

# PINN-EM: Physics-Guided Disease Progression Model of Geographic Atrophy

Dmitrii Lachinov, Thomas Pinetz, and Hrvoje Bogunović

**Abstract—Objective:** To construct a personalizable spatio-temporal disease progression model of patients with a late dry form of Age-Related Macular Degeneration (AMD), known as Geographic Atrophy (GA). **Methods:** From a series of retinal optical coherence tomography (OCT) scans, we infer the coefficients for the parametrized partial differential equation (PDE), such that the parametrized PDE best describes the observed imaging data. Acting as a soft constraint, the recovered PDE helps to extrapolate an implicit neural representation (INR) of the GA segmentation map progression. To enable efficient training, we propose an iterative method - PINN-EM, designed to recover coefficients of non-linear PDEs. At each iteration, the method decouples the problem into PDE coefficients fitting and data fitting steps, resembling Expectation Maximization algorithm. **Results:** We extensively tested the proposed method in large-scale experiments using the open-source PDEBench benchmark to validate its performance. Furthermore, we applied the method to the challenging problem of GA progression modeling, where patients exhibit a high variance in GA growth patterns and speed. The proposed spatio-temporal disease progression model outperformed the baselines, even outperforming posterior knowledge models in Dice score for newly affected growth areas. **Conclusion:** We demonstrated that the proposed spatio-temporal disease progression model fitted with introduced PINN-EM outperforms existing baselines in synthetic and real clinical applications, highlighting the extrapolation capabilities of the INR models. **Significance:** The proposed spatio-temporal disease progression model and PINN-EM fitting procedure can be applied across diverse domains facing the challenge of fitting parametrized PDE to the empirical datasets.

**Index Terms—**Disease Progression, DPM, Geographic Atrophy, Implicit Neural Representation, PINNs

## I. INTRODUCTION

This work was supported in part by the Christian Doppler Research Association, Austrian Federal Ministry of Economy, Energy and Tourism, the National Foundation for Research, Technology and Development. (Corresponding author: Dmitrii Lachinov)

Dmitrii Lachinov, Thomas Pinetz and Hrvoje Bogunović are with Institute of Artificial Intelligence, Center for Medical Data Science, and also with Comprehensive Center for AI in Medicine, Medical University of Vienna, 1090 Vienna, Austria. (e-mail: dmitrii.lachinov@meduniwien.ac.at, thomas.pinetz@meduniwien.ac.at, hrvoje.bogunovic@meduniwien.ac.at)

Dmitrii Lachinov and Hrvoje Bogunović are also with Christian Doppler Lab for Artificial Intelligence in Retina, Institute of Artificial Intelligence, Center for Medical Data Science, Medical University of Vienna, 1090 Vienna, Austria.

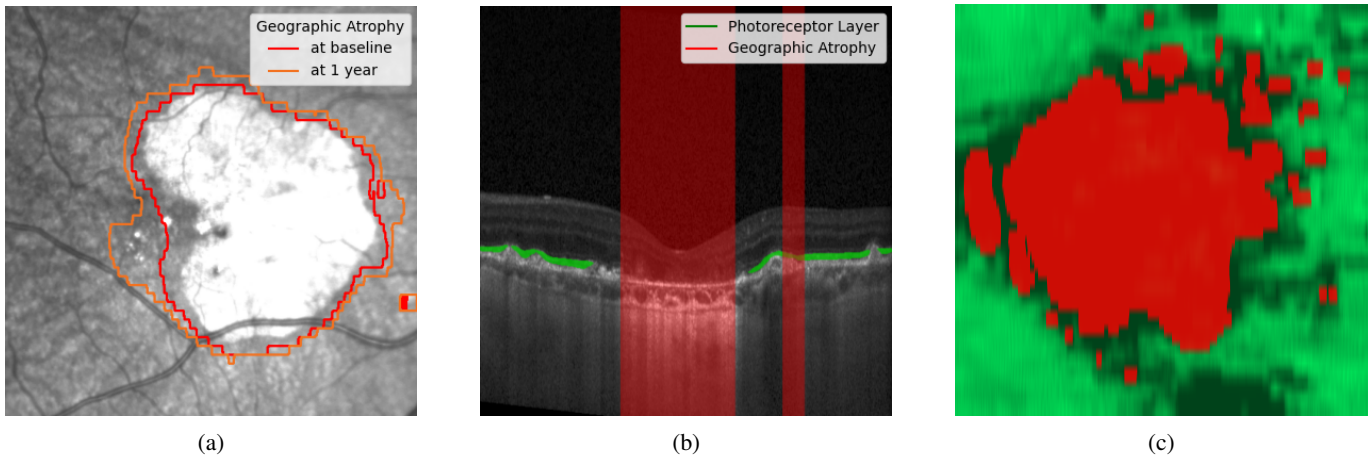
Code Available: [https://github.com/lachinov/pinn\\_em](https://github.com/lachinov/pinn_em)

PREDICTING disease progression from medical imaging data remains a challenging and largely an unsolved problem with a limited number of prior works in *spatio-temporal* biomedical domains. The main challenges in addressing disease progression from the perspective of modeling *spatio-temporal* changes are high non-linearity of the progression dynamics, including arbitrary pathological shapes, heterogeneity of the normal and pathological alterations, stochastic nature of the underlying processes, rendering long-term predictions ineffective, scarcity and sampling frequency of the imaging data, and, oftentimes, access exclusively to limited data of specific imaging modality.

Depending on the strength of the underlying assumptions, disease progression models can be roughly classified into *analytical*, forcing physical constraints on the model, and *empirical*, inferring hidden dependencies from the data. *Analytical* models typically describe the disease progression as a dynamical system, e.g., as a bio-chemical process, and utilize hard-wired prior knowledge about disease progression in the form of ordinary (ODEs) [1] or partial (PDEs) [2] differential equations. *Empirical* models, on the other hand, infer the progression dynamics directly from the observed data, incorporating little to no prior assumption. The class of *empirical* disease progression models is dominated by statistical and machine learning methods such as linear regression [3], multivariate Cox proportional hazards model [4], Conditional Restricted Boltzmann Machine [5], and recurrent neural networks [6].

The two classes of models have their own set of benefits and downsides. For instance, *analytical* models offer a robust framework, typically through the use of differential equations, describing the pathological processes in detail. However, such models often lack flexibility in modeling real data and are notoriously hard to construct, as they require deep domain knowledge. On the other hand, *empirical* models are designed to fit the observed data, completely ignoring the underlying processes. These models are much easier to construct, as they capture hidden correlations (e.g. proximity to tumor) from the data through automated tools. *Empirical* models naturally excel at approximating the observed data, however, they often struggle with generalization and extrapolation, in addition to failing to provide deep insights into the disease progression as they are in general black-box models.

In this paper, we aim to bring the benefits of both *analytical* and *empirical* models together through the use of Physics Informed Neural Networks (PINNs) [7]. Our objective is



**Fig. 1:** (a) En-face view of the retina highlighting the progression of the Geographic Atrophy (GA); (b) Cross-section (B-scan) of the retina on OCT, highlighting areas affected by GA (loss of RPE cells in red) and the photoreceptor layer (green); (c) En-face view of the areas affected by GA (red) and thickness of the surrounding photoreceptor layer (shades of green).

to expand the class of *spatio-temporal* disease models with implicit neural representations (INR) of the progression profile constrained by a surrogate PDE using the PINNs framework. We apply this to construct a spatio-temporal disease progression model of Geographic Atrophy (GA)(Fig. 1-a), a late dry form of Age-related Macular Degeneration (AMD). To this end, we consider a series of prior optical coherence tomography (OCT) scans where retinal layers were segmented (Fig. 1-c). Through processing the series of prior segmentations and modeling disease progression with INR constrained by reaction-diffusion PDE, we attempt to predict the disease segmentation at arbitrary future time points (Fig. 5).

#### A. Related work

We aim to leverage imaging data to construct a *spatio-temporal* disease progression model, incorporating soft constraints imposed by the PDE. Hence, in this subsection, we first examine existing methodologies in *spatio-temporal* disease progression modeling. Then, we present an overview of a key component of our approach by exploring prior research on Physics Informed Neural Networks (PINNs).

**1) Disease progression models:** As the main topic of interest, we review *spatio-temporal* disease progression models. For *analytical* approaches, PDE-based models are commonly used and play a key role in therapy planning. Frieboes et al. [8] introduced reaction-diffusion equations describing tumor progression, accounting for tumor mass, cell velocity, local pressure, nutrients, and oxygen uptake. Later studies incorporated richer data, such as DTI derivatives, into anisotropic diffusion models [9]. Recent developments include chemotaxis terms directed toward vasculature [10]. Expanding on brain tumor modeling, Le et al. [11] developed a system for automatic radiotherapy planning, using a personalized reaction-diffusion solver based on individual MRI scans. Konukoglu et al. [12] personalized prediction models by recovering patient-specific PDE coefficients from MRI series. Ezhov et al. [13] proposed a computationally efficient learning-based method to estimate these parameters from a single scan, validated on synthetic and

real rat brain tumor data. Subramanian et al. [14] approached tumor growth in reverse, restoring individual PDE coefficients and initial conditions to identify tumor origin.

*Empirical* models are primarily based on machine learning. Before the rise of deep learning, Morris et al. [15] developed a model predicting glioma growth from patient-specific features and tumor attributes, incorporating noise reduction, registration, intensity standardization, tissue segmentation, and diffusion components. More recently, Zhang et al. [16] used CNNs for tumor growth prediction, with one CNN modeling cell invasion and another capturing mass-effect from tumor expansion. They later employed Convolutional LSTMs [17] to capture temporal dynamics from multiple patient scans [18]. Petersen et al. [19] further advanced glioma growth modeling by extending Neural Processes [20] with spatio-temporal attention and multi-scale representations.

