PRESS: PHYSICS-REGULARIZED PARAMETER ESTIMATION FROM STEADY-STATE TURING PATTERNS

Zi-Yan Shi*, **Jaron Yeh***, **Chien-Yu Hung***
Department of Engineering Science and Ocean Engineering, National Taiwan University
pegasus2737@gmail.com, yeh.jaron@gmail.com, hunjasper9@gmail.com
*Equal contribution.

ABSTRACT

The Turing mechanism explains how biological systems leverage spontaneous symmetry breaking within reaction-diffusion processes to generate complex spatial patterns. Despite the widespread recognition of this mechanism, accurately recovering the underlying control parameters solely from these static patterns remains a major bottleneck in inverse problem research. This study challenges this limitation by proposing **PRESS** (Physics-Regularized Estimation from Steady-State Turing Patterns), an efficient parameter discovery framework that integrates a novel steady-state physics-informed residual loss to constrain neural network architectures. Unlike traditional approaches that rely on temporal derivatives, our key contribution is a regularization strategy that leverages the equilibrium condition of reaction-diffusion systems, enabling robust inference directly from single snapshots. We systematically evaluate the impact of this physical constraint on both CNNs and Multilayer Perceptrons (MLPs) for the simultaneous inference of four governing parameters. Experimental results demonstrate that integrating the steady-state physics loss significantly enhances model performance. Crucially, PRESS is capable of accurately extracting parameters directly from raw pattern images, substantially outperforming unconstrained baseline models. Our findings confirm that steady-state physical laws serve as a potent regularization tool, providing an efficient, end-to-end, and physically consistent solution for quantifying Turing systems.

1 INTRODUCTION

The spontaneous emergence of spatial patterns in nature—from zebra stripes to the arrangement of hair follicles—has fascinated scientists for decades. Since Alan Turing's seminal work in 1952 Turing (1952), reaction–diffusion systems have served as a foundational mathematical framework for understanding morphogenesis. While the *forward problem*—generating patterns from known dynamical parameters—is now well understood through numerical simulation, the *inverse problem* of inferring governing parameters solely from an observed, static pattern remains a formidable challenge. Solving this inverse problem is crucial for quantitatively characterizing biological models and uncovering the mechanistic principles underlying developmental processes.

In practice, the inverse problem faces two major obstacles. First, biological experiments are costly and time-consuming, resulting in an acute scarcity of labeled data with known ground-truth parameters. Second, most experimental observations are static snapshots lacking temporal resolution. This poses a fundamental difficulty for standard physics-informed approaches, which typically rely on time-series data to compute temporal derivatives ($\partial u/\partial t$) for residual minimization. Without temporal information, traditional deep learning approaches fall back to purely data-driven mapping. However, existing literature—notably Schnörr & Schnörr (2023)—reports that convolutional neural networks (CNNs) perform poorly in such small-data regimes compared to kernel-based methods, suggesting that CNNs inherently require large datasets to extract meaningful representations from Turing patterns.

In this work, we challenge this prevailing assumption by demonstrating that the perceived inefficacy of CNNs in small-data regimes is a result of insufficient regularization, rather than architectural limitations. We introduce **PRESS** (Physics-Regularized Estimation from Steady-State Turing Pat-

terns), a framework that bridges the gap between data-driven learning and mathematical modeling. Instead of requiring temporal trajectories, PRESS leverages the steady-state equilibrium condition of reaction-diffusion systems (where reaction kinetics balance diffusion) as a physics-informed residual loss. This effectively constrains the hypothesis space of the neural network, allowing it to infer parameters that are not only visually consistent with the input pattern but also physically valid according to the governing Partial Differential Equations (PDEs).

In summary, our main contributions are as follows:

- We formulate a steady-state physics-informed loss specifically designed for snapshot data, eliminating the need for temporal derivatives or time-series observations in parameter inference.
- We propose PRESS, a unified framework that integrates this physical regularization with convolutional architectures. We demonstrate that imposing steady-state PDE constraints significantly boosts parameter discovery performance in data-scarce regimes.
- We provide a systematic analysis of the interaction between physical regularization, dataset size, and network depth, confirming that physics-informed constraints effectively compensate for the lack of massive labeled datasets.

2 Related Work

Recovering reaction–diffusion parameters from observed Turing patterns has been approached through four major methodological directions.

Handcrafted Feature Methods. A prominent line of work, exemplified by Schnörr & Schnörr (2023), relies on geometric descriptors to summarize Turing pattern morphology. These hand-engineered features demonstrate that steady-state patterns inherently contain recoverable information about underlying parameters. However, they depend heavily on manual design choices and often generalize poorly across different pattern regimes or stochastic initial conditions. In contrast, our method operates end-to-end on raw images, extracting physically meaningful representations without requiring manually crafted preprocessing or domain-specific feature construction.

Data-Driven CNN Approaches. The comparative study by Schnörr & Schnörr (2023) also investigated convolutional neural networks as a data-driven alternative, though their experiments were limited to single-parameter estimation. Subsequent research has expanded this scope; for instance, Shin et al. (2026) demonstrated the potential of CNNs in classifying spatial heterogeneities, while Scholz & Scholz (2022) utilized CNNs to efficiently scan high-dimensional parameter spaces for complex morphologies. Furthermore, Li et al. (2020) developed an encoder-decoder CNN to directly predict concentration distributions, bypassing expensive numerical simulations. While these models successfully exploit local spatial structure, their predictions lack explicit enforcement of the governing PDEs. This absence of physical regularization often leads to limited consistency and reduced robustness in low-data settings, a limitation we address by directly embedding the steady-state equilibrium condition into the loss landscape.

