2{,}000/500$, maintaining an 80:20 ratio, with larger-scale results reported in Appendix F.1. This experiment is designed to examine how the utility of representations changes as labeled data become more abundant. GL-EnRep uses $m = 10$, corresponding to 201 features, while the raw image contains 784 pixels.

Colorectal histology. We consider two publicly distributed eight-class colorectal histology[1] collections associated with the benchmark introduced by Kather et al. (Kather et al., 2016). Histology serves as a complex real-world texture benchmark, testing whether the prescribed GL representation remains useful beyond canonical images and physics-aligned synthetic morphology. We use grayscale images at resolutions of 28×28 and 150×150, respectively. We treat them as separate benchmark realizations because they are not assumed to be paired resizings of the same images. Both contain 5,000 samples with 625 images per tissue class. GL-EnRep uses the same $m = 14$ truncation in both experiments, giving 393 features.

[1] `https://www.kaggle.com/datasets/kmader/colorectal-histology-mnist`

4.4 Convolutional Neural Networks

The neural baseline is a CNN with one convolutional layer and six 3×3 convolutional filters, ReLU, 2×2 max pooling, flattening, and a linear output layer. Its architecture and training protocol follow Shin et al. (Shin et al., 2026): mini-batch SGD with batch size 32, learning rate 0.004, cross-entropy loss, 10,000 epochs, and $K = 6$ convolutional filters. We retain the 10,000-epoch schedule for consistency with the reference protocol rather than tuning the stopping point. The model is used as a compact learned-feature baseline rather than a state-of-the-art neural upper bound. We evaluate it on both Turing tasks, MNIST at $2{,}000/500$, and 28×28 histology at $4{,}000/1{,}000$, omitting the 150×150 case due to its substantially higher computational cost.

5 Results

5.1 Turing Patterns

Table 1 reports the Turing results. On the 4-class task, both GL modes reach approximately 99% accuracy, compared with approximately 37% for raw-pixel ridge classification and 82% for the CNN. On the 36-class task, Energy-Density mode and Energy-Decomposition mode reach around 73% and 70%, respectively, versus 10% for raw pixels and 23% for the CNN, showing that the representation remains informative for the finer 36-class task.

Table 1: Test accuracy of Turing pattern classification over 10 runs (mean $\pm$ standard deviation).

Dataset	Method	Dim.	Accuracy (%)
Turing 4-class	GL Energy-Density + Ridge	76	98.98 ± 1.72
	GL Energy-Decomposition + Ridge	76	99.63 ± 0.65
	Raw Pixels + Ridge	4096	37.31 ± 6.08
	CNN ($K = 6$)	5766	82.41 ± 3.98
Turing 36-class	GL Energy-Density + Ridge	76	72.78 ± 0.79
	GL Energy-Decomposition + Ridge	76	69.93 ± 1.99
	Raw Pixels + Ridge	4096	10.07 ± 1.24
	CNN ($K = 6$)	5766	22.99 ± 3.71

5.2 Handwritten Digits

Figure 3 shows that the energy-based representations generally outperform raw pixels in the low-data regime, with the gap narrowing as more training data become available. The CNN model reaches about 93% at the $2{,}000/500$ split. Larger-scale results are reported in Appendix F.1, with nonlinear ridge results in Appendix G.

Table 2: Test accuracy of MNIST handwritten digits classification at the 2,000/500 split over 10 runs (mean $\pm$ standard deviation).

Method	Dim.	Accuracy (%)
GL Energy-Density + Ridge	201	83.42 ± 1.24
GL Energy-Decomposition + Ridge	201	81.14 ± 0.99
Raw Pixels + Ridge	784	76.34 ± 1.27
CNN ($K = 6$)	1014	92.82 ± 0.98

5.3 Colorectal Histology

Figures 4 and 5 show the histology results, and Table 3 reports the largest evaluated split. Energy-Decomposition mode achieves the highest accuracy among the evaluated methods, reaching about 56% at 28×28 and 60% at 150×150, compared with approximately 26% for the raw-pixel ridge baseline.