In retinal imaging, Zhang et al. [21] used Bidirectional LSTMs for GA progression and U-Net to refine GA segmentation masks. Gigon et al. [22] applied Taylor series decomposition and CNNs to predict temporal derivatives of GA from retinal layer thickness maps. Implicit deep learning approaches were explored by Lachinov et al. [23], who framed disease progression as a learnable dynamical system using NeuralODEs [24] to forecast spatio-temporal AMD changes. Liu [25] extended this framework with NeuralODEs and NeuralSDEs [26] to model structural changes directly from imaging data, producing time series of images from a single scan and validating on retinal, multiple sclerosis, and glioblastoma datasets. In contrast, our study introduces a unique approach to GA modeling that, to the best of our knowledge, has not been previously explored.

**2) Physics Informed Neural Networks (PINNs):** PINNs [7] take a unique role in modeling dynamical processes, combining features of *analytical* models, as they solve boundary value problems, with those of *empirical* models, since they can learn from large datasets. Using modern auto-differentiation tools, Raissi et al. [7] computed the residual of a differential equation with a solution parameterized by a multilayer perceptron,

demonstrating strong performance on several benchmark problems. However, applying PINNs to boundary value problems presents challenges. Sampling is critical for training accuracy, and Wu et al. [27] showed the benefits of adaptive residual-based sampling. Wang et al. [28] highlighted issues with spatio-temporal causality and proposed loss-function modifications. Jagtap et al. [29] and Wang et al. [30] studied adaptive and cumulative activation functions, demonstrating improved performance across problems. Sitzmann et al. [31] addressed the inability of neural networks to capture high-frequency signals and their derivatives, proposing periodic activations that enhance representation quality.

In the domain of disease progression modeling, specifically spatio-temporal modeling, PINNs have a rather limited number of applications. Shaier et al. [32] employed PINNs for creating population-level models for predicting the spread of infectious diseases and estimating unique parameters such as the death rate. Zhang et al. [33] utilized PINNs and symbolic regression for modeling the progression of Alzheimer's disease from tau positron emission tomography data. Another group of authors, Zhang et al. [34], proposed a method to estimate patient-specific parameters of the reaction-diffusion PDE model of glioblastoma growth from a single MRI scan using PINNs.

### B. Clinical Background - Geographic Atrophy

We focus on AMD, which is a leading cause of blindness in the developed world. In particular, GA, the dry late stage of AMD, is characterized by the gradual loss of the Photoreceptor (PR) layer (Fig. 1) and the underlying Retinal Pigment Epithelium (RPE) cells of the retina. Fig. 1 depicts the key biomarkers - PR thickness and RPE loss, and GA progression over a one year. The disease impairs vision and contribute to total and irreversible vision loss. GA is widespread in individuals over 85 years of age with 9.8% prevalence in the European population [35]. Recently, the very first therapeutic treatments for GA became available [36], [37]. Thus, active research now aims to uncover the causes and risk factors of GA progression [38], where precise and robust progression models would contribute towards patient management, forecasting vision loss, and aiding in drug prescription by stratifying the patients into different progression risk groups.

### C. Contribution

In this paper, we address the problem of disease progression from the novel perspective of Implicit Neural Representation learning. In the example of GA, we construct an individualized representation of the disease progression constrained by reaction-diffusion PDE through PINN.

First, we introduce a novel optimization framework for PINNs to estimate unknown coefficients in PDEs from observational data. Adapting the Expectation-Maximization algorithm, our approach iteratively refines parameter estimates and neural network weights to enhance convergence. Through extensive experiments on synthetic data from the publicly available PDEBench dataset, we demonstrate the framework's efficacy and robustness.

Second, we leverage the proposed PINN-EM framework to train a spatio-temporal disease progression model, encoding disease trajectories as PDE-constrained implicit neural representations. We validate the framework's effectiveness using real-world clinical imaging data for GA progression. In experiments comparing the proposed method to multiple strong baselines, including PINNs-based approaches, our framework outperforms prior state-of-the-art methods.

The application to retinal imaging represents a novel methodological advancement, addressing the problem of GA progression through PDE-constrained computational framework, a perspective previously unexplored in the field of retinal disease progression. Notably, the methodological contribution of the proposed optimization scheme extends beyond disease progression modeling, and is applicable across diverse domains facing the challenge of fitting parameterized partial differential equations to the empirical datasets as shown in our extensive synthetic experiments.

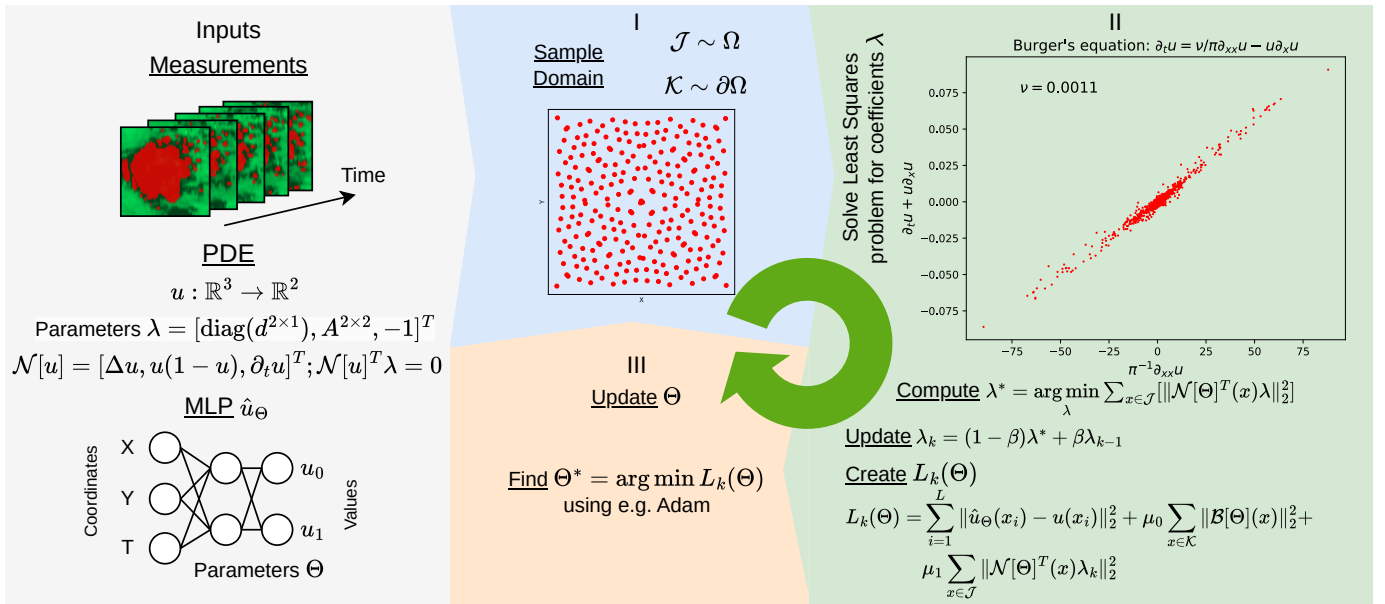
## II. METHODS

As a starting point, we consider a series of prior 3D OCT scans of a single patient. We segment each 3D volume using the segmentation CNNs [39], [40] and extract a planar map of GA and the surrounding PR layer thickness (Fig. 1(c)). We treat these segmentation maps as an input to our algorithm to predict how these feature maps evolve over time.

For convenience, we represent GA segmentation and PR thickness as a function  $u(x, t) : \mathbb{R}^{N+1} \rightarrow \mathbb{R}^M$ , taking spatial coordinates and additional time component, and map it to the vector of corresponding measurements at the spatial location and time. Next, we try to recover a differential operator  $\mathcal{N}$  parameterized by  $\lambda$  s.t.  $\mathcal{N}[u, \lambda] = 0$ . Once we reconstructed the differential operator  $\mathcal{N}$ , we are able to extrapolate from the last visit by learning the implicit neural representation of  $u$  subject to  $\mathcal{N}[u, \lambda] = 0$ . In the next sections, we discuss the choice of the differential operator  $\mathcal{N}[u, \lambda]$  specific to the problem of GA progression modeling, and the proposed method to recover coefficients  $\lambda$  from the observed data. The workflow of the proposed method is depicted in Fig. 2.

### A. Selection of the appropriate PDE

We use PINNs to learn an implicit neural representation of GA progression, which requires prior knowledge of the PDE governing the process. However, computational models of Geographic Atrophy formulated as dynamical systems are currently lacking. Developments in Retinitis Pigmentosa [41], a disease leading to degeneration of Photoreceptor cells, similar to GA, suggest to use reaction-diffusion models. In turn, observational studies of AMD and GA indicate that AMD affects the photoreceptor layer, which precedes GA development, and that its thinning accelerates GA progression [42], [43] - a pattern that can be represented by reaction-diffusion model. Furthermore, the research in brain tumor imaging [8], [11]–[14] highlights successful applications of reaction-diffusion equation to brain tumor modeling, which can be translated to our use-case.



**Fig. 2:** We consider a series of spatio-temporal measurements, e.g. thickness of PR layer at multiple visits, parameterized partial-differential equation, e.g. reaction-diffusion, and parameterized solution  $u$ , e.g. neural network. During the iterative process we repeat three steps: (I) we sample the domain and boundary points for unsupervised loss, (II) we solve the least squares problem for parameters  $\lambda$ , update the estimate of  $\lambda$  and create the optimization objective  $L_k(\Theta)$ , (III) we optimize function  $L_k$  and update parameters  $\Theta$  of the neural network.

Based on these insights, we hypothesize that GA progression can be modeled using reaction-diffusion equations incorporating photoreceptor thickness and the retinal pigment epithelium layer. In this work, we showcase that this simple formulation already provides good results, even when utilizing FEM coupled with adjoint sensitivity analysis. Discovering the exact governing equations is beyond the scope of this study. Instead, we use the PDE to guide the INR to ensure generalization over a specified future time interval.

