

電気インピーダンストモグラフィ問題の計測データを用いた囲い込みニューラルネットワークの適用

Enclosure Neural Networks for Experimental Electrical Impedance Tomography

井手 貴範 (Takanori Ide)¹, Samuli Siltanen²,
¹ 城西大学 (Josai University), ²University of Helsinki

e-mail : tide@josai.ac.jp

1 概要

Electrical Impedance Tomography(EIT) は物体の表面に電圧を与えた時, 物体表面に発生した電流の計測値から内部包含物の位置を画像で再構成する非破壊検査法として知られている. 医療画像診断の応用先として, 人体における癌の位置の同定がある. EIT は次の偏微分方程式の逆問題として記述することができる [1]. 特に, \mathbb{R}^2 内の物体 Ω (有界集合) の電気伝導度を $\sigma = \sigma(x, y)$ とし $u = u(x, y)$ を電圧のポテンシャルとすると次の偏微分方程式で記述される数理モデルは EIT の連続体モデルと呼ばれる.

$$\begin{cases} \nabla \cdot \sigma \nabla u = 0 \text{ in } \Omega, \\ u = f \text{ on } \partial\Omega. \end{cases} \quad (1)$$

ここで, n を境界 ($\partial\Omega$) 上の単位法線ベクトルとし, 以下のような Dirichlet–Neumann 写像を与える.

$$\Lambda : f \rightarrow \sigma \frac{\partial u}{\partial n} \Big|_{\partial\Omega} \quad (2)$$

また, 電気伝導度を次のように与える.

$$\sigma(x, y) = 1 + \chi_D(x, y)h(x, y) \quad (3)$$

ここで, $D \subset \Omega$ を内部包含物の集合とし, 次式が成り立つとする.

$$\chi_D(x, y) = \begin{cases} 1, & (x, y) \in D \\ 0, & (\text{その他}) \end{cases}. \quad (4)$$

この時, Dirichlet–Neumann 写像 Λ から内部包含物 D を再構成できるのかという Calderón の問題を解くために, 数理解析的な再構成法として知られている囲い込み法 [2] とニューラルネットワークを融合した手法を提案した [3]. 提案した方法は現実の測定データに適用するためには現象を支配する数理モデルを変更する必要がある. 本公演では完全電極モデルと呼ばれる偏微分方程式で記述される EIT の問題を考え, 囲い込みニューラルネットワークを現実の測定データに適用する手法と結果を紹介する.

2 完全電極体モデル

実験における計測データは雑音を含む。また、所望する箇所のデータを全て取得することが出来ないことが多い。つまり、不完全な情報を用いて内部包含物の情報を正確に抽出するのは極めて困難な問題であり、理論と実用の観点から極めて重要である。本講演において、有限個の観測データから内部包含物を再構成する囲い込みニューラルネットワークを紹介する。

EIT の実験においても物体の周囲全体に測定装置である電極をつけることは難しい。そのため、物体の周囲に有限個の測定装置をつけることが現実的である。本研究において、物体の境界において有限個の電極をつけた数理モデルとして完全電極体モデルを採用する。物体 Ω 内の内部包含物の集合 D の電気伝導度 σ が連続体モデルと同じ場合、完全電極体モデルは次の偏微分方程式で記述される [4]。

$$\begin{cases} \nabla \cdot \sigma \nabla u = 0 \text{ in } \Omega, \\ u + \zeta_l \sigma \frac{\partial u}{\partial n} = V_l \text{ on } e_l, l = 1, \dots, L. \end{cases} \quad (5)$$

ここで、 e_l は電極、 ζ_l は接触抵抗、 V_l は電極 e_l における電圧とし、電圧は

$$\sum_{l=1}^L V_l = 0 \quad (6)$$

を満たすとする。

3 実験データへの適用

完全電極モデルの実験データとして 2 次元 EIT のオープンデータセットを用いる [5]。囲い込みニューラルネットワークを実験データへ適用した結果は講演時に紹介する [6]。