The same 14×14 spectral truncation keeps the GL-EnRep dimension fixed at 393 while the raw dimension changes from 784 to 22,500. Because the two public collections are not assumed to

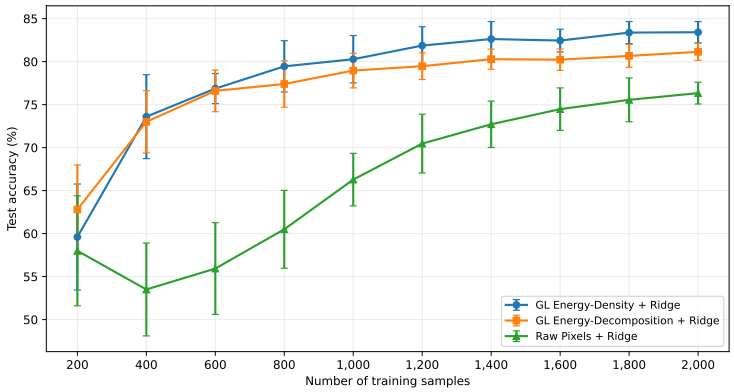

Figure 3: Test accuracy of MNIST handwritten digits classification over 10 runs (mean ± standard deviation) across different training-set sizes.

contain paired resized images, this is not a controlled resizing study; rather, it evaluates the same GL-EnRep design on separate datasets at two resolutions.

Table 3: Test accuracy of colorectal histology classification at the 4,000/1,000 split over 10 runs (mean ± standard deviation).

Resolution	Method	Dim.	Accuracy (%)
28×28	GL Energy-Density + Ridge	393	41.17 ± 1.59
	GL Energy-Decomposition + Ridge	393	56.45 ± 1.32
	Raw Pixels + Ridge	784	25.84 ± 1.68
	CNN ($K = 6$)	1014	49.54 ± 1.42
150×150	GL Energy-Density + Ridge	393	43.98 ± 1.06
	GL Energy-Decomposition + Ridge	393	59.92 ± 1.72
	Raw Pixels + Ridge	22500	25.57 ± 1.62

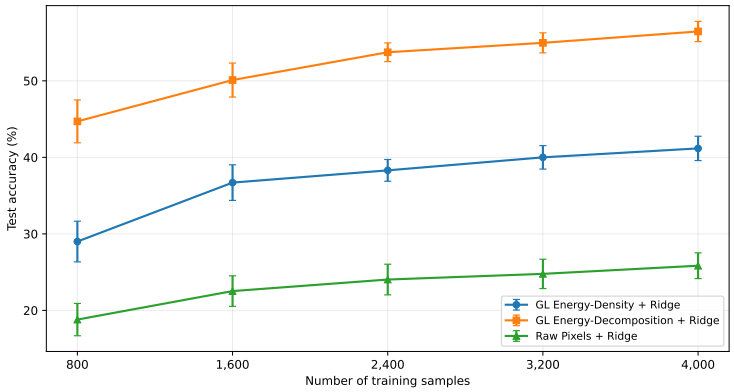

Figure 4: Test accuracy of colorectal histology (28×28) classification over 10 runs (mean ± standard deviation) across different training-set sizes.

These benchmarks span physics-aligned morphology, canonical nonphysical images, and complex real-world texture. Appendix B introduces a raw spectral mode that removes the energy transformation while retaining the corresponding spectral organization. The results show a task-dependent benefit of prescribed energy structure: spectral organization alone is already strong on digit images, whereas energy-derived representations provide additional discriminative structure on Turing patterns and histology.

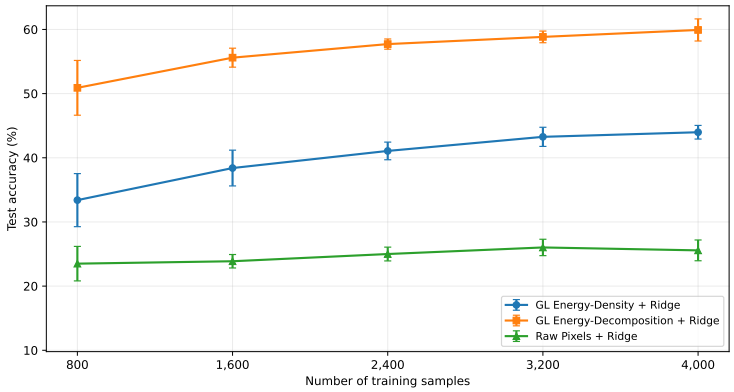

Figure 5: Test accuracy of colorectal histology (150×150) classification over 10 runs (mean $\pm$ standard deviation) across different training-set sizes.