### B. Intuition

For an intuitive explanation, we refer to spatio-temporal data representing Geographic Atrophy progression, provided as an mp4 animation in the supplementary materials, or as a single slice at  $t = 0$  in Fig. 1c. Ultimately, from a few prior observations, we aim to reconstruct the full patient trajectory, including the portion that has not yet been observed. To achieve this, we represent the trajectory as a vector function  $u(x, t)$ , which describes PR thickness (green) and GA state (red) at pixel location  $x$  and visit time  $t$  (Fig. 1c).

Since the function  $u$  can only be observed through actual patient scans, which are costly; we approximate it with a neural network  $\hat{u}_\Theta$ , which we can manipulate more flexibly. For  $\hat{u}_\Theta$  to be accurate, it must obey the same physical laws as the original  $u$ . These laws are encoded in a differential operator  $\mathcal{N}[u, \lambda]$  parameterized by  $\lambda$ . One example is the heat dissipation operator,  $\mathcal{N}[u, \lambda] = \lambda \Delta u - \partial_t u$ , where  $\lambda$  controls the rate of heat dissipation, and the Laplacian  $\Delta u$  reflects how much "hotter" a location is compared to its immediate surroundings.

Boundary conditions  $\mathcal{B}[u]$ , defined later, describe how the function  $u$  behaves at the edges of the domain (e.g., the edges

of the 2D image scan). For example, they may specify that no flux enters or leaves through the boundary. Such boundary conditions are necessary to ensure the problem has a unique solution.

### C. Recovering Differential Operator

We represent the disease progression trajectory as a function  $u(x, t) : \mathbb{R}^{N+1} \rightarrow \mathbb{R}^M$  that maps spatio-temporal coordinates to the measurements. We consider the following parameterized non-linear boundary value problem:

$$\begin{aligned} \mathcal{N}[u, \lambda] &= 0, & \forall x \in \Omega, \forall t \in \mathcal{T}; \\ \mathcal{B}[u] &= 0, & \forall x \in \partial\Omega, \forall t \in \mathcal{T}; \end{aligned}$$

with solution  $u(x, t)$ , differential operator  $\mathcal{N}$  parameterized by  $\lambda$ , spatial and temporal coordinates  $x, t$ ,  $N$ -dimensional study domain  $\Omega \subseteq \mathbb{R}^N$  coupled with optional time dimension  $\mathcal{T} \subseteq \mathbb{R}^+$ , and operator of the boundary conditions  $\mathcal{B}$ . Additionally, we have a finite set of points  $\mathcal{I} \subseteq \Omega \times \mathcal{T}$  with known or measured values  $u(x, t) : x, t \in \mathcal{I}$ . Our goal is both to restore the coefficients  $\lambda$  of the differential operator  $\mathcal{N}$  and to approximate solution  $u$  with  $\hat{u}$ , given the observed data.

The approach introduced by Raissi et al. [7] as PINNs parameterizes  $\hat{u}_\Theta$  as a neural network with parameters  $\Theta$  and proposes to optimize the following functions on three finite sets of collocation points: a set with known values  $\mathcal{I}$ , set of

boundary points  $\mathcal{K} \subseteq \partial\Omega$  and points inside the domain  $\mathcal{J} \subseteq \Omega$ :

$$f(\Theta, \lambda) = \frac{1}{|\mathcal{I}|} \sum_{x,t \in \mathcal{I}} \|\hat{u}_\Theta(x, t) - u(x, t)\|_2^2 \quad (1)$$

$$+ \frac{1}{|\mathcal{J}|} \sum_{x,t \in \mathcal{J}} \|\mathcal{N}[\hat{u}_\Theta, \lambda](x, t)\|_2^2 \quad (2)$$

$$+ \frac{1}{|\mathcal{K}|} \sum_{x,t \in \mathcal{K}} \|\mathcal{B}[\hat{u}_\Theta](x, t)\|_2^2 \quad (3)$$

During training, (2) enforces the constraints imposed by the PDE, (1) corresponds to the observed data with known values, while (3) focuses on boundary conditions. The loss function  $f(\Theta, \lambda)$  is minimized with respect to the parameters  $\Theta$  of the neural network and coefficients  $\lambda$  of the PDE equation using a quasi-Newton full-batch gradient-based optimization method.

In this work, we revise the optimization part of the PINNs approach of recovering the unknown coefficients  $\lambda$  and neural network parameters  $\Theta$  in inverse problems. We propose to interpret this task as a constrained optimization problem defined as:

$$\begin{aligned} \min_{\Theta, \lambda} \sum_{i=1}^L \|\hat{u}_\Theta(x_i, t_i) - u(x_i, t_i)\|_2^2 \\ \text{subject to } \left[ \int_{\Omega} \|\mathcal{N}[\hat{u}_\Theta, \lambda]\|_2^2 dx \right] = 0 \\ \left[ \int_{\partial\Omega} \|\mathcal{B}[\hat{u}_\Theta]\|_2^2 dx \right] = 0 \end{aligned}$$

where we minimize Eq. 1, representing known measurements, under the constraints of setting the system of differential equations to zero. As the number of domain measurements is known beforehand, Eq. 1 is computed over a finite set of points. The constraints should be satisfied over the whole domain  $\Omega$ . To estimate the constraint values over  $\Omega$  we again rely on sampling, thus transforming the problem into:

$$\begin{aligned} \min_{\Theta, \lambda} \sum_{i=1}^L \|\hat{u}_\Theta(x_i, t_i) - u(x_i, t_i)\|_2^2 \quad (4) \\ \text{subject to } \left[ \begin{array}{c} \mathbb{E}_{x,t \sim \Omega} [\|\mathcal{N}^T[\hat{u}_\Theta](x, t)\lambda\|_2^2] \\ \mathbb{E}_{x,t \sim \partial\Omega} [\|\mathcal{B}[\hat{u}_\Theta](x, t)\|_2^2] \end{array} \right] = 0 \end{aligned}$$

under additional assumption, that the operator  $\mathcal{N}$  is linear with respect to  $\lambda$  and can be represented as  $\mathcal{N}^T[\hat{u}_\Theta]\lambda$ .

Having Eq. 4 in mind, we propose the Expectation-Maximization iterative method for estimating both parameters  $\Theta$  of the function  $\hat{u}_\Theta$  and parameters  $\lambda$  of the operator  $\mathcal{N}$ . In the proposed method, each iteration  $k$  can be decoupled into two steps: during the first step, we estimate the value of the PDE parameters  $\lambda$  by solving the least squares problem  $\lambda^* = \argmin_{\lambda} \mathbb{E}_{x,t \sim \Omega} [\|\mathcal{N}^T[\hat{u}_\Theta](x, t)\lambda\|_2^2]$  with exponential running average  $\lambda_k = (1 - \beta)\lambda^* + \beta\lambda_{k-1}$ , and construct unconstrained objective function

$$\begin{aligned} L_k(\Theta) = \sum_{i=1}^L \|\hat{u}_\Theta(x_i, t_i) - u(x_i, t_i)\|_2^2 + \\ \mu_0 \mathbb{E}_{x,t \sim \partial\Omega} [\|\mathcal{B}[\hat{u}_\Theta](x, t)\|_2^2] + \\ \mu_1 \mathbb{E}_{x,t \sim \Omega} [\|\mathcal{N}^T[\hat{u}_\Theta](x, t)\lambda_k\|_2^2], \end{aligned} \quad (5)$$

where  $\mu_0$  and  $\mu_1$  are scalar hyperparameters controlling strength of the constraints.

In the second step, we optimize function  $L_k$  w.r.t parameters  $\Theta$  using Adam optimizer. The method is described in Algorithm 1.

---

**Algorithm 1:** Proposed PINN-EM optimization procedure

---

**Set algorithm variables:**

$N$  - maximum total iteration count:  $10^5$   
 $N_\Theta$  - inner loop iteration count for  $L_k(\Theta)$ : 100  
 $\beta$  - exponential running average coefficient: 0.95  
 $\mu_0, \mu_1$  - hyperparameters in  $L_k(\Theta)$ : 0.1

**Initialize:**  $\Theta, \lambda$ , Adam state for parameters  $\Theta$

**while**  $k \leq N$  **do**

**Sample set of points**

$\mathcal{J}$  inside domain  $\Omega$

$\mathcal{K}$  on the boundary  $\partial\Omega$

    1. **Estimate**  $\lambda_k$  and **construct**  $L_k(\Theta)$

**Solve**  $\lambda^* = \argmin_{\lambda} \sum_{x \in \mathcal{J}} [\|\mathcal{N}^T[\Theta](x)\lambda\|_2^2]$

**Update** parameters  $\lambda_k = (1 - \beta)\lambda^* + \beta\lambda_{k-1}$

**Construct** function  $L_k(\Theta)$ :

$L_k(\Theta) = \sum_{i=1}^L \|\hat{u}_\Theta(x_i) - u(x_i)\|_2^2 +$

$\mu_0 \sum_{x \in \mathcal{K}} \|\mathcal{B}[\Theta](x)\|_2^2 +$

$\mu_1 \sum_{x \in \mathcal{J}} \|\mathcal{N}^T[\Theta](x)\lambda_k\|_2^2$

    2. **Optimize**  $L_k(\Theta)$  **using Adam optimizer**

        over  $N_\Theta$  iterations.