Physics-Constrained Convolutional Architectures. Integrating CNNs with physical constraints has become a dominant paradigm for solving PDEs on grid-structured domains. Existing literature extensively utilizes PI-CNNs as efficient surrogate models for *forward problems*—mapping known parameters to solution fields. Applications range from spatiotemporal dynamics in reaction-diffusion systems Yuan et al. (2024) and transient Darcy flows Zhang (2022), to electromagnetic simulations Emekci (2024) and steady-state thermal field prediction Zhao et al. (2023). To enhance learning under data scarcity, Rao et al. (2023) proposed a framework unifying measurement data with PDE priors to learn spatiotemporal evolution. A unifying theme is the use of discretized residual losses (e.g., Finite Difference) computed via fixed convolutional kernels. While some studies have extended this to real-time optimization Ge & Khazaei (2024) and state estimation Kelshaw & Magri (2024), a critical distinction exists between reconstructing a physical field (*state estimation*) and inferring governing coefficients (*parameter discovery*). Standard approaches for the latter typically rely on computationally expensive per-instance optimization, lacking a framework designed specifically for amortized parameter inference from static snapshots.

Physics-Informed Optimization Methods. Physics-Informed Neural Networks (PINNs) and PDE-constrained optimization approaches incorporate reaction–diffusion equations directly into the learning objective. PINNs have shown strong performance in forward simulation for classical Turing models Hariri et al. (2025). Extensions to inverse problems, such as the formulation by Kho et al. (2022), jointly optimize parameters and network weights to satisfy steady-state PDE constraints. However, these approaches function as *per-instance solvers*: they require solving a full optimization problem for each individual observed pattern, leading to substantial computational costs. RBF-enhanced PINN variants Matas-Gil & Endres (2024) share this limitation. Unlike these optimization-based methods, our approach learns a generalized mapping function, allowing for instant inference on new data without re-training.

Positioning of PRESS. Our framework, PRESS, bridges these paradigms to address the steady-state inverse problem. Unlike handcrafted descriptors, PRESS requires no manual feature engineering. Unlike purely data-driven CNNs, it enforces explicit physical consistency through PDE residuals. Unlike classical PINN-based inverse solvers which perform per-instance optimization, PRESS amortizes the inference cost by training a unified encoder. Crucially, unlike methods requiring time-series data to compute temporal derivatives, we formulate a physics loss specifically derived from the *steady-state equilibrium* of the reaction-diffusion system. The result is a lightweight, physically regularized method that achieves efficient, end-to-end parameter inference.

3 METHODOLOGY

3.1 THE CHALLENGE OF STATIC SNAPSHOT DATA

A primary obstacle in quantifying reaction–diffusion systems is the reliance on static experimental observations. Unlike time-series data where temporal derivatives ($\partial u/\partial t$) can be directly approximated to guide physics-informed learning, biological experiments often yield only *steady-state snapshots*—single images of the final pattern. Furthermore, large-scale public datasets containing Turing patterns paired with their ground-truth kinetic parameters are virtually non-existent, and generating such labeled data via wet-lab experiments is cost-prohibitive. To overcome these bottlenecks, we employ a physics-based synthetic data generation strategy to construct a comprehensive dataset for training our PRESS framework.

3.2 DATA GENERATION AND NUMERICAL SIMULATION

To construct a high-fidelity dataset of Turing patterns for supervised learning, we implement a numerical pipeline based on the Gierer–Meinhardt reaction–diffusion model. Our mathematical formulation and numerical integration follow the framework of Schnörr & Schnörr (2023), ensuring consistency with established theoretical benchmarks.

3.2.1 MATHEMATICAL FORMULATION

The Gierer-Meinhardt Model Gierer & Meinhardt (1972) The dynamics of the activator concentration $u(x,t)$ and inhibitor concentration $v(x,t)$ are governed by the following system of coupled reaction-diffusion equations on a two-dimensional domain Ω:

$$\frac{\partial u}{\partial t} = D_u \nabla^2 u + \rho \left(\frac{u^2}{v(1+\kappa u^2)} - \mu_u u \right) + \rho_u \tag{1}$$

$$\frac{\partial v}{\partial t} = D_v \nabla^2 v + \rho \left(u^2 - \mu_v v \right) + \rho_v \tag{2}$$

where D_u and D_v denote the diffusion coefficients for the activator and inhibitor, respectively. The reaction terms include a source production rate ρ, decay rates μ_u and μ_v, and basic production terms ρ_u and ρ_v. The saturation constant κ regulates the activator's autocatalytic growth.

Dimensionless Formulation Following the non-dimensionalization procedure described in Murray (1982), we rescale the system to reduce the parameter space to four key dimensionless quantities: the decay rate of the activator a, the basic production rate of the activator b, the saturation constant c, and

the ratio of diffusion coefficients δ. A global scaling parameter s is also introduced to control the effective diffusion magnitude.

The resulting dimensionless system is given by:

$$\frac{\partial u}{\partial t} = s\Delta u + \frac{u^2}{v(1+cu^2)} - au + b \tag{3}$$

$$\frac{\partial v}{\partial t} = s\delta\Delta v + u^2 - v \tag{4}$$

where $\Delta = \nabla^2$ represents the Laplacian operator. In this formulation:

- s is the fixed baseline diffusion coefficient.
- $\delta = D_v/D_u$ represents the relative diffusion scale of the inhibitor.
- The parameter set $\theta = \{a, b, c, \delta\}$ fully determines the pattern selection regime.

3.2.2 NUMERICAL IMPLEMENTATION

Parameter Space Sampling and Stability Filtering To generate a diverse dataset, we sample the kinetic parameters (a, b, c) and the diffusion ratio δ uniformly from the empirical ranges defined in Schnörr & Schnörr (2023), as summarized in Table 1.

Since not all parameter combinations lead to stable Turing patterns, we apply a linear stability analysis (LSA) filter prior to simulation. For each parameter set θ, we compute the homogeneous steady state (u^*, v^*) and verify the Turing instability conditions (existence of positive steady states, stability to homogeneous perturbations, and instability to spatial perturbations). This filtering process ensures that computational resources are focused solely on pattern-forming regimes.

Table 1: Parameter ranges for data generation, adopted from Schnörr & Schnörr (2023).

Parameter	**Symbol**	**Range**
Decay rate of activator	a	$[0.01, 0.7]$
Basic production of activator	b	$[0.4, 2.0]$
Saturation constant	c	$[0.02, 7.0]$
Diffusion scale factor	δ	$[20.0, 200.0]$

Spectral Numerical Simulation We solve the system under periodic boundary conditions using a semi-implicit spectral method, as utilized in Schnörr & Schnörr (2023). This approach employs the Fast Fourier Transform (FFT) to diagonalize the Laplacian operator, treating the stiff diffusion terms implicitly in the frequency domain while evaluating the nonlinear reaction terms explicitly in the spatial domain.