5.4 Computational Cost

Table 4 summarizes wall-clock time under the timing scope defined in Section 4.1. All reported values were measured on the NVIDIA RTX A6000 setup described there. The ridge-based pipelines finish in under 0.13 s across the reported settings, whereas the CNN model requires hundreds to thousands of seconds for its 10,000-epoch iterative training. These measurements reflect different optimization procedures and the specific hardware/software execution environment and should not be interpreted as hardware-independent speedup factors.

Table 4: Wall-clock time in seconds over 10 runs (mean $\pm$ standard deviation)

Dataset	Energy-Density	Energy-Decomposition	Raw Pixels	CNN ($K = 6$)
Turing 4-class	0.0111 ± 0.0002	0.1280 ± 0.0054	0.0040 ± 0.0001	575.7 ± 42.6
Turing 36-class	0.0113 ± 0.0002	0.1202 ± 0.0016	0.0048 ± 0.0001	727.0 ± 54.5
MNIST	0.0034 ± 0.0002	0.0031 ± 0.0002	0.0044 ± 0.0000	2329.4 ± 75.8
Histology 28×28	0.0065 ± 0.0001	0.0063 ± 0.0002	0.0062 ± 0.0002	4598.9 ± 113.7
Histology 150×150	0.0862 ± 0.0003	0.0776 ± 0.0006	0.1111 ± 0.0008	–

6 Conclusion

We introduced EnRep as a framework for constructing representations from energy structure. Using the GL functional, GL-EnRep yields two complementary spectral encodings based on local energy densities and mode-wise energy decomposition. The representation map contains no trainable parameters and requires no backpropagation, while classification is performed by a closed-form ridge readout.

The results show that these energy-derived representations are highly effective for Turing pattern classification, remain informative on MNIST where raw spectral organization is itself strong, and remain informative on two colorectal histology benchmarks at different resolutions. Together, the experiments support the view that an energy functional can serve not only as an optimization objective, but also as a source of compact representations for classification.

AI USE STATEMENT

In this work, we used generative AI tools to assist with feedback on research methodology and experimental design, software implementation and debugging, interpretation and presentation of experimental results, and manuscript drafting and editing. We also used generative AI tools to assist with literature search and summarization, citation-consistency checks, and improvements to the clarity and organization of the manuscript. The synthetic datasets and numerical results reported in this paper were produced by the authors' simulation and experimental pipelines rather than generated directly by an AI model. All AI-assisted text, code, mathematical expressions, and research-related suggestions were reviewed and verified by the authors before inclusion. The authors take full responsibility for the final content of this work, including all claims, results, and artifacts produced with the aid of generative AI.

REFERENCES

Nasir Ahmed, T. Raj Natarajan, and K. R. Rao. Discrete cosine transform. *IEEE Transactions on Computers*, C-23(1):90–93, 1974. doi: 10.1109/T-C.1974.223784.

Samuel Audia, Soheil Feizi, Matthias Zwicker, and Dinesh Manocha. How learnable grids recover fine detail in low dimensions: A neural tangent kernel analysis of multigrid parametric encodings. In *International Conference on Learning Representations*, 2025.

Yoshua Bengio, Aaron Courville, and Pascal Vincent. Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1798–1828, 2013. doi: 10.1109/TPAMI.2013.50.

Joan Bruna and Stéphane Mallat. Invariant scattering convolution networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8):1872–1886, 2013. doi: 10.1109/TPAMI.2012.230.

Rebekka Burkholz. Batch normalization is sufficient for universal function approximation in CNNs. In *International Conference on Learning Representations*, 2024.

John W. Cahn and John E. Hilliard. Free energy of a nonuniform system. I. Interfacial free energy. *The Journal of Chemical Physics*, 28(2):258–267, 1958. doi: 10.1063/1.1744102.

Cong Cong, Shiyu Xuan, Sidong Liu, Maurice Pagnucco, Shiliang Zhang, and Yang Song. Dataset distillation for histopathology image classification, 2024.

Chinmay Datar, Taniya Kapoor, Abhishek Chandra, Qing Sun, Erik L. Bolager, Iryna Burak, Anna Veselovska, Massimo Fornasier, and Felix Dietrich. Fast training of accurate physics-informed neural networks without gradient descent. In *International Conference on Learning Representations*, 2026.

Jonathan Frankle, David J. Schwab, and Ari S. Morcos. Training batchnorm and only batchnorm: On the expressive power of random features in CNNs. In *International Conference on Learning Representations*, 2021.

Arthur E. Hoerl and Robert W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970. doi: 10.1080/00401706.1970.10488634.