**end**

---

#### D. Training details

For training, we need to sample sets  $\mathcal{I}$ ,  $\mathcal{J}$  and  $\mathcal{K}$  from the input data. For the supervised dataset,  $\mathcal{I}$ , 90% of the points are uniformly sampled among all the annotated points available for training using the Sobol sequence [44]. The remaining points are sampled from Gaussian Mixture with variances equal to  $10^{-2}$  with centers at the locations, where the Laplacian of the initial condition is larger than the empirically selected value of 10. The Laplacian is computed using the finite difference approximation of derivatives. For the unsupervised dataset  $\mathcal{J}$ , the sampling domain is limited by the time  $\tilde{t}$ , which is defined dynamically and increasing from one epoch to another until it reaches its maximum value at half of the maximum number of training epochs  $\tilde{t} = t_{max} \cdot \min(2 \cdot \text{epoch}/\text{epoch}_{max}, 1)$ . Taking into account dynamic temporal domain extension, the points inside the domain are sampled uniformly using the Sobol sequence. For parametric function  $\hat{u}$  we select multi-layer perceptron with sine activation functions [31].

During training, we use  $10^5$  collocation points, with varying proportions of the points sampled from the initial conditions or domain measurements. Such proportion depends on the experiment we are performing. The points are resampled at each iteration unless an L-BFGS optimizer is used. For the L-BFGS optimizer, we additionally perform pretraining using the Adam optimizer with a learning rate of  $10^{-3}$  for 100 epochs with 1000 iterations each. The domain uniformly increases in the temporal dimension and reaches the requested time  $T$  by half of the epochs. We don't run pretraining for other optimizers. The overall compute budget is limited to 200

epochs or 2 hours of compute time on RTX A6000, whatever happens first. The experiments were performed using JAX framework.

### III. EXPERIMENTS

#### A. Dataset

In the experiments, we utilized two sources of data. First, we performed experiments in a controlled environment with synthetic data to validate the correctness of the PINN-EM method and highlight its advantages. Furthermore, we applied the proposed algorithm to the real-world problem on a dataset collected at the Medical University of Vienna.

1) *Synthetic data*: In the experiments with synthetic data, we utilize the open-source PDEBench dataset [45]. Since we are interested in using reaction-diffusion type models with non-linear dynamics, we limit ourselves to the following equations:

a) *Burger's 1D equation*: Defined with the equations below

$$\begin{aligned}\partial_t u &= \nu / \pi \partial_{xx} u - \partial_x (u^2/2), x \in (0; 1), t \in (0; 2], \\ u(x, 0) &= u_0(x), x \in (0; 1),\end{aligned}$$

with periodic boundary conditions and initial conditions generated randomly as a superposition of sine waves  $u_0(x) = \sum_{k_i=k_1, \dots, k_N} A_i \sin(k_i x + \phi_i)$ . For experiments, we utilize all 12 values of  $\nu$  provided as part of PDEBench dataset with 10 different initial conditions per  $\nu$  value. We consider time interval  $t \in (0; 1]$  as training data and reserve  $t \in (1; 2]$  as test data to evaluate the extrapolation capabilities of the methods. As a differential operator and parameter set, we select:

$$\begin{aligned}\lambda &= [\nu, -1]^T \\ \mathcal{N} &= [\frac{1}{\pi} \partial_{xx} u, \partial_x (u^2/2) + \partial_t u]^T\end{aligned}$$

b) *Reaction-Diffusion 2D equation*: The reaction-diffusion equation is defined as following:

$$\begin{aligned}\partial_t u &= D_u \Delta u + u - u^3 - k - v, \\ \partial_t v &= D_v \Delta v + u - v, \\ x &\in (-1; 1) \times (-1; 1), t \in (0; 5], \\ u(x, 0) &= u_0(x), x \in (-1; 1) \times (-1; 1),\end{aligned}$$

with Neumann boundary conditions and initial conditions generated as a random noise sampled from the normal distribution  $u_0(x) \sim \mathcal{N}(0, 1)$ . The values of parameters remain fixed at  $D_u = 10^{-3}$ ,  $D_v = 5 \cdot 10^{-3}$  and  $k = 5 \cdot 10^{-3}$ . In the experiments, we use 120 different initial conditions. For training, we limit the time interval to  $t \in (0; 4]$  and reserve remaining  $t \in (4; 5]$  as test data. As a differential operator and parameter set, we select:

$$\begin{aligned}\lambda &= [D_u, 0, 1, 1, -1, 0]^T \\ \mathcal{N} &= [\Delta u, \Delta v, u - v, -u^3 - k, \partial_t u, \partial_t v]^T\end{aligned}$$

The data was downloaded from the official repository available to the public. More details on generating the synthetic ground truth can be found in the work of Takamoto et al. [45].

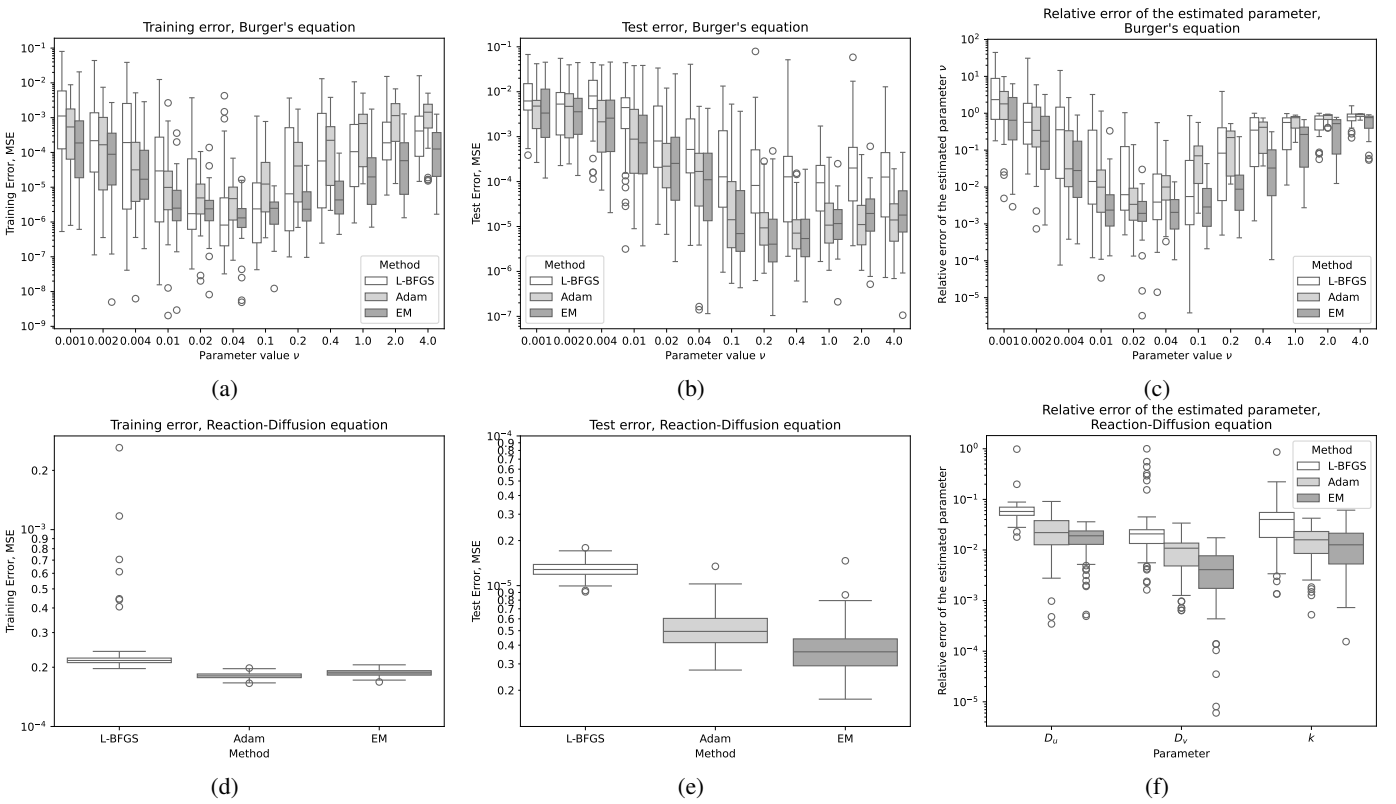
2) *AMD Imaging Data*: We used a longitudinal medical imaging dataset from a clinical observational study of patients with GA [46]. In this study, patients of at least 50 years were enrolled. The patients were diagnosed with GA secondary to AMD and followed for a minimum of 12 months. Patients were excluded if there was a history of other ocular diseases. Visits with insufficient FAF quality, visits with irreparable registration errors, eyes with less than 4 visits or without 1 year followup were also excluded. The dataset features the progression of 75 eyes of 51 patients with 553 scans in total. The average age of the patients is 78 years with IQR [73.66; 83.11]. The dataset has 39% male and 61% female patients affected by GA with predominant patterns listed in the Supplementary Table I. The patients are requested to perform visits 3 months apart. If the visit is missing, we ignore it, as the method doesn't require regular visits. If possible, both eyes are imaged and processed independently. In this retrospective study, we select the baseline visit such that it has at least 4 prior visits and 1 year follow-up visit. The data was split for each eye into the prior visits utilized during the model fitting, and the follow-up visits held for testing. We selected 1 year as the primary endpoint, as it corresponds to the highest number of visits (Supplementary Table II). Additional endpoints include 6 months, 18 months, and beyond 18 months. The dataset consists of 3D OCT scans of the retina with 49 B-scans, each the size of 1024x496 px. The size of the en-face projection is 49x1024 px covering the central  $6 \times 6 \text{ mm}^2$  macular region of the retina. The spatial alignment of the follow-up scans was performed using the image registration algorithm of Arikan et al. [47].

In the experiments, we use photoreceptor thickness and RPE integrity maps extracted automatically from 3D OCT scans. The thickness of photoreceptors is a continuous quantity measured in micrometers and is obtained with an image segmentation algorithm [40]. The second quantity is the RPE integrity, which is a binary quantity, measured as the presence (healthy) or absence (pathology) of the RPE cells indicating the presence of GA. The segmentation of the GA lesion is performed using an image segmentation algorithm [39]. An example of the data provided in a single visit is shown in Figure 1.