参考文献

- [1] A.P. Calderón, On an inverse boundary value problem, Seminar on Numerical Analysis and its Applications to Continuum Physics (1980), 65–73.
- [2] M. Ikehata, Reconstruction of the support function for inclusion from boundary measurements, Journal of Inverse and Ill-Posed Problems, 8 (2000), 367–378.
- [3] S. Siltanen, T. Ide, Electrical impedance tomography, enclosure method and machine learning, Proceeding of 2020 IEEE 30th International Workshop on Machine Learning for Signal Processing, (2020), 1–6.
- [4] M. Cheney, D. Isaacson and J. C. Newell, Electrical impedance tomography, SIAM Review, 41 (1999), 85–101.
- [5] A. Hauptmann, V. Kolehmainen, N. M. Mach, T. Savolainen, A. Seppanen and S. Siltanen, Open 2d electrical impedance tomography data archive, arXiv preprint, arXiv:1704.01178, (2017).
- [6] S. Sippola, S. Rautio, A. Hauptmann, T. Ide, S. Siltanen, Learned enclosure method for experimental EIT data, Applied Mathematics for Modern Challenges, 4, (2025), 47–65

On NeuralODE and neural network

Hirotsada Honda¹

¹ Faculty of Information and Networking for Innovation and Design, Toyo University
e-mail : honda.hirotsada@iniad.org

1 Introduction

In this talk, we will demonstrate that *every* C^1 -diffeomorphism of a compact smooth manifold M can be implemented **exactly** by the composition of a finite feed-forward neural network and a single, time-dependent Neural-ODE flow. Specifically, we will focus on the following questions:

- (Q1) *Expressive power*: Which self - maps of M can be realized *exactly* as the combination of time-1 flows of a (possibly time - dependent) vector fields and neural networks (NNs)?
- (Q2) *Learnability*: Under which geometric or analytic conditions does the associated hypothesis class admit uniform or non-uniform learnability?

Notation and conventions

■ **Manifolds.** Throughout the paper, M denotes a connected, smooth (C^∞) d -dimensional manifold.

Unless stated otherwise, we assume *compactness*. Charts are written $(U_\lambda, \phi_\lambda)$ with $\phi_\lambda : U_\lambda \rightarrow V_\lambda \subset \mathbb{R}^d$ and inverse $\psi_\lambda := \phi_\lambda^{-1}$. A smooth partition of unity $\{\rho_\lambda\}_{\lambda=1}^L$ subordinate to the finite atlas satisfies $\sum_\lambda \rho_\lambda \equiv 1$ and $\text{supp} \rho_\lambda \subseteq U_\lambda$.

■ **Neural networks.** A *one - hidden - layer network* (MLP) with activation $\sigma \in C^1(\mathbb{R})$ is the map of the form $N(\theta)(\cdot) = \sum_{i=1}^m a_i \sigma(w_i^\top x + b_i)$, $\theta = (a_i, w_i, b_i)_{i=1}^m$, that is,

$$N(\theta)(x) \in \text{Span}\langle \sigma(w^\top x + b) \mid w \in \mathbb{R}^d, b \in \mathbb{R} \rangle.$$

Its gradient is $\nabla N(\theta)(x) = \sum_i a_i \sigma'(w_i^\top x + b_i) w_i$. For manifolds, we regard $h : M \rightarrow \mathbb{R}$ as *neural* if $h = \sum_\lambda \rho_\lambda (N_\lambda \circ \phi_\lambda)$ with finitely many Euclidean MLPs N_λ .

■ **Neural - ODE flow.** Given a time - dependent vector field $X \in C^1(TM)$,

$$\dot{x}_t = X(x_t), \quad x_0 = x,$$

admits a unique flow $\Phi_X^{(t)} \in \text{Diff}^1(M)$. We call the map $x \mapsto \Phi_X^{(1)}(x)$ a *Neural - ODE block*; it is *exact* (no discretization) in all proofs.

■ **Mapping-class notation.** $\text{Diff}(M)$ is the group of orientation-preserving C^1 -diffeomorphisms on M , $\text{Diff}_0(M)$ its identity component, and the mapping - class group [1] is $\text{MCG}(M) := \text{Diff}(M)/\text{Diff}_0(M)$. Generators (e.g. Dehn twists T_γ) are regarded as fixed smooth maps.

2 Main results

The following theorem is the primary result of our paper. We fix a compact, connected, orientable d -manifold M and a bounded activation function $\sigma \in C^1(\mathbb{R})$. For a time - dependent

vector field $V(t) \in C^1(TM)$, $\Phi_V^{(t)} \in \text{Diff}^1(M)$ represents its exact flow.

Theorem 1. *Let M be a connected, orientable, compact d -manifold, and fix a smooth, non-polynomial activation $\sigma \in C^1(\mathbb{R})$. Choose any