Zahra Kadkhodaie, Florentin Guth, Eero Simoncelli, and Stéphane Mallat. Generalization in diffusion models arises from geometry-adaptive harmonic representations. In *International Conference on Learning Representations*, 2024.

Jakob Nikolas Kather, Cleo-Aron Weis, Francesco Bianconi, Susanne M. Melchers, Lothar R. Schad, Timo Gaiser, Alexander Marx, and Frank Gerrit Zöllner. Multi-class texture analysis in colorectal cancer histology. *Scientific Reports*, 6:27988, 2016. doi: 10.1038/srep27988.

Masanori Koyama, Kenji Fukumizu, Kohei Hayashi, and Takeru Miyato. Neural Fourier transform: A general approach to equivariant representation learning. In *International Conference on Learning Representations*, 2024.

Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998. doi: 10.1109/5.726791.

István Lengyel and Irving R. Epstein. Modeling of Turing structures in the Chlorite–Iodide–Malonic Acid–Starch reaction system. *Science*, 251(4994):650–652, 1991. doi: 10.1126/science.251.4994.650.

Seoyoung Oh and Seunggyu Lee. Extracting insights of classification for Turing pattern with feature engineering. *Journal of the Korean Society for Industrial and Applied Mathematics*, 24(3):321–330, 2020. doi: 10.12941/jksiam.2020.24.321.

Francesco Piatti, Thomas Cass, and William F. Turner. Random controlled differential equations. In *International Conference on Learning Representations*, 2026.

Ali Rahimi and Benjamin Recht. Random features for large-scale kernel machines. In *Advances in Neural Information Processing Systems*, volume 20, pp. 1177–1184, 2007.

Zhenmei Shi, Junyi Wei, and Yingyu Liang. A theoretical analysis on feature learning in neural networks: Emergence from inputs and advantage over fixed features. In *International Conference on Learning Representations*, 2022.

Jaemin Shin, Junyoung Park, Minhwan Ji, and Seunggyu Lee. Exploring potential of Turing pattern classification through convolution maps. *Scientific Reports*, 16:3008, 2026. doi: 10.1038/s41598-025-32911-0.

Alan M. Turing. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 237(641):37–72, 1952. doi: 10.1098/rstb.1952.0012.

A Discrete GL-EnRep Construction

A.1 DCT Construction for Non-Periodic Images

For an $N \times N$ image, let $C = \mathrm{DCT}(u)$ and $R = \mathrm{DCT}(u^3)$ denote the orthonormal two-dimensional DCT-II coefficients. The discrete Laplacian eigenvalues used in the implementation are

$$\lambda_{pq} = 4\sin^2\Big(\frac{\pi p}{2N}\Big) + 4\sin^2\Big(\frac{\pi q}{2N}\Big), \qquad 0 \le p, q < N.$$

The energy-decomposition arrays are

$$D_{pq} = \frac{1}{2}\lambda_{pq}C_{pq}^2, \qquad P_{pq} = -\frac{1}{2}C_{pq}^2 + \frac{1}{4}C_{pq}R_{pq}.$$

For the energy-density construction, finite differences produce

$$e_g = \frac{1}{2}|\nabla u|^2, \qquad e_p = \frac{1}{4}(u^2-1)^2,$$

followed by orthonormal DCT-II transforms $A_g = \mathrm{DCT}(e_g)$ and $A_p = \mathrm{DCT}(e_p)$. Centered differences are used in the interior and first-order one-sided differences at the image boundary. The final vectors retain the low-frequency $m \times m$ block of the two arrays and append the image mean:

$$z_{\mathrm{dec}} = [\bar{u}, \mathrm{vec}(D_{0:m-1,0:m-1}), \mathrm{vec}(P_{0:m-1,0:m-1})],$$
$$z_{\mathrm{dens}} = [\bar{u}, \mathrm{vec}(A_{g,0:m-1,0:m-1}), \mathrm{vec}(A_{p,0:m-1,0:m-1})].$$

Thus both have dimension $1 + 2m^2$.