Following the section II-A, we select the reaction-diffusion equation as a constraint for the neural representation of the GA growth model. For GA growth, we define the set of parameters  $\lambda$ , differential operators  $\mathcal{N}$  and  $\mathcal{B}$  as following:

$$\begin{aligned}\lambda &= [\text{diag}(d^{2 \times 1}), A^{2 \times 2}, -1]^T \\ \mathcal{N} &= [\nabla^2 \hat{u}, \hat{u}(1 - \hat{u}), \partial_t \hat{u}]^T \\ \mathcal{B} &= \nabla \hat{u}(x, t) \cdot \hat{n}\end{aligned}$$

where  $\hat{n}$  is the normal directed outward from the boundary  $\partial\Omega$ . The set  $\mathcal{I}$  of measured values consists of initial conditions and available prior visits. The boundary conditions for the problem are assumed to have zero flux at the rectangular boundary of the domain. Provided the arguably narrow field of view of the retina scan centered at the macula (Figure 1a), we assume that no anatomical structures sufficiently influence the domain area.



**Fig. 3:** Comparison of different optimization methods for training PINNs for inverse problems. For Burger's equation (first row): (a) MSE on training time interval, evaluating interpolation capabilities, (b) MSE on testing time interval, evaluating extrapolation capabilities, (c) relative error of the estimated parameter depending on its value. For reaction-diffusion equation (second row): (d) MSE on training time interval, (e) MSE on testing time interval, (f) relative error of the three estimated parameters

## B. Experimental Design

In the experiments, we test the capabilities of the methods to solve inverse problems for PDE parameter estimation. In addition, we evaluate the abilities to robustly extrapolate provided data. In all experiments, we train separate models for each boundary value problem.

**Data split:** The proposed method is designed to model individual trajectories independently, thus we do not perform cross-validation. Instead, we censor future visits beyond the defined time interval. For example, we use  $t \in (0, 1]$  for training and  $t \in (1, 2]$  for testing. After training the model, we evaluate the estimated PDE coefficients, the interpolation performance on the training interval, and the extrapolation performance on the test interval. To this end, we report the relative error of the coefficients (synthetic data only), training Mean Squared Error (MSE), and test MSE.

For NODE [23], FNO [48], nnUNet [49], RNN [6] and Taylor [22] methods, we used 5 Fold CV and then pooled validation results.

**1) Synthetic Data:** For Burger's equation, we vary the parameter  $\nu$ , the number of in-domain samples with known measurements, and the initial conditions. Again, we consider  $t \in (0, 1]$  to be the training time interval and  $t \in (1, 2]$  as the time interval for testing. We compare the MSE of the solution to the respective ground truth for training  $MSE_{train}$

and testing  $MSE_{test}$  intervals, as well as relative error for parameter  $\nu$ :  $Rel_{\nu} = |\hat{\nu} - \nu|/\nu$ . In the reaction-diffusion equation, we consider  $t \in (0, 4]$  to be the training time interval and  $t \in (4, 5]$  as the testing time interval. We fix the three parameters  $D_u$ ,  $D_v$ ,  $k$  and change the initial conditions only. We track  $MSE_{train}$ ,  $MSE_{test}$ ,  $Rel_{D_u}$ ,  $Rel_{D_v}$ ,  $Rel_k$ .

**2) AMD Imaging Data:** In the experiments with GA, we train a single model separately for each patient. We vary the number of prior visits available for training, ranging from 2 to 4 visits. We use the Kolmogorov-Fisher reaction-diffusion equation [50] as a constraint for the INR, as discussed in section III-A.2.

**Metrics:** The model is evaluated on the task of predicting the segmentation map of GA at the primary endpoint of 12 months. In addition, we assessed performance at 6 months, 18 months, and beyond 18 months. The evaluation metrics include:

- Dice, Dice index of the GA lesion. This metric ranges  $[0, 1]$  (higher is better) and measures how well the predicted lesion overlaps with the ground truth;
- Dice growth, Dice index of newly converted GA lesions. This metric is computed only on regions that were not present at the previous visit and indicates how well the predicted growth overlaps with the observed growth (higher is better);

- MSE PR, Mean Squared Error of photoreceptors. This metric takes values in the range  $[0, \infty)$  (lower is better) and quantifies how close the predicted PR values are to the ground truth;
- Pearson's correlation  $r$  of GA growth rate (square root-transformed,  $mm/y$ ). This metric ranges from  $[0, 1]$  (higher is better) and indicates how well the predicted growth rate correlates with the observed rate;
- Coefficient of determination  $R^2$  of GA growth rate (square root-transformed,  $mm/y$ ). This metric ranges from  $(-\infty, 1]$  (higher is better) and measures how well the predicted growth values fit the observed ones. Negative values indicate that the predictions perform worse than a constant mean-growth baseline.

Individual metrics should be interpreted with care. For example,  $R^2$  is sensitive to outliers, while Dice can yield high values for large regions. A combination of metrics provides a more reliable assessment. In practice, we recommend focusing on  $r$  as the primary measure of performance,  $R^2$  for robustness and outlier sensitivity, Dice as a sanity check, and Dice growth for evaluating prediction of lesion progression.

**Clinical Significance:** To estimate the potential impact on clinical management, we evaluated how well the methods can identify fast-progressing eyes - those at highest risk. For a given time interval, we defined fast progressors as eyes with a GA growth rate above the top 25% quantile threshold, estimated from the entire dataset. In the experiments, an eye was counted as a fast progressor if the tested method predicted a GA growth rate above this threshold. To assess performance on this classification task, we report F-score, sensitivity, and specificity.

### C. Statistical Analysis

For paired data, e.g. Dice scores, MSE and Rel, we employed one-tailed paired sample t-test. For  $r$  and  $R^2$  we performed one-tailed independent sample t-test over 1000 bootstrap estimates. Asterisk (\*) denotes significance value  $p < 0.05$ . In GA experiments, we calculate the metrics assuming independence of the eye. However, this is not always the case. In the supplementary materials, we extend the evaluation by addressing the intra-patient correlation (Supplementary Table V).

### D. Baselines

1) **Synthetic Data:** For experiments with synthetic data, we employed two baseline PINNs trained with different optimization methods (L-BFGS and Adam). These benchmarks serve as a tool to evaluate the proposed method on the inverse problems and to compare the performance with the existing methodologies in training PINNs.

2) **AMD Imaging Data:** For experiments with GA patients, we compared the proposed method with two groups of methods: those trained with and without posterior knowledge. The Finite Elements Method (FEM), Linear Regression, and Linear Growth models were implemented in both variants. In the normal variant, models were fitted to the training data and required to extrapolate, whereas in the "Oracle"

**TABLE I:** Reaction-Diffusion's Equation. Comparison between different optimization methods. *MSE* stands for Mean Squared Error, *Rel* is a relative error of the PDE coefficient, multiplied by the corresponding factor. \* denotes that the result is significantly better than all of the baselines.

| Method        | $MSE \cdot 10^{-5}$ |                  | $Rel \cdot 10^{-2}$ |          |         |
|---------------|---------------------|------------------|---------------------|----------|---------|
|               | Train               | Test             | $D_u$               | $D_v$    | $k$     |
| PINN (L-BFGS) | $28.1 \pm 31.1$     | $19.5 \pm 99.1$  | 7.19                | 5.89     | 5.86    |
| PINN (Adam)   | $18.1 \pm 0.6$      | $0.53 \pm 0.2$   | 2.79                | 1.04     | 1.67    |
| PINN (EM)     | $18.7 \pm 0.8$      | $0.38 \pm 0.2^*$ | $1.8^*$             | $0.51^*$ | $1.5^*$ |

variant, models had access to test data containing future visits to indicate ceiling performance. Other methods had no access to test data. Simple constant prediction (baseline), nnUNet [49], Recurrent Neural Network (RNN) [51], Fourier Neural Operator (FNO) [48], NeuralODE [23], and Taylor [22] required only a single baseline scan for inference. Linear Regression, Linear Growth, FEM, and PINNs used the full set of visits listed in Table II. Linear Regression was fitted to the root-transformed GA lesion area, while the Linear Growth model assumed constant-speed lesion expansion in the normal direction. The FEM model used all available visits with the same reaction-diffusion equation as PINNs (Section III-A.2), implemented using FEniCS and dolfin-adjoint frameworks. We implemented nnUNet v2 [49] using official github repository<sup>1</sup> for primary endpoint of 1 year. We generated the training data as image pairs: baseline SLO image + GA annotation at baseline and 1 year. We reimplemented RNN described by Nguyen et al. [51] and adopted it to handle imaging data of patients with Geographic Atrophy by introducing convolutions. We reimplemented the work of Gigon et al. [22] utilizing Taylor series decomposition, thus we named it Taylor. Our reimplementation follows the original paper, except that we employed the publicly available method of Garvin et al. [52] for layer segmentation. We also implemented the FNO as described by Li et al. [48] and trained it on GA data by sampling random pairs of visits that are one year apart. Finally, we trained the NeuralODE model using the GA imaging data as described in [23].