Let $\hat{u}_k^{(n)}$ and $\hat{v}_k^{(n)}$ denote the Fourier coefficients of u and v at time step n for wave vector k. The explicit update rules used in our implementation are:

$$\hat{u}_k^{(n+1)} = \frac{\hat{u}_k^{(n)} + \Delta t \cdot \left[f_u(u^{(n)}, v^{(n)})\right]_k}{1 + \Delta t \cdot s \cdot |k|^2} \tag{5}$$

$$\hat{v}_k^{(n+1)} = \frac{\hat{v}_k^{(n)} + \Delta t \cdot \left[f_v(u^{(n)}, v^{(n)})\right]_k}{1 + \Delta t \cdot s \cdot \delta \cdot |k|^2} \tag{6}$$

where $[\cdot]$ represents the FFT operator, and the nonlinear reaction terms are given by $f_u = \frac{u^2}{v(1+cu^2)} - au + b$ and $f_v = u^2 - v$.

Implementation Details The simulation was implemented using CuPy for GPU acceleration. The spatial domain was discretized into a 128×128 grid. Time integration was performed with a step size $\Delta t = 0.2$ for a total of $T_{max} = 5000$ units. Initial conditions were set to the homogeneous steady state (u^*, v^*) with additive Gaussian noise ($\sigma = 0.001$) to induce symmetry breaking. The final dataset consists of 20,000 simulation samples, stratified into training, validation, and testing sets. Examples of the generated Turing patterns are shown in Figure 1 in the appendix.

3.3 Steady-State Physics Regularization

The core of the PRESS framework is the enforcement of the steady-state equilibrium condition as a regularization term. Instead of requiring temporal trajectories to compute time derivatives, we derive a residual loss directly from the stationary form of the governing reaction–diffusion equations. This constrains the neural network to predict parameters that satisfy the delicate balance between reaction kinetics and diffusion.

3.3.1 Residual Formulation

Let $\hat{\theta} = \{\hat{a}, \hat{b}, \hat{c}, \hat{\delta}\}$ be the parameters predicted by the neural network given an input pattern (u, v). We define the physical residuals $\mathcal{R}_u$ and $\mathcal{R}_v$ based on the dimensionless Gierer-Meinhardt model (Eqs. 3-4) at steady state (i.e., $\frac{\partial u}{\partial t} \approx 0$, $\frac{\partial v}{\partial t} \approx 0$).

The residuals at each spatial location $x_{i,j}$ are given by:

$$\mathcal{R}_u(x_{i,j}) = s\Delta u_{i,j} + \frac{u_{i,j}^2}{v_{i,j}(1 + \hat{c}u_{i,j}^2)} - \hat{a}u_{i,j} + \hat{b} \tag{7}$$

$$\mathcal{R}_v(x_{i,j}) = s\hat{\delta}\Delta v_{i,j} + u_{i,j}^2 - v_{i,j} \tag{8}$$

where $s = 0.4$ is the fixed diffusion scaling parameter. Ideally, for ground-truth parameters, the system should be at equilibrium, driving these residuals to zero.

3.3.2 Numerical Approximation via Convolution

To efficiently compute the Laplacian term Δ within the neural network's computational graph, we employ a finite difference approximation implemented as a fixed convolutional layer. Specifically, we utilize the standard five-point stencil kernel K_{lap}:

$$K_{lap} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \tag{9}$$

The Laplacian of the concentration field u is computed as:

$$\Delta u \approx \frac{1}{(\Delta x)^2}(K_{lap} * u) \tag{10}$$

where $*$ denotes the convolution operation. To maintain consistency with the data generation process, we apply circular padding to the input feature maps prior to convolution, thereby enforcing periodic boundary conditions.

The final physics-regularized objective $\mathcal{L}_{phy}$ utilized in PRESS is computed as the mean squared residual over the spatial domain:

$$\mathcal{L}_{phy} = \frac{1}{N}\sum_{i,j}\left(\mathcal{R}_u(x_{i,j})^2 + \mathcal{R}_v(x_{i,j})^2\right) \tag{11}$$

where $N = 128 \times 128$ is the number of pixels in the grid.

3.3.3 Total Training Objective

The complete training objective of PRESS combines the supervised data loss with the steady-state physics regularization term. Given the predicted parameters $\hat{\theta} = \{\hat{a}, \hat{b}, \hat{c}, \hat{\delta}\}$ and the normalized ground-truth parameters θ, the total loss is defined as:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{data}} + \lambda_{\text{phy}} \cdot \mathcal{L}_{\text{phy}} \tag{12}$$

where $\mathcal{L}_{\text{data}} = \|\hat{\theta} - \theta\|^2$ is the mean squared error between predicted and ground-truth parameters in the normalized space, $\mathcal{L}_{\text{phy}}$ is the steady-state PDE residual defined in Eq. (11), and $\lambda_{\text{phy}} \geq 0$ is a weighting hyperparameter that controls the trade-off between data fidelity and physical consistency. A larger λ_{phy} enforces stricter adherence to the governing equations at the cost of potentially degrading parameter accuracy, as analyzed in Appendix A.1.

3.4 Architecture Implementation

The PRESS framework is architecture-agnostic and can be integrated into various backbone networks. To systematically evaluate the impact of network depth and spatial feature extraction strategies on parameter estimation, we implement the framework on a series of convolutional neural network architectures with varying depths (1, 2, and 4 layers). Specifically, we compare two distinct spatial sampling mechanisms to encode the Turing patterns: *Maxpooling* and *Dilated Convolution*.

In all configurations, circular padding is employed to align with the periodic boundary conditions inherent in Turing patterns. For architectures utilizing maxpooling, the model aggregates features through spatial downsampling to extract macro-structural information. Conversely, in the dilated convolution models, the dilation rate increases exponentially with layer depth. This allows the model to effectively expand its receptive field to a global scale without compromising spatial resolution—a critical feature for accurately computing the finite-difference Laplacian in the physics loss.