A.2 Periodic Fourier Aggregation for Turing Patterns

For periodic Turing patterns, let c_k and r_k denote the orthonormal Fourier coefficients of u and u^3, respectively. For a Fourier mode $k = (k_x, k_y)$, where k_x and k_y denote the corresponding signed frequency indices on an $N \times N$ periodic grid, the discrete Laplacian eigenvalue used in the implementation is

$$\lambda_k = 4\sin^2\left(\frac{\pi k_x}{N}\right) + 4\sin^2\left(\frac{\pi k_y}{N}\right).$$

Because these coefficients are generally complex, the energy-decomposition terms are

$$g_k = \frac{1}{2}\lambda_k|c_k|^2, \qquad p_k = -\frac{1}{2}|c_k|^2 + \frac{1}{4}\,\mathrm{Re}(\overline{c_k}r_k)\,.$$

We retain the first 25 nearest-integer radial shells,

$$S_j = \left\{k : \left\lfloor |k| + \frac{1}{2} \right\rfloor = j\right\}, \qquad j = 0, \ldots, 24,$$

and aggregate

$$G_j = \sum_{k \in S_j} g_k, \qquad P_j = \sum_{k \in S_j} p_k.$$

For energy-density, we Fourier transform the two local energy maps and aggregate spectral power,

$$G_j = \sum_{k \in S_j} |a_k|^2, \qquad P_j = \sum_{k \in S_j} |b_k|^2.$$

Directional information is summarized from the gradient-energy component by the second angular moment

$$C_j = \sum_{k \in S_j} w_k \cos(2\theta_k), \qquad H_j = \sum_{k \in S_j} w_k \sin(2\theta_k),$$

$$A_j = \frac{\sqrt{C_j^2 + H_j^2}}{\sum_{k \in S_j} w_k + 10^{-12}},$$

where $w_k = |g_k|$ for energy-decomposition and $w_k = |a_k|^2$ for energy-density. Modes outside the retained shells are discarded. The final Turing representation is

$$z(u) = [\bar{u}, G_0, P_0, A_0, \ldots, G_{24}, P_{24}, A_{24}] \in \mathbb{R}^{76}.$$

B RAW SPECTRAL MODE

To assess how much discriminative information is already available from spectral organization alone, we evaluate a raw spectral mode using the same spectral basis and cutoff as the corresponding GL-EnRep construction but without forming any energy-derived quantity.

For MNIST and colorectal histology, let $C = \mathrm{DCT}(u)$ be the orthonormal two-dimensional DCT-II of the normalized image. The raw spectral representation retains the same cutoff,

$$z_{\text{raw-spec}} = \mathrm{vec}(C_{0:m-1,0:m-1}) \,.$$

The DC coefficient C_{00} is retained, so no separate image mean is appended. This gives 100 features for MNIST ($m = 10$) and 196 features for both histology resolutions ($m = 14$).

For periodic Turing patterns, let c_k denote the orthonormal Fourier coefficients of u and use the same radial shells S_j defined in Appendix A.2. We aggregate raw spectral power as

$$R_j = \sum_{k \in S_j} |c_k|^2, \qquad j = 0, \ldots, 24,$$

and use

$$z_{\text{raw-spec}} = [R_0, R_1, \ldots, R_{24}] \in \mathbb{R}^{25}.$$

No separate image mean is appended.

Table 5 reports representative linear ridge results. Raw spectral organization is already strong on MNIST, particularly in the low-data regime. In contrast, the energy-derived representations provide clear gains beyond the raw spectrum on the Turing tasks and on both histology benchmarks, supporting the task-dependent interpretation in the main text.

Table 5: Representative results of the raw spectral mode over 10 runs (mean $\pm$ standard deviation accuracy, %).

Dataset	Train / Test	Raw Spectral	Energy-Density	Energy-Decomposition	Raw Pixels
Turing 4-class	432 / 108	96.67 ± 1.81	98.98 ± 1.72	99.63 ± 0.65	37.31 ± 6.08
Turing 36-class	576 / 144	56.04 ± 1.91	72.78 ± 0.79	69.93 ± 1.99	10.07 ± 1.24
MNIST	200 / 50	74.60 ± 5.25	59.60 ± 6.17	62.80 ± 5.18	58.00 ± 6.39
	2,000 / 500	84.32 ± 1.43	83.42 ± 1.24	81.14 ± 0.99	76.34 ± 1.27
	16,000 / 4,000	85.37 ± 0.43	85.48 ± 0.24	82.51 ± 0.58	84.66 ± 0.69
	48,000 / 12,000	85.35 ± 0.24	85.65 ± 0.22	82.56 ± 0.46	85.18 ± 0.31
Histology 28×28	4,000 / 1,000	25.59 ± 1.28	41.17 ± 1.59	56.45 ± 1.32	25.84 ± 1.68
Histology 150×150	4,000 / 1,000	27.46 ± 1.24	43.98 ± 1.06	59.92 ± 1.72	25.57 ± 1.62