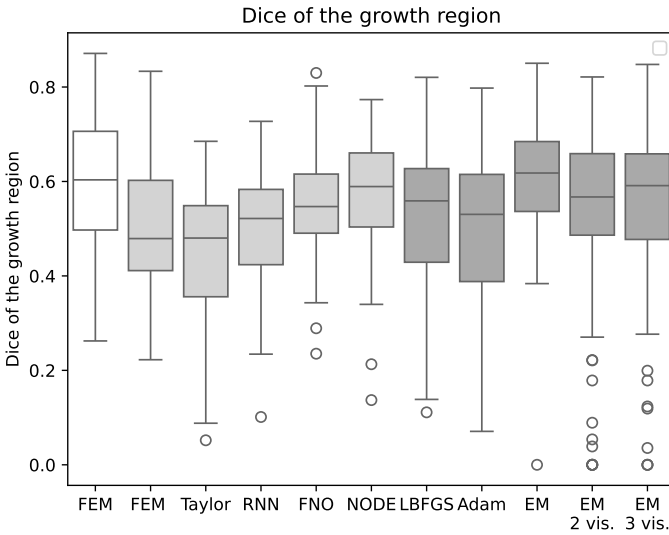
## IV. RESULTS

1) **Synthetic Data:** The results of the experiments with Burger's equation are summarized in Fig. 3a-c. We compared three different optimization algorithms in this ablation study. Interestingly, in all experiments, the quasi-Newton L-BFGS method demonstrates worse results compared to the first-order methods. This can be attributed to the feature of L-BFGS, which requires full batch training, while Adam and the proposed EM utilize randomly sampled mini-batches, giving a better approximation to the mean estimate in Eq. 4. Comparing Adam and EM, the latter often produces more accurate parameter estimates, thus, increasing both interpolation and extrapolation performance. Supplementary Table III further highlights the advantages of the proposed method in parameter estimation. The values in the table are averaged across multiple initial conditions and parameter values. The proposed method,

<sup>1</sup><https://github.com/MIC-DKFZ/nnUNet>

**TABLE II:** Results on estimating spatio-temporal progression of Geographic Atrophy. The + signifies the visits used for inference, e.g., 3 visits including the baseline, and the numbers represent the number of days between visits in a relative time frame; \* denotes that the result is significantly better than all of the baselines.

| Model                           | -540 | -360 | -180 | 0 | 360 | Test: Day 360 |       |                 |                   |                       |
|---------------------------------|------|------|------|---|-----|---------------|-------|-----------------|-------------------|-----------------------|
|                                 |      |      |      |   |     | r             | $R^2$ | Dice            | Dice growth       | MSE (PR)              |
| No Change                       |      |      |      | + |     | n/a           | -2.15 | $0.87 \pm 0.08$ | 0.0               | 0.0138                |
| <i>Data-driven</i>              |      |      |      |   |     |               |       |                 |                   |                       |
| Linear Regression               | +    | +    | +    | + |     | 0.76          | 0.51  | n/a             | n/a               | n/a                   |
| Linear Growth                   | +    | +    | +    | + |     | 0.67          | 0.06  | $0.91 \pm 0.06$ | $0.39 \pm 0.26$   | n/a                   |
| nnUNet [49]                     |      |      |      | + |     | 0.07          | -1.38 | $0.85 \pm 0.14$ | $0.45 \pm 0.15$   | n/a                   |
| RNN [51]                        |      |      |      | + |     | 0.57          | -0.54 | $0.87 \pm 0.09$ | $0.50 \pm 0.12$   | n/a                   |
| Taylor [22]                     |      |      |      | + |     | 0.62          | -0.45 | $0.85 \pm 0.13$ | $0.45 \pm 0.14$   | n/a                   |
| FNO [48]                        |      |      |      | + |     | 0.65          | 0.35  | $0.91 \pm 0.05$ | $0.55 \pm 0.11$   | $0.0109 \pm 0.0054$   |
| NODE [23]                       |      |      |      | + |     | 0.68          | 0.23  | $0.90 \pm 0.07$ | $0.57 \pm 0.12$   | n/a                   |
| <i>Physical</i>                 |      |      |      |   |     |               |       |                 |                   |                       |
| FEM                             | +    | +    | +    | + |     | 0.65          | 0.14  | $0.89 \pm 0.08$ | $0.50 \pm 0.13$   | $0.0108 \pm 0.0056$   |
| <i>Proposed (Hybrid)</i>        |      |      |      |   |     |               |       |                 |                   |                       |
| PINN (L-BFGS)                   | +    | +    | +    | + |     | 0.76          | 0.48  | $0.91 \pm 0.05$ | $0.52 \pm 0.15$   | $0.0089 \pm 0.0047$   |
| PINN (Adam)                     | +    | +    | +    | + |     | 0.79          | 0.56  | $0.91 \pm 0.06$ | $0.50 \pm 0.16$   | $0.0085 \pm 0.0045$   |
| PINN (EM)                       | +    | +    | +    | + |     | 0.83*         | 0.56  | $0.91 \pm 0.10$ | $0.61 \pm 0.12^*$ | $0.0081 \pm 0.0045^*$ |
| PINN (EM)                       |      | +    | +    | + |     | 0.62          | 0.41  | $0.87 \pm 0.19$ | $0.54 \pm 0.20$   | $0.0081 \pm 0.0045^*$ |
| PINN (EM)                       |      |      | +    | + |     | 0.47          | -1.41 | $0.87 \pm 0.18$ | $0.51 \pm 0.21$   | $0.0084 \pm 0.0045$   |
| <i>With posterior knowledge</i> |      |      |      |   |     |               |       |                 |                   |                       |
| Linear Regression               | +    | +    | +    | + | +   | 0.91          | 0.83  | n/a             | n/a               | n/a                   |
| Linear Growth                   | +    | +    | +    | + | +   | 0.94          | 0.86  | $0.92 \pm 0.04$ | $0.52 \pm 0.20$   | n/a                   |
| FEM                             | +    | +    | +    | + | +   | 0.97          | 0.92  | $0.91 \pm 0.05$ | $0.59 \pm 0.15$   | $0.0091 \pm 0.0048$   |



**Fig. 4:** Results on estimating the Geographic Atrophy growth region over the one-year period. The white box highlights the method had access to posterior data, light grey boxes denote prior art methods, dark gray boxes highlight physics informed methods, and vis. stands for the number of prior visits used.

using just 128 training samples is similar or often outperforms the baselines that use 8192 training samples.

For the reaction-diffusion equation, we observe the same trend as in the previous experiment, where the quasi-Newton method falls behind first-order methods in all experiments. The proposed method again excelled in the parameter estimation, surpassing the baselines. It can be assumed that more accurate PDE parameter estimates lead to better performance in interpolation and extrapolation. Table I highlights the superiority of

the proposed optimization method for the multiple parameter estimation. Overall, the ablation experiments with synthetic data demonstrate the applicability of the proposed method to inverse problems and its great potential in further applications on clinical data.

**2) AMD Imaging Data:** The qualitative examples of the prediction outputs for primary endpoint are shown in Fig. 5. For quantitative comparison (Table II, Fig. 4 and Supplementary Table V), we additionally included the models trained using the posterior data. Such overfitted results indicate the ceiling performance for the disease progression methods. The overfitted FEM and Linear Growth model was able to achieve Dice index of the growth region of 0.59 and 0.52 respectively, while having very few outliers and producing a high  $R^2$  index of 0.92. The proposed model, while having access only to the prior visits, was able to surpass the performance of overfitted FEM and Linear Growth in Dice index of the growth region achieving a score of 0.61. This can be attributed to multiple factors such as soft constraints imposed on the implicit representation model in contrast to FEM with hard constraints, suitability of the reaction-diffusion equation for modeling GA progression. The FEM is explicitly designed for solving PDEs with error as low as possible, a few magnitudes lower than PINNs can achieve. However, this downside of PINNs serves as a great advantage for disease progression modeling because of two reasons. First, we don't know the exact model of disease progression, thus closely following the reaction-diffusion equation can deviate the solution from the observed values. On the second note, PINNs can fit the observed data by modifying initial conditions, for example introducing a small perturbation that will eventually evolve into new lesions, while FEM with coupled reaction-diffusion equation has no means of explaining new lesions if they were

not present in the initial conditions. Overall, the proposed optimization method achieved the highest Dice score of the growth region compared to the prior art methods, even when considering the methods that utilize the posterior information.

Analyzing the  $r$  and  $R^2$  metrics, we can observe that PINN models achieve the highest score with the proposed method excelling in  $r$  metric. A high number of outliers affect the  $R^2$  results of the model utilizing only two prior visits, often leading to negative values. The most extreme outlier is shown in Supplementary Fig. 3. When disease dynamics are inferred from just two visits spaced 60 days apart, the method becomes prone to errors caused by over- or underestimation. In the presented outlier case, rapid progression occurring within the preceding 60 days, which the proposed method attempted to reproduce, strongly influenced the extrapolation results. Providing more prior visits improved the results for this particular outlier, though performance still remained far from perfect. In general, however, the availability of additional training data leads to a substantial improvement in performance. For the Dice score of the one-year lesion progression, the performance eventually saturates at 0.91 for all the methods. This can be attributed to the noise present in the ground truth and uncertainty involved in the disease progression modeling process. The MSE scores for predicting photoreceptor thickness remain rather high for all of the models that utilize this information, however, the proposed method achieved the lowest error.

When analyzing the extended results with additional study endpoints at 6, 18, and beyond 18 months (Supplementary Table IV; Supplementary Fig. 2 and and Supplementary Fig. 4), we observe that extrapolation over shorter time intervals presents a particularly challenging task. The proposed method follows this trend; however, it remains among the best-performing approaches across all time points. For longer intervals, such as beyond 18 months, the proposed method substantially outperforms the baselines, achieving a Dice growth of 0.69. As shown in Supplementary Fig. 2, it also exhibits higher metric values and lower variance compared to other methods.

In the experiment with fast GA progressing eyes, we aimed to highlight the clinical relevance of the proposed method for patient management (Supplementary Table VI). At the primary endpoint of 1 year, the proposed method achieved the highest performance among all approaches, with an F1 score of 0.78, sensitivity of 0.79, and specificity of 0.93. Across different time intervals, simpler methods such as Linear Regression and Linear Growth also performed strongly, in some cases surpassing baseline models. Notably, at longer forecasting horizons (540 days), the proposed method successfully identified 100% of patients with rapidly progressing GA.