The final extracted features are processed through two fully connected layers to regress the predicted parameters $\hat{\theta}$. By comparing these architectures, we assess whether spatial abstraction (pooling) or high-resolution geometric preservation (dilation) is more effective when constrained by the PRESS regularization.

4 Experiments

In this section, we evaluate the effectiveness of the proposed PRESS framework for parameter estimation. We systematically compare the performance of different neural network architectures (CNN vs. MLP) and investigate the impact of the steady-state physics regularization across various dataset sizes.

4.1 Experimental Setup

4.1.1 Dataset and Preprocessing

The dataset consists of 20,000 successful simulation instances generated via the spectral numerical pipeline (Section 3), ensuring all samples exhibit valid Turing patterns. The data is partitioned as follows:

- **Training and Validation Sets**: A pool of up to 16,000 samples, partitioned with a 3:1 ratio (75% training, 25% validation). To evaluate data efficiency, models were trained on subsets of sizes $N_{train+val} \in \{2000, 4000, 8000, 16000\}$.
- **Test Set**: A dedicated hold-out set of 4,000 samples reserved for final performance evaluation.

4.1.2 Data Normalization

The governing parameters $\theta = \{a, b, c, \delta\}$ span disparate physical scales (e.g., $a \sim 10^{-2}$ vs. $\delta \sim 10^2$). To ensure numerical stability and prevent gradient dominance by larger-magnitude parameters, we apply Min-Max normalization to map each parameter $y^{(j)}$ into the unit interval $[0, 1]$, where the bounds are defined in Table 1. All subsequent loss computations and evaluation metrics are performed in this normalized space to ensure equal weighting across all physical dimensions.

4.1.3 EVALUATION METRICS

We evaluate performance using two metrics: (i) **Normalized Root Mean Square Error (NRMSE)** for parameter accuracy, and (ii) **Mean PDE Residual ($\mathcal{R}_{avg}$)**, calculated as the average pixel-wise squared residual of the governing equations (Eq. 11) over the test set. While NRMSE measures ground-truth alignment, $\mathcal{R}_{avg}$ quantifies the physical consistency of the predicted parameters. To measure the efficacy of PRESS, we report the percentage improvement (Imp.) in these metrics compared to the data-driven baseline.

4.1.4 TRAINING CONFIGURATION

Both the purely data-driven baselines and the PRESS-regularized models are trained under the same set of hyperparameters to ensure a fair comparison. Models are trained for 100 epochs (batch size 32) using Adam ($lr = 10^{-3}$) with a two-stage schedule: a 10% linear warm-up (10^{-6} to 10^{-3}) followed by cosine annealing. Random seeds are fixed across all experiments for reproducibility. No additional weight decay or parameter-level smoothing are used.

For the baseline supervised training, the loss consists solely of the mean squared error (MSE) between predicted and normalized target parameters. For PRESS training, the steady-state PDE residual loss derived from the Gierer–Meinhardt equations is added to the MSE, weighted by $\lambda_{\text{phy}} = 10^{-n}$, where $n \in \{1, 2, 3, 4, 5\}$.

4.1.5 MLP COMPARISON

In addition, we construct a multilayer perceptron (MLP) as a non-convolutional baseline. Instead of using convolutional operations that preserve spatial topology, the MLP flattens the input Turing pattern image into a one-dimensional vector and processes it through several fully connected layers to regress the four normalized system parameters. Since this architecture does not preserve local spatial neighborhood information, it serves as a reference point for evaluating the model's inference capability *without* exploiting local pattern structures. This comparison allows us to verify whether the performance gains are driven by the PRESS framework's ability to leverage spatial derivatives (Laplacians) within the CNN architecture.

4.2 RESULTS AND DISCUSSION

4.2.1 IMPACT OF STEADY-STATE PHYSICS REGULARIZATION

We evaluate the effect of the proposed **PRESS** framework across a wide range of CNN architectures, varying both network depth (L) and spatial operations, as summarized in Table 2. The results consistently demonstrate that incorporating steady-state physical constraints improves parameter inference accuracy across most architectural and data regimes, though the magnitude of improvement depends strongly on model capacity and dataset size.

Impact of Physics-Informed Loss Across Data Scales. We first evaluate the baseline single-layer architecture ($L = 1$, none) to establish the foundational impact of physical constraints. As summarized in Table 2, the PRESS-regularized model consistently outperforms the purely data-driven baseline across all evaluated dataset sizes. The physical constraint acts as a robust regularizer, particularly in the low-data regime; at 2,000 samples, PRESS achieves an NRMSE of 0.1868, marking a 14.23% improvement. Crucially, it yields a massive 56.56% reduction in PDE residual ($0.8276 \rightarrow 0.3595$), demonstrating that the model successfully internalizes the reaction-diffusion dynamics even when data is scarce. This trend persists as the data scale increases to 8,000 samples, where the improvement peaks at 23.08%. These results demonstrate that steady-state physics regularization effectively guides the network to internalize the underlying reaction-diffusion dynamics even when supervision is relatively scarce.

Influence of Architecture Depth and Model Capacity. Our experiments reveal that the advantages of PRESS scale proportionally with network complexity ($L = 2$ and $L = 4$). A remarkable performance leap is observed in the $L = 4$ dilated convolution setting: at 8,000 samples, PRESS reduces the NRMSE from 0.1299 to 0.0507, representing a substantial 60.97% improvement. This suggests that as the hypothesis space expands with depth, the physical loss becomes increasingly

Table 2: Comprehensive performance comparison. We report NRMSE (parameter error) and PDE Residual (physical inconsistency). PRESS consistently reduces the physical residual (positive Imp.), confirming adherence to governing laws even when NRMSE gains saturate. λ_{phy} denotes the optimal physical loss weight.