C MODE-COUNT SENSITIVITY

The sensitivity experiments use $2{,}000/500$ for MNIST and $4{,}000/1{,}000$ for the two histology benchmarks. MNIST is evaluated over $m = 4, 6, \ldots, 28$. The 28×28 histology benchmark uses the same range, while the 150×150 benchmark additionally evaluates $m = 32, 40, 48, 64, 80, 100, 120, 150$. The same sweeps include raw spectral mode using the corresponding DCT cutoff m. The main experiments use $m = 10$ for MNIST and a common $m = 14$ for both histology resolutions. These values provide compact representations with strong performance in the corresponding sensitivity ranges; they are not presented as validation-selected optima.

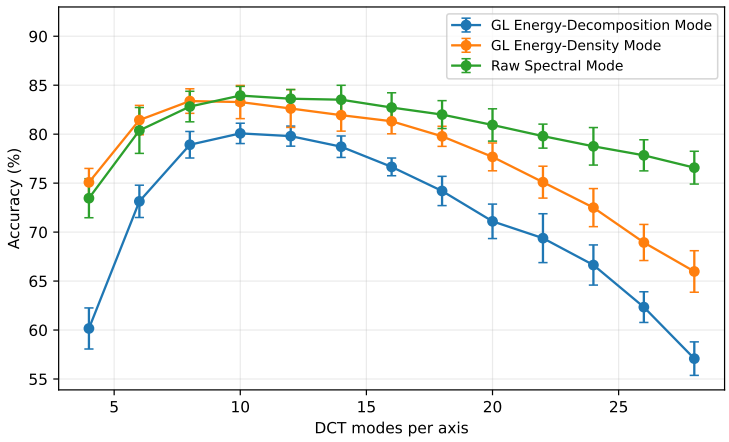

Figure 6: Mode-count sensitivity of the three fixed representations on MNIST at the $2{,}000/500$ split. Test accuracy is shown as a function of the number of retained DCT modes per axis; points and error bars denote the mean $\pm$ standard deviation over 10 runs.

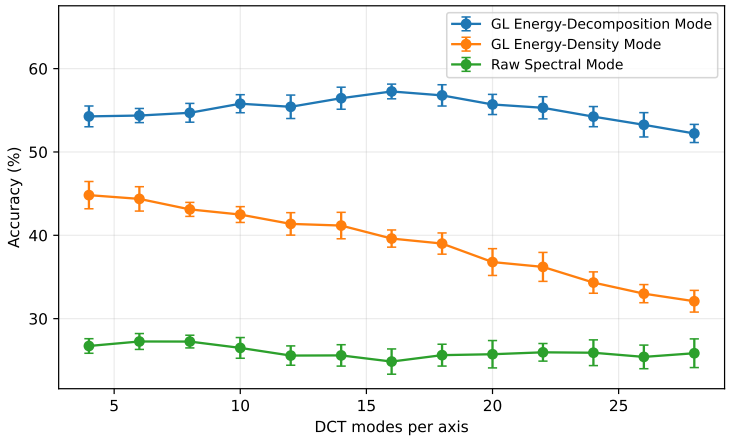

Figure 7: Mode-count sensitivity of the three fixed representations on the colorectal histology dataset (28×28) at the $4{,}000/1{,}000$ split. Test accuracy is shown as a function of the number of retained DCT modes per axis; points and error bars denote the mean $\pm$ standard deviation over 10 runs.

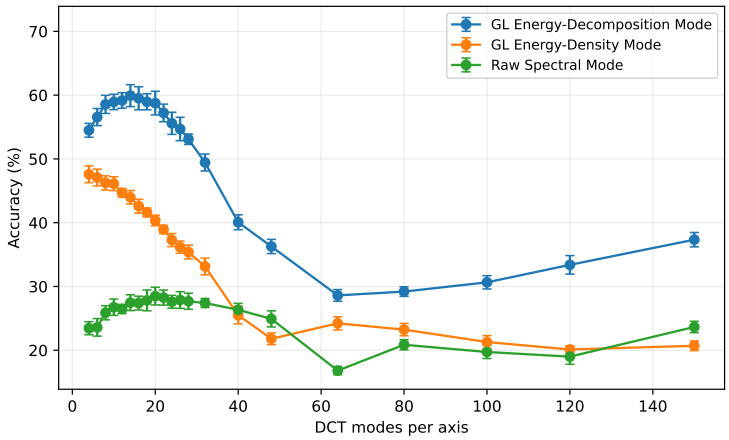

Figure 8: Mode-count sensitivity of the three fixed representations on the colorectal histology dataset (150×150) at the $4{,}000/1{,}000$ split. Test accuracy is shown as a function of the number of retained DCT modes per axis; points and error bars denote the mean $\pm$ standard deviation over 10 runs.