## V. DISCUSSION

*Alternative PDE Families:* We use a system of reaction-diffusion PDEs to constrain the implicit neural representation of each patient's GA progression. Despite employing a simple Fisher-Kolmogorov model, the INR inferred via PINNs successfully extrapolated expected growth. The method could be enhanced with a more suitable PDE constraint, with

anisotropic diffusion [9] guided by patient-specific anatomy as a natural extension, and data-driven PDE discovery [53], [54] presenting a promising approach.

*Alternative Boundary Conditions:* In this work, we model a  $6 \times 6$  mm<sup>2</sup> section of the retina centered on the macula, using Neumann boundary conditions with zero flux along the domain boundary. We believe that Neumann boundary conditions are the best fit, as e.g. Dirichlet BC may potentially bias the results, due to presence of cases where GA grows large enough that it traverses the boundary. The choice of Neumann boundary conditions reflects the simplified design and absence of major anatomical structures near the scan, except for the optic nerve head, present in only a minority of cases. While incorporating vessel boundaries or optic nerve head geometry [55] could improve realism, it would also increase system complexity. As our focus is on evaluating PINNs for GA modeling, exploration of more complex boundary conditions is left for future work.

*Population level Models:* The proposed method models individual patients' trajectories without relying on population-level statistics. Each patient has a personalized parametric neural network, fine-tuned as new visits become available. While this increases adaptability, it ignores global trends. One potential improvement is to first infer a population "average" trajectory and then fine-tune it using patient-specific data [56]. A more sophisticated approach could introduce a parameter  $\gamma$  in  $u(x, t, \gamma)$  to encode the trajectory [57], updating only  $\gamma$  during fine-tuning to identify the best-fitting trajectory. We leave these extensions for future research.

*Uncertainty Quantification:* The reliability and explainability of the method are critical for clinical translation. Beyond achieving high accuracy, reporting uncertainty is essential for clinical decision support systems [58]. This can be addressed via ensembling, dropout, test-time augmentation, or Bayesian neural networks [59]. In particular, our method could be extended with a Bayesian neural network [60], following the B-PINNs framework [61]. Evaluation procedure can also benefit from uncertainty quantification with ground truth produced by multiple graders coupled with inter-observer variability estimation. Despite this, constrained resources made us refrain from producing such annotation. Accounting towards the works [62] comparing AI and Human experts for GA progression prediction, we can deduct that AI methods typically has lower or comparable variability. While recognizing the benefits of uncertainty estimation, we consider it a current limitation and an important direction for future research.

*Runtime requirements:* A limitation of the proposed method is that each eye requires a separate training run, resulting in forecasting latencies of 15 minutes on consumer hardware or 4 minutes on server-grade hardware. This may hinder clinical use. Workload could be reduced by optimizing FLOPs per backward pass or using next-generation hardware, while a more robust long-term solution would be adopting population-level models, as discussed earlier.

*Variable Number of Visits:* Variable and irregular sampling intervals still able to affect performance, despite the method is inherently continuous and makes no assumptions about visit timing. This can be seen on the performance when only 2 visits

are used. In ablation experiments, the method performed well on longer sequences but was unsuitable with only two visits. Applicability could be improved by interpolating intermediate visits or ensembling forecasts with NODE [23] or FNO [48]. Ultimately, integrating FNO and PINNs into a unified framework [63] with test-time fine-tuning may overcome this limitation, which we leave for future work.

**Domain Shift:** Although the proposed algorithm is a generic modeling tool that can be extended beyond GA, its validation is limited by available data. We expect performance to remain stable under domain shifts, such as differences in segmentation, provided all visits are segmented with the same tool and residuals are normally distributed. This also applies to changes in OCT scanner type, which may affect segmentation quality. We hypothesize the method is robust under these conditions, however, no quantitative domain-shift analysis has yet been performed.

**Translation to Patient Management:** In clinical settings, the proposed method currently has limited use, mainly for identifying patients requiring urgent care. Its applicability is expected to improve with uncertainty quantification, population-level knowledge, and personalized diffusion models with lower computational demands.

## VI. CONCLUSION

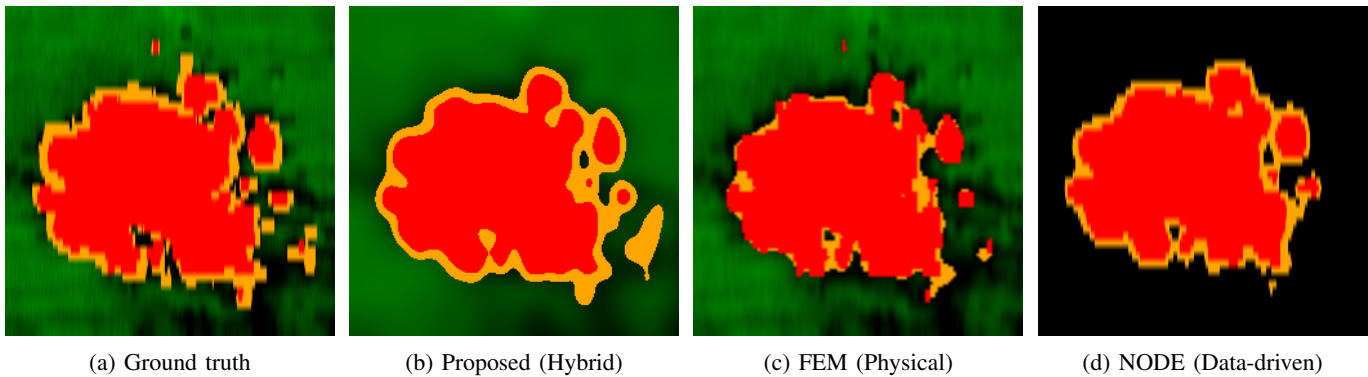
In this paper, we addressed spatio-temporal disease progression modeling from multiple prior visits by learning an Implicit Neural Representation constrained by a parameterized PDE. Our contributions include an optimization method that jointly trains a physics-informed model and recovers the coefficients of the PDE constraints; and a novel application to retinal imaging that introduces a previously unexplored PDE-constrained optimization for GA progression.

Applied to GA progression, the method captured observed data, extrapolated spatial changes, and outperformed existing approaches. Large-scale synthetic and clinical experiments confirmed its robustness, particularly for 1-year predictions. Key limitations remain: it requires at least two visits (three for better accuracy), per-patient fine-tuning with high computational cost, a potentially suboptimal PDE constraint, and the absence of uncertainty quantification and population-level information needed for clinical integration.

Future work includes recovering the PDE system that best describes GA dynamics, incorporating eye geometry through anisotropic diffusion and anatomy-aware boundary conditions, including a study validating the effectiveness of anisotropic diffusion; introducing population-level models or integrating population information during fine-tuning; extending the method with uncertainty quantification by incorporating *B-PINNs* as a natural extension of the framework; reducing computational burden by lifting the need for extensive per-patient fine-tuning, while keeping the requirements of the new features, e.g. *B-PINNs*, under reasonable constraint; and validating the vendor and device output variability through artificial image perturbations and larger multi-center multi-device cohort. The proposed method demonstrates strong potential for accurately modeling and forecasting GA progression, providing a foundation for future research.

## REFERENCES

- [1] M. Peterson and M. Riggs, "Predicting nonlinear changes in bone mineral density over time using a multiscale systems pharmacology model," *CPT: Pharmacometrics & Systems Pharmacology*, vol. 1, 2012.
- [2] S. Habib *et al.*, "Complex dynamics of tumors: modeling an emerging brain tumor system with coupled reaction-diffusion equations," *Physica A: Statistical Mechanics and Its Applications*, vol. 327, pp. 501–524, 2003.
- [3] A. Exuzides *et al.*, "Statistical modeling of disease progression for chronic obstructive pulmonary disease using data from the eclipse study," *Medical Decision Making*, vol. 37, pp. 453 – 468, 2017.
- [4] C. Zoccali *et al.*, "Left ventricular mass monitoring in the follow-up of dialysis patients: Prognostic value of left ventricular hypertrophy progression," *Kidney International*, vol. 65, no. 4, pp. 1492–1498, 2004.
- [5] C. K. Fisher *et al.*, "Machine learning for comprehensive forecasting of alzheimer's disease progression," *Scientific Reports*, vol. 9, 2018.
- [6] G. Lee *et al.*, "Predicting alzheimer's disease progression using multi-modal deep learning approach," *Scientific Reports*, vol. 9, 2019.
- [7] M. Raissi, P. Perdikaris, and G. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computational Physics*, vol. 378, pp. 686–707, 2019.
- [8] H. B. Friboes *et al.*, "An integrated computational/experimental model of tumor invasion," *Cancer research*, vol. 66 3, pp. 1597–604, 2006.
- [9] A. Swan *et al.*, "A patient-specific anisotropic diffusion model for brain tumour spread," *Bulletin of Mathematical Biology*, vol. 80, pp. 1259 – 1291, 2017.
- [10] Fernández-Romero, Antonio, Guillén-González, Francisco, and Suárez, Antonio, "A glioblastoma pde-ode model including chemotaxis and vasculature," *ESAIM: M2AN*, vol. 56, no. 2, pp. 407–431, 2022.
- [11] M. Lê *et al.*, "Personalized radiotherapy planning based on a computational tumor growth model," *IEEE Transactions on Medical Imaging*, vol. 36, pp. 815–825, 2017.
- [12] E. Konukoglu *et al.*, "Personalization of reaction-diffusion tumor growth models in mr images: Application to brain gliomas characterization and radiotherapy planning," in *Multiscale Cancer Modeling*. CRC Press, 2010.
- [13] I. Ezhov *et al.*, "Neural parameters estimation for brain tumor growth modeling," in *MICCAI 2019*, 2019, pp. 787–795.
- [14] S. Subramanian *et al.*, "Where did the tumor start? an inverse solver with sparse localization for tumor growth models," *Inverse Problems*, vol. 36, 2019.
- [15] M. Morris *et al.*, "Learning a classification-based glioma growth model using mri data," *Journal of Computers*, vol. 1, pp. 21–31, 2006.
- [16] L. Zhang *et al.*, "Convolutional invasion and expansion networks for tumor growth prediction," *IEEE Transactions on Medical Imaging*, vol. 37, pp. 638–648, 2018.
- [17] X. Shi *et al.*, "Convolutional lstm network: A machine learning approach for precipitation nowcasting," in *Neural Information Processing Systems*, 2015.
- [18] L. Zhang *et al.*, "Spatio-temporal convolutional lstms for tumor growth prediction by learning 4d longitudinal patient data," *IEEE Transactions on Medical Imaging*, vol. 39, no. 4, pp. 1114–1126, 2020.
- [19] J. Petersen *et al.*, "Continuous-time deep glioma growth models," in *MICCAI*, 2021.
- [20] M. Garnelo *et al.*, "Neural processes," *ArXiv*, vol. abs/1807.01622, 2018.
- [21] Y. Zhang *et al.*, "An integrated time adaptive geographic atrophy prediction model for sd-oct images," *Medical Image Analysis*, vol. 68, p. 101893, 2020.
- [22] A. Gigon *et al.*, "Personalized atrophy risk mapping in age-related macular degeneration," *Translational Vision Science & Technology*, vol. 10, 2021.
- [23] D. Lachinov *et al.*, "Learning spatio-temporal model of disease progression with neuralodes from longitudinal volumetric data," *IEEE Transactions on Medical Imaging*, vol. 43, no. 3, pp. 1165–1179, 2024.
- [24] T. Q. Chen *et al.*, "Neural ordinary differential equations," in *Neural Information Processing Systems*, 2018.
- [25] C. Liu *et al.*, "Imageflownet: Forecasting multiscale image-level trajectories of disease progression with irregularly-sampled longitudinal medical images," 2024.
- [26] P. Kidger *et al.*, "Neural sdes as infinite-dimensional gans," in *International Conference on Machine Learning*, 2021.
- [27] C. Wu *et al.*, "A comprehensive study of non-adaptive and residual-based adaptive sampling for physics-informed neural networks," *Computer Methods in Applied Mechanics and Engineering*, vol. 403, 2023.