Depth	Spatial	Size	NRMSE ↓			PDE Residual ↓			λ_{phy}
(L)	Op.		Base	PRESS	Imp.	Base	PRESS	Imp.	
$L = 1$	None	2000	0.2178	**0.1868**	+14.23	0.8276	**0.3595**	+56.56	0.1
		4000	0.1826	**0.1526**	+16.42	0.5141	**0.2394**	+53.43	0.1
		8000	0.1616	**0.1243**	+23.08	0.3894	**0.1627**	+58.23	0.1
		16000	0.1445	**0.1119**	+22.61	0.2545	**0.1626**	+36.11	0.1
$L = 2$	Dilated	2000	0.1713	**0.1409**	+17.77	0.4459	**0.0612**	+86.28	0.1
		4000	0.1545	**0.1080**	+30.08	0.3552	**0.0784**	+77.93	0.1
		8000	0.1222	**0.0811**	+33.61	0.3708	**0.0740**	+80.04	0.1
		16000	0.1172	**0.1170**	+0.17	0.1065	**0.1010**	+5.14	0.00001
	Maxpool	2000	0.1421	**0.0879**	+38.17	0.4578	**0.1568**	+65.74	0.1
		4000	0.1253	**0.0847**	+32.46	0.3563	**0.1246**	+49.12	0.001
		8000	0.0645	**0.0541**	+16.16	0.1343	**0.0827**	+61.17	0.01
		16000	0.0497	**0.0492**	+1.00	0.1032	**0.1014**	+1.82	0.0001
$L = 4$	Dilated	2000	0.1059	**0.0777**	+26.63	0.2312	**0.1261**	+45.46	0.1
		4000	0.0952	**0.0692**	+27.28	0.1802	**0.0871**	+51.71	0.1
		8000	0.1299	**0.0507**	+60.97	0.5369	**0.0440**	+91.81	0.1
		16000	0.0538	**0.0455**	+15.51	0.1013	**0.0434**	+57.16	0.1
	Maxpool	2000	0.1121	**0.1100**	+1.83	0.2804	**0.1741**	+37.91	0.01
		4000	0.0735	**0.0702**	+4.47	0.1481	**0.1429**	+3.51	0.001
		8000	0.0466	**0.0379**	+18.54	0.0496	**0.0184**	+62.97	0.001
		16000	**0.0303**	0.0319	-5.24	0.0182	**0.0181**	+0.27	0.001

critical in penalizing non-physical solutions that a deep baseline might otherwise overfit. By enforcing the Gierer–Meinhardt equations, PRESS effectively "prunes" the learning space, ensuring that high-capacity models remain grounded in the governing physics.

Behavior in High-Data Regimes. While PRESS yields significant gains in low- and intermediate-data regimes, we observe diminishing returns at 16,000 samples. The purely data-driven signal becomes strong enough for high-capacity models to approximate the solution manifold. For instance, in the $L = 4$ Maxpool configuration at 16k samples, the PRESS model exhibits a slightly higher NRMSE (0.0319) compared to the baseline (0.0303). However, the PDE Residual still decreases ($0.0182 \rightarrow 0.0181$, +0.27%). This confirms that PRESS regularization successfully minimized the physical error as intended, even when numerical approximation biases prevent further NRMSE gains. Similarly, for the $L = 2$ Dilated model at 16k samples, while NRMSE improvement is negligible ($+0.17\%$), the PDE Residual improves significantly ($0.1065 \rightarrow 0.1010$, a $+5.14\%$ reduction). This indicates that even when parameter gains saturate, PRESS continues to enforce better physical consistency.

4.2.2 Architecture Comparison: MLP vs. CNN

To examine the impact of spatial inductive bias on inverse parameter inference, we compare a Multi-Layer Perceptron (MLP), which operates on flattened inputs, with the aforementioned single-layer convolutional model that preserves local spatial structure. Results across all dataset sizes are summarized in Table 3.

Across all data regimes, the MLP exhibits stable behavior but benefits only marginally from the PRESS framework, with NRMSE reductions limited to $+0.52\%$–$+4.98\%$. This indicates that, without spatial inductive bias, physical constraints alone provide limited improvement for fully connected architectures.

Table 3: Comparison between MLP and single-layer CNN ($L = 1$). While MLP shows limited improvement, CNN achieves significantly higher residual reduction (Imp.), demonstrating superior physical consistency.

Model	Size	NRMSE ↓			PDE Residual ↓			λ_{phy}
		Base	PRESS	Imp.	Base	PRESS	Imp.	
MLP	2000	0.2096	**0.2049**	+2.26	0.5448	**0.4432**	+18.66	0.1
	4000	0.1919	**0.1824**	+4.98	0.4338	**0.3448**	+20.52	0.1
	8000	0.1767	**0.1745**	+1.25	0.3824	**0.3293**	+13.89	0.00001
	16000	0.1683	**0.1674**	+0.52	0.3211	**0.2801**	+12.77	0.00001
CNN ($L = 1$)	2000	0.2178	**0.1868**	+14.23	0.8276	**0.3595**	+56.56	0.1
	4000	0.1826	**0.1526**	+16.42	0.5141	**0.2394**	+53.43	0.1
	8000	0.1616	**0.1243**	+23.08	0.3894	**0.1627**	+58.23	0.1
	16000	0.1445	**0.1119**	+22.61	0.2545	**0.1626**	+36.11	0.1

In contrast, the single-layer CNN responds strongly to the steady-state physics loss. Although the unconstrained CNN is comparable to or slightly worse than the MLP in the low-data regime (e.g., 2,000 samples), incorporating physics constraints yields substantial improvements, outperforming both baseline and physics-informed MLPs. This advantage also holds for other data sizes.

Overall, these results demonstrate that the efficacy of PRESS is strongly architecture-dependent: MLPs show limited sensitivity to physical constraints, whereas CNNs leverage them to unlock spatial inductive bias, resulting in substantially larger performance gains in inverse parameter inference.

5 Conclusion

In this work, we revisited the inverse problem of recovering reaction–diffusion parameters from static Turing patterns through the lens of physics-informed deep learning. Focusing on the Gierer–Meinhardt model, we constructed a synthetic dataset of steady-state patterns and demonstrated that convolutional neural networks can reliably infer governing parameters directly from raw snapshot images when constrained by the PRESS framework.

Our experiments reveal that incorporating steady-state physics regularization fundamentally enhances the accuracy and robustness of inference, with the most significant gains observed in low- and intermediate-data regimes. By explicitly enforcing the equilibrium condition of the reaction-diffusion system, PRESS leverages both local spatial derivatives and global dynamical constraints, significantly outperforming baselines that lack this physical grounding. These findings suggest that the previously reported limitations of CNNs for inverse Turing problems stem not from intrinsic architectural flaws, but from the absence of suitable physical regularization for snapshot data. By combining a convolutional encoder with a steady-state residual loss, our framework achieves efficient, end-to-end parameter inference without computationally expensive per-instance optimization, providing a scalable and physically consistent tool for quantitatively analyzing pattern-forming systems.