D Radial Spectral Contributions of Turing Patterns

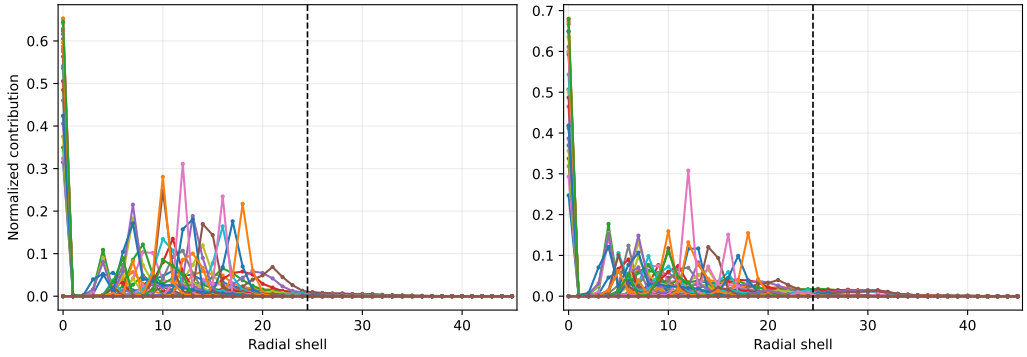

Figure 9: Normalized radial spectral contributions of the GL representation for the Turing 36-class dataset. Each curve represents the mean radial contribution for one (D_v, k_1) parameter configuration over its realizations. The left and right panels show the gradient-energy and potential-energy components, respectively. The dashed line marks the cutoff; shells $j = 0, \ldots, 24$ are retained.

The radial contributions are concentrated predominantly in the low-order shells. We therefore retain the first 25 shells as a compact spectral truncation rather than as a validation-selected optimum.

E Label Construction of Turing Patterns

The parameter-pair morphology assignments used in the 4-class task are shown in Table 6. The four retained numeric labels are $0 =$ small spot, $1 =$ inverted spot, $2 =$ large spot, and $3 =$ stripe. All 20 realizations generated from a given parameter pair inherit the same morphology label. Morphology labels are assigned by visual inspection of the representative final patterns for each parameter configuration.

Table 6: Morphology labels assigned to the Turing 4-class parameter configurations; dashes indicate spatially uniform configurations (stable states) excluded from the benchmark.

$k_1 \backslash D_v$	0.01	0.02	0.03	0.04	0.05	0.06
11	0	3	–	–	–	–
9	0	3	3	–	–	–
7	0	3	3	1	–	–
5	0	0	3	3	3	1
3	0	0	2	2	3	3
1	2	2	2	2	2	2

F Full Results

F.1 Handwritten Digits

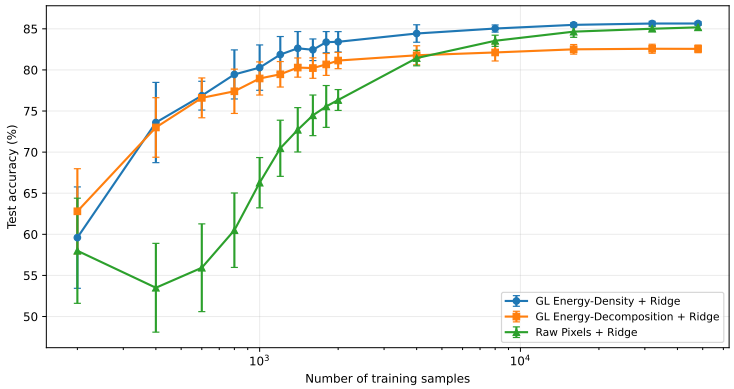

Figure 10: Test accuracy of full MNIST handwritten digits classification over 10 runs (mean ± standard deviation) across different training-set sizes.

Table 7: Test accuracy of MNIST handwritten digits classification in all splits over 10 runs (mean ± standard deviation, %).