**Fig. 5:** Qualitative comparison of the Geographic Atrophy (GA) progression models. GA at baseline visit is depicted in red, GA progression over one year is in orange, and photoreceptor thickness at the baseline is shown in green. NODE utilizes raw imaging data, hence it has no photoreceptor component. The animated version can be found in the supplementary materials.

- [28] S. Wang, S. Sankaran, and P. Perdikaris, "Respecting causality is all you need for training physics-informed neural networks," *ArXiv*, vol. abs/2203.07404, 2022.
- [29] A. D. Jagtap, K. Kawaguchi, and G. E. Karniadakis, "Adaptive activation functions accelerate convergence in deep and physics-informed neural networks," *Journal of Computational Physics*, vol. 404, p. 109136, 2020.
- [30] H. Wang *et al.*, "Learning specialized activation functions for physics-informed neural networks," *ArXiv*, vol. abs/2308.04073, 2023.
- [31] V. Sitzmann *et al.*, "Implicit neural representations with periodic activation functions," in *NIPS*, vol. 33, 2020, pp. 7462–7473.
- [32] S. Shaier, M. Raissi, and P. Seshaiyer, "Data-driven approaches for predicting spread of infectious diseases through dinns: Disease informed neural networks," 2021.
- [33] Z. Zhang *et al.*, "Discovering a reaction–diffusion model for alzheimer's disease by combining PINNs with symbolic regression," *Comput. Methods Appl. Mech. Eng.*, vol. 419, 2024.
- [34] R. Z. Zhang *et al.*, "Personalized predictions of glioblastoma infiltration: Mathematical models, physics-informed neural networks and multimodal scans," *ArXiv*, 2023.
- [35] N. M. Schultz *et al.*, "Global burden of dry age-related macular degeneration: A targeted literature review," *Clinical Therapeutics*, vol. 43, no. 10, pp. 1792–1818, 2021.
- [36] J. S. Heier *et al.*, "Pegcetacoplan for the treatment of geographic atrophy secondary to age-related macular degeneration (oaks and derby): two multicentre, randomised, double-masked, sham-controlled, phase 3 trials," *The Lancet*, vol. 402, no. 10411, pp. 1434–1448, 2023.
- [37] A. M. Khanani *et al.*, "Efficacy and safety of avacincaptad pegol in patients with geographic atrophy (gather2): 12-month results from a randomised, double-masked, phase 3 trial," *The Lancet*, vol. 402, no. 10411, pp. 1449–1458, 2023.
- [38] S. Vujosevic *et al.*, "Epidemiology of geographic atrophy and its precursor features of intermediate age-related macular degeneration," *Acta Ophthalmologica*, vol. 101, no. 8, pp. 839–856, 2023.
- [39] D. Lachinov *et al.*, "Projective skip-connections for segmentation along a subset of dimensions in retinal oct," in *MICCAI*, 2021, pp. 431–441.
- [40] J. I. Orlando *et al.*, "Automated quantification of photoreceptor alteration in macular disease using optical coherence tomography and deep learning," *Scientific Reports*, vol. 10, 2020.
- [41] P. A. Roberts, "Mathematical models of retinitis pigmentosa: The trophic factor hypothesis," *J. Theor. Biol.*, vol. 534, Feb. 2022.
- [42] S. Riedl *et al.*, "The effect of pegcetacoplan treatment on photoreceptor maintenance in geographic atrophy monitored by artificial intelligence–based oct analysis," *Ophthalmology Retina*, vol. 6, no. 11, pp. 1009–1018, 2022.
- [43] U. Schmidt-Erfurth *et al.*, "Disease activity and therapeutic response to pegcetacoplan for geographic atrophy identified by deep learning-based analysis of oct," *Ophthalmology*, vol. 132, pp. 181–193, 2025.
- [44] I. Sobol', "On the distribution of points in a cube and the approximate evaluation of integrals," *USSR Computational Mathematics and Mathematical Physics*, vol. 7, no. 4, pp. 86–112, 1967.
- [45] M. Takamoto *et al.*, "PDEBENCH: An extensive benchmark for scientific machine learning," 2024.
- [46] P. Bui *et al.*, "Fundus autofluorescence and optical coherence tomography biomarkers associated with the progression of geographic atrophy secondary to age-related macular degeneration," *Eye*, vol. 36, pp. 2013 – 2019, 2021.
- [47] M. Arikan *et al.*, "Deep learning based multi-modal registration for retinal imaging," in *IMMIC 2019, MICCAI*, 2019, p. 75–82.
- [48] Z.-Y. Li *et al.*, "Fourier neural operator for parametric partial differential equations," *ArXiv*, vol. abs/2010.08895, 2020.
- [49] F. Isensee *et al.*, "nnU-net revisited: A call for rigorous validation in 3D medical image segmentation," in *MICCAI 2024*. Springer Nature Switzerland, 2024, pp. 488–498.
- [50] V. M. Tikhomirov, "A study of the diffusion equation with increase in the amount of substance, and its application to a biological problem," 1991.
- [51] M. Nguyen *et al.*, "Predicting Alzheimer's disease progression using deep recurrent neural networks," *NeuroImage*, vol. 222, p. 117203, 2020.
- [52] Garvin *et al.*, "Automated 3-d intraretinal layer segmentation of macular spectral-domain optical coherence tomography images," *IEEE Transactions on Medical Imaging*, vol. 28, no. 9, pp. 1436–1447, 2009.
- [53] S. H. Rudy *et al.*, "Data-driven discovery of partial differential equations," *Science Advances*, vol. 3, 2016.
- [54] K. Kacprzyk, Z. Qian, and M. van der Schaar, "D-cipher: discovery of closed-form partial differential equations," in *Proceedings of the 37th International Conference on Neural Information Processing Systems*, ser. NIPS '23, 2023.
- [55] I. A. Sigal *et al.*, "Reconstruction of human optic nerve heads for finite element modeling," *Technology and Health Care*, vol. 13, no. 4, pp. 313–329, 2005.
- [56] Y. Wang *et al.*, "Transfer learning in physics-informed neural networks: Full fine-tuning, lightweight fine-tuning, and low-rank adaptation," *International Journal of Mechanical System Dynamics*, vol. 5, no. 2, pp. 212–235, 2025.
- [57] X. Huang *et al.*, "Meta-auto-decoder for solving parametric partial differential equations," in *Proceedings of the 36th International Conference on Neural Information Processing Systems*, ser. NIPS '22, 2022.
- [58] U.S. FDA, "Artificial intelligence program: Research on AI/ML-based medical devices," 2024.
- [59] B. Lambert *et al.*, "Trustworthy clinical ai solutions: A unified review of uncertainty quantification in deep learning models for medical image analysis," *Artificial Intelligence in Medicine*, vol. 150, p. 102830, 2024.
- [60] E. Goan and C. Fookes, *Bayesian Neural Networks: An Introduction and Survey*. Springer, 2020, pp. 45–87.
- [61] L. Yang, X. Meng, and G. E. Karniadakis, "B-pinns: Bayesian physics-informed neural networks for forward and inverse pde problems with noisy data," *Journal of Computational Physics*, vol. 425, p. 109913, 2021.
- [62] G. S. Reiter *et al.*, "A novel management challenge in age-related macular degeneration: Artificial intelligence and expert prediction of geographic atrophy," *Ophthalmology Retina*, vol. 9, no. 5, pp. 421–430, 2025.
- [63] Z. Li *et al.*, "Physics-informed neural operator for learning partial differential equations," *ACM / IMS J. Data Sci.*, vol. 1, no. 3, May 2024.