References

Selma Emekci. A physics informed convolutional neural network for simulating carbon nanotube wires in infinite spatial dimensions and arbitrary frequency. In *2024 IEEE International Conference on Big Data (BigData)*, pp. 7308–7312, 2024. doi: 10.1109/BigData62323.2024.10825182.

Xiaoyu Ge and Javad Khazaei. Physics-informed convolutional neural network for microgrid economic dispatch. *Sustainable Energy, Grids and Networks*, 40:101525, 2024. ISSN 2352-4677. doi: 10.1016/j.segan.2024.101525.

Alfred Gierer and Hans Meinhardt. Gierer, a. meinhardt, h. a theory of biological pattern formation. kybernetik 12, 30- 39. *Biological Cybernetics*, 12:30–39, 01 1972. doi: 10.1007/BF00289234.

I. Hariri, A. Radid, and K. Rhofir. Physics-informed neural networks methodology for the resolution of some turing's type models. *Advanced Mathematical Models & Applications*, 10(1):61–80, 04 2025. doi: 10.62476/amma.10161.

Daniel Kelshaw and Luca Magri. Physics-constrained convolutional neural networks for inverse problems in spatiotemporal partial differential equations. *Data-Centric Engineering*, 5:e43, 2024. doi: 10.1017/dce.2024.46.

Jordon Kho, Winston Koh, Jian Cheng Wong, Pao-Hsiung Chiu, and Chin Chun Ooi. Design of turing systems with physics-informed neural networks. In *2022 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1180–1186, 2022. doi: 10.1109/SSCI51031.2022.10022026.

Angran Li, Ruijia Chen, Amir Barati Farimani, and Yongjie Jessica Zhang. Reaction diffusion system prediction based on convolutional neural network. *Scientific Reports*, 10(1):3894, mar 2020. doi: 10.1038/s41598-020-60853-2.

Antonio Matas-Gil and Robert G. Endres. Unraveling biochemical spatial patterns: Machine learning approaches to the inverse problem of stationary turing patterns. *iScience*, 27(6):109822, 2024. ISSN 2589-0042. doi: 10.1016/j.isci.2024.109822.

J.D. Murray. Parameter space for turing instability in reaction diffusion mechanisms: A comparison of models. *Journal of Theoretical Biology*, 98(1):143–163, 1982. ISSN 0022-5193. doi: 10.1016/0022-5193(82)90063-7.

Chengping Rao, Pu Ren, Qi Wang, Oral Buyukozturk, Hao Sun, and Yang Liu. Encoding physics to learn reaction–diffusion processes. *Nature Machine Intelligence*, 5(7):765–779, Jul 2023. doi: 10.1038/s42256-023-00685-7. URL https://doi.org/10.1038/s42256-023-00685-7.

David Schnörr and Christoph Schnörr. Learning system parameters from turing patterns. *Machine Learning*, 112(9):3151–3190, June 2023. ISSN 0885-6125. doi: 10.1007/s10994-023-06334-9.

Christian Scholz and Sandy Scholz. Exploring complex pattern formation with convolutional neural networks. *American Journal of Physics*, 90(2):141–151, 02 2022. ISSN 0002-9505. doi: 10.1119/5.0065458. URL https://doi.org/10.1119/5.0065458.

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 Mathison Turing. The chemical basis of morphogenesis. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 237(641):37–72, 08 1952. ISSN 0080-4622. doi: 10.1098/rstb.1952.0012.

Biao Yuan, He Wang, Ana Heitor, and Xiaohui Chen. f-picnn: A physics-informed convolutional neural network for partial differential equations with space-time domain. *Journal of Computational Physics*, 515:113284, 2024. ISSN 0021-9991. doi: 10.1016/j.jcp.2024.113284.

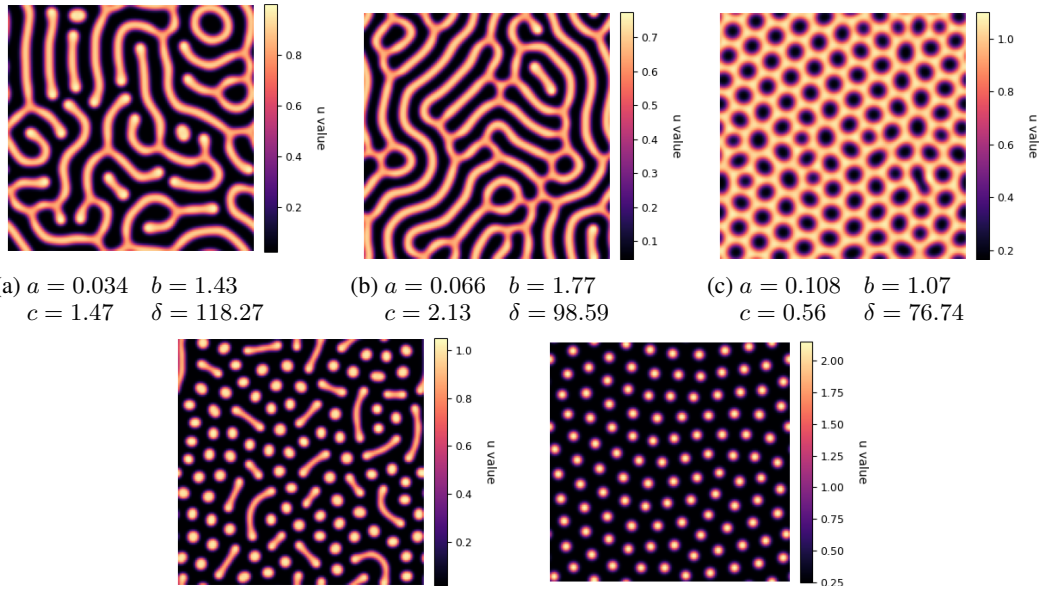
Zhao Zhang. A physics-informed deep convolutional neural network for simulating and predicting transient darcy flows in heterogeneous reservoirs without labeled data. *Journal of Petroleum Science and Engineering*, 211:110179, 2022. ISSN 0920-4105. doi: 10.1016/j.petrol.2022.110179.