Train / Test	Energy-Density	Energy-Decomposition	Raw Pixels
200 / 50	59.60 ± 6.17	62.80 ± 5.18	58.00 ± 6.39
400 / 100	73.60 ± 4.88	73.00 ± 3.62	53.50 ± 5.40
600 / 150	76.87 ± 1.75	76.60 ± 2.42	55.93 ± 5.34
800 / 200	79.45 ± 2.99	77.40 ± 2.70	60.50 ± 4.53
1,000 / 250	80.28 ± 2.76	78.96 ± 2.02	66.28 ± 3.06
1,200 / 300	81.87 ± 2.20	79.47 ± 1.55	70.47 ± 3.41
1,400 / 350	82.63 ± 2.04	80.29 ± 1.17	72.71 ± 2.70
1,600 / 400	82.45 ± 1.32	80.23 ± 1.24	74.48 ± 2.47
1,800 / 450	83.38 ± 1.29	80.67 ± 1.33	75.56 ± 2.54
2,000 / 500	83.42 ± 1.24	81.14 ± 0.99	76.34 ± 1.27
4,000 / 1,000	84.43 ± 1.06	81.78 ± 1.16	81.44 ± 0.94
8,000 / 2,000	85.04 ± 0.43	82.12 ± 1.04	83.55 ± 0.68
16,000 / 4,000	85.48 ± 0.24	82.51 ± 0.58	84.66 ± 0.69
32,000 / 8,000	85.64 ± 0.27	82.57 ± 0.55	85.01 ± 0.27
48,000 / 12,000	85.65 ± 0.22	82.56 ± 0.46	85.18 ± 0.31

F.2 COLORECTAL HISTOLOGY

Table 8: Test accuracy of colorectal histology (28×28) classification in all splits over 10 runs (mean $\pm$ standard deviation, %).

Train / Test	Energy-Density	Energy-Decomposition	Raw Pixels
800 / 200	29.00 ± 2.66	44.70 ± 2.80	18.80 ± 2.11
1,600 / 400	36.70 ± 2.33	50.10 ± 2.22	22.53 ± 1.99
2,400 / 600	38.30 ± 1.42	53.73 ± 1.22	24.03 ± 2.00
3,200 / 800	40.00 ± 1.53	54.96 ± 1.31	24.77 ± 1.91
4,000 / 1,000	41.17 ± 1.59	56.45 ± 1.32	25.84 ± 1.68

Table 9: Test accuracy of colorectal histology (150×150) classification in all splits over 10 runs (mean $\pm$ standard deviation, %).

Train / Test	Energy-Density	Energy-Decomposition	Raw Pixels
800 / 200	33.40 ± 4.14	50.90 ± 4.27	23.50 ± 2.69
1,600 / 400	38.40 ± 2.79	55.60 ± 1.48	23.88 ± 1.05
2,400 / 600	41.07 ± 1.37	57.72 ± 0.80	25.00 ± 1.07
3,200 / 800	43.26 ± 1.49	58.84 ± 0.91	26.02 ± 1.28
4,000 / 1,000	43.98 ± 1.06	59.92 ± 1.72	25.57 ± 1.62

G NONLINEAR RIDGE READOUT

The main experiments use a linear ridge readout to test how directly linearly discriminative the prescribed representation is. As a supplementary experiment, we replace the linear readout by an RBF kernel ridge classifier with $\alpha = 1$ and

$$K(z_i, z_j) = \exp\left(-\frac{\|z_i - z_j\|_2^2}{2\sigma^2}\right), \qquad \sigma = \mathrm{median}_{i<j} \|z_i - z_j\|_2,$$

after training-feature standardization. No validation-based bandwidth search is used. For large training sets, the median-distance heuristic is estimated from up to 200,000 reproducibly sampled training pairs, while the full RBF kernel matrix and kernel-ridge solve are retained. Representative MNIST results are shown in Table 10.

Table 10: Representative MNIST RBF kernel ridge results over 10 runs (mean $\pm$ standard deviation, %).

Train / Test	Energy-Density	Energy-Decomposition	Raw Pixels
200 / 50	80.00 ± 3.65	69.20 ± 5.01	78.80 ± 5.35
2,000 / 500	88.92 ± 1.30	84.74 ± 1.66	87.72 ± 1.14
16,000 / 4,000	93.82 ± 0.47	91.01 ± 0.45	93.48 ± 0.51