Xiaoyu Zhao, Zhiqiang Gong, Yunyang Zhang, Wen Yao, and Xiaoqian Chen. Physics-informed convolutional neural networks for temperature field prediction of heat source layout without labeled data. *Engineering Applications of Artificial Intelligence*, 117:105516, 2023. ISSN 0952-1976. doi: 10.1016/j.engappai.2022.105516.

A Appendix

A.1 Case Study: Sensitivity Analysis of Regularization Weight

We systematically investigated the impact of the hyperparameter λ_{phy}, which controls the trade-off between the data-driven objective and the steady-state physics constraint. Table 4 presents the



(a) $a = 0.034$ $b = 1.43$ $c = 1.47$ $\delta = 118.27$

(b) $a = 0.066$ $b = 1.77$ $c = 2.13$ $\delta = 98.59$

(c) $a = 0.108$ $b = 1.07$ $c = 0.56$ $\delta = 76.74$

(d) $a = 0.030$ $b = 1.91$ $c = 1.44$ $\delta = 115.81$

(e) $a = 0.316$ $b = 1.78$ $c = 0.21$ $\delta = 175.70$

Figure 1: Examples of Turing patterns generated on a 128×128 spatial grid. These examples illustrate the diverse morphologies inherent to the system, including isolated spots and labyrinthine structures, emerging from variations in the governing parameters. The color intensity corresponds to the concentration of species u, where brighter colors denote higher concentrations.

performance of the L4 Maxpool architecture trained on 8,000 samples across varying magnitudes of λ_{phy}. The results reveal a distinct non-monotonic relationship, indicating that the strength of the physical regularization must be carefully tuned to achieve synergy within the PRESS framework.

Table 4: Sensitivity analysis of the **steady-state physics loss weight** λ_{phy} for the $L = 4$ Maxpool architecture trained on $N = 8,000$ samples. The baseline model corresponds to $\lambda_{phy} = 0$. We report NRMSE (parameter error) and PDE Residual (physical inconsistency). The improvement is calculated relative to the baseline.

Loss Weight (λ_{phy})	**NRMSE ↓**	**Imp.**	**PDE Residual ↓**	**Imp.**
0 (Baseline)	0.0466	-	0.0698	-
0.1	0.0573	-22.96%	0.0205	+70.56%
0.01	0.0642	-37.91%	0.0586	+16.04%
0.001	**0.0379**	**+18.54%**	**0.0258**	**+62.97%**
0.0001	0.0480	-3.06%	0.0709	-1.58%
0.00001	0.0544	-16.71%	0.1217	-74.44%

Optimal Regularization. The model achieves peak performance at $\lambda_{phy} = 0.001$, yielding a remarkable 18.54% reduction in NRMSE compared to the pure data-driven baseline (0.0379 vs. 0.0466). Crucially, this is accompanied by a massive 62.97% reduction in PDE Residual (0.0258 vs. 0.0698). In this regime, the physical loss effectively functions as a domain-specific regularizer, guiding the optimization landscape toward physically consistent parameter regions without dominating the gradient descent trajectory.

Over-Regularization. Conversely, setting λ_{phy} too high (≥ 0.01) leads to an instructive trade-off. At $\lambda_{phy} = 0.1$, the PDE Residual is minimized to its lowest value (0.0205, a 70.56% improvement), yet the NRMSE degrades significantly (-22.96%). This explicitly confirms that excessive weighting

causes the optimization to prioritize satisfying the PDE constraint at the expense of accurately regressing the ground-truth parameters, effectively "overfitting" the physics loss.

Under-Regularization. When λ_{phy} is insufficiently small (≤ 0.0001), the regularization effect vanishes. The network behaves similarly to or worse than the unconstrained baseline (e.g., at $\lambda_{phy} = 0.00001$, both NRMSE and Residual degrade), likely due to the introduction of uninformative gradients that fail to leverage the structural information encoded in the reaction-diffusion equations.

These findings underscore that while physics-informed constraints are powerful, they require a balanced integration where physical laws complement, rather than override, the observational data.

To qualitatively assess the model's performance, we visualize the inference results for one of the representative test sample from the $L = 4$ Maxpool models trained on 8,000 samples. Figure 2 compares the prediction quality between the purely data-driven baseline and the PRESS-regularized model ($\lambda_{phy} = 0.001$).

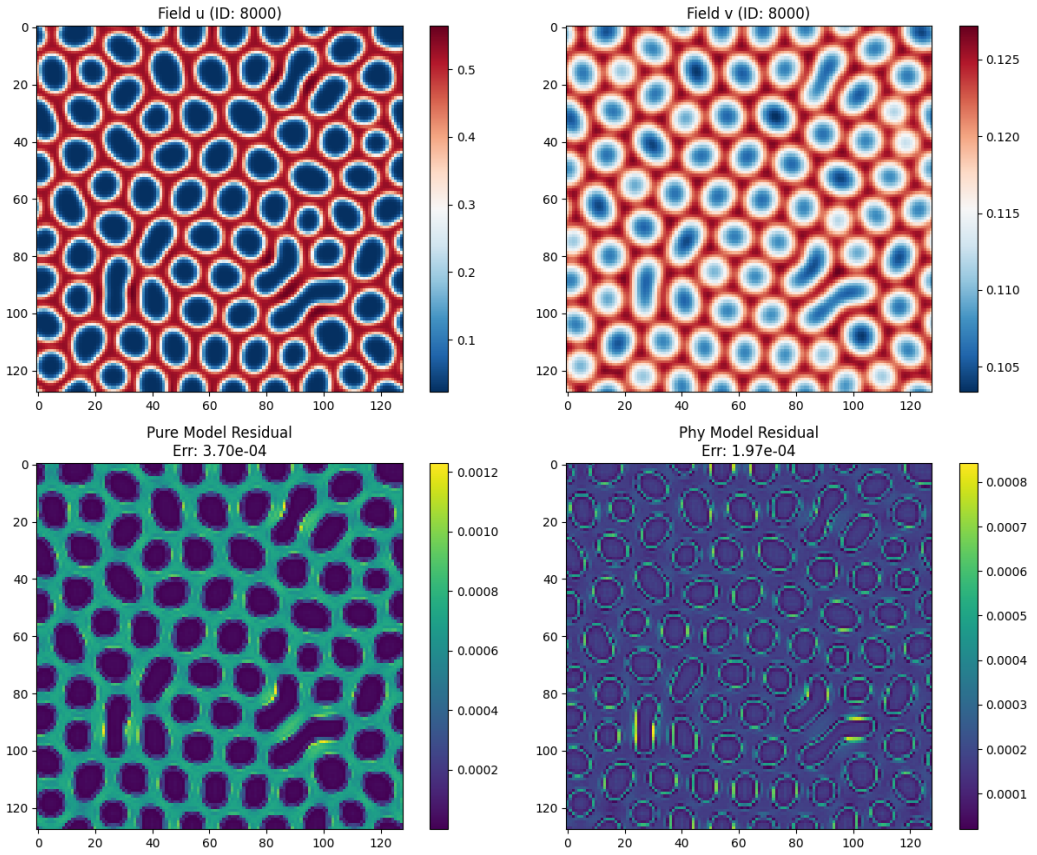


Figure 2: **Visual comparison of parameter inference.** The ground truth parameters are $a = 0.0384, b = 1.9097, c = 4.5398, \delta = 185.9997$. Left: The unconstrained baseline often predicts parameters that lead to distorted or structurally incorrect patterns. Right: The PRESS framework ($\lambda_{phy} = 0.001$) recovers parameters that faithfully reproduce the ground truth morphology, demonstrating superior physical consistency.

A.2 Metric Definitions

To rigorously evaluate the performance of PRESS, we employ two complementary metrics: one measuring the parameter inference accuracy in the normalized space, and the other quantifying the physical consistency of the predictions with respect to the governing laws.

Normalized Root Mean Square Error (NRMSE). Since the governing parameters span different orders of magnitude, we evaluate accuracy using the normalized parameter vectors. Let $\theta^{(i)} \in [0,1]^4$ denote the ground-truth normalized parameter vector for the i-th test sample, and $\hat{\theta}^{(i)} \in [0,1]^4$ denote the corresponding prediction by the network. The NRMSE is calculated as the average Euclidean distance over the test set of size $N_{test} = 4,000$:

$$\text{NRMSE} = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} \left\| \hat{\theta}^{(i)} - \theta^{(i)} \right\|_2 = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} \sqrt{\sum_{j=1}^{4} \left(\hat{\theta}_j^{(i)} - \theta_j^{(i)} \right)^2} \tag{13}$$

This metric provides a direct measure of how close the inferred parameters are to the ground truth manifold in the normalized hypercube.

Mean PDE Residual ($\mathcal{R}_{avg}$). The PDE residual serves as a proxy for physical validity, independent of the ground truth labels. For a given sample i with predicted parameters $\hat{\theta}^{(i)}$, we compute the steady-state residuals $\mathcal{R}_u$ and $\mathcal{R}_v$ (as defined in Eqs. 7-8) at each spatial location $\mathbf{x}$ on the discretized grid Ω. The residual metric for a single sample, $\mathcal{L}_{res}^{(i)}$, is the mean squared residual over all 128×128 pixels:

$$\mathcal{L}_{res}^{(i)} = \frac{1}{|\Omega|} \sum_{\mathbf{x} \in \Omega} \left(\mathcal{R}_u(\mathbf{x}; \hat{\theta}^{(i)})^2 + \mathcal{R}_v(\mathbf{x}; \hat{\theta}^{(i)})^2 \right) \tag{14}$$

The reported **Mean PDE Residual** is then the average of this quantity across the entire test set:

$$\mathcal{R}_{avg} = \frac{1}{N_{test}} \sum_{i=1}^{N_{test}} \mathcal{L}_{res}^{(i)} \tag{15}$$

A lower $\mathcal{R}_{avg}$ indicates that the predicted parameters generate a reaction-diffusion field that is closer to the steady-state equilibrium condition required by the Gierer-Meinhardt model.

A.3 Future Work

While this study demonstrates the efficacy of the PRESS framework for parameter discovery in the Gierer–Meinhardt model, several avenues for future research remain:

- **Application to Real Biological Data:** The current framework relies on high-fidelity synthetic data generated under controlled numerical conditions. A critical next step is to validate the model on experimental biological images, such as pigmentation patterns in zebrafish or skin textures. This will require addressing observational challenges such as non-uniform lighting, perspective distortion, and complex measurement noise. Furthermore, distinct from the Gaussian perturbations we employed solely to trigger symmetry breaking, the intrinsic stochastic fluctuations that initiate morphogenesis in real biological systems may follow non-Gaussian distributions.
- **Generalization to Other Reaction–Diffusion Systems:** We focused exclusively on the Gierer–Meinhardt model. Future work should assess the universality of PRESS by applying it to other classic Turing systems, such as the Schnakenberg, Gray–Scott, or FitzHugh–Nagumo models. Investigating whether the steady-state assumption remains sufficient for systems exhibiting complex or chaotic regimes would be particularly valuable.
- **Adaptive Balancing of Physical Constraints:** In our current implementation, the weight of the physical loss λ_{phy} is treated as a fixed hyperparameter. This static approach may not be optimal throughout the entire training trajectory. Future work should explore adaptive curriculum learning strategies, where the influence of the physical constraint is dynamically adjusted based on the training stability or gradient norms. Systematically analyzing the optimal timing and proportion of injecting physical laws could provide a generalizable guideline for balancing data-driven fitting and physical consistency.

- **Complex Geometries and Boundary Conditions:** Our experiments utilized square domains with periodic boundary conditions to facilitate spectral data generation. However, biological pattern formation often occurs on irregular, non-periodic domains (e.g., limb buds or animal skins). Extending the framework to handle Dirichlet or Neumann boundary conditions, or adapting the architecture (e.g., using Graph Neural Networks or Physics-Informed DeepONets) to process patterns on non-Euclidean manifolds or 3D volumetric data, represents a significant direction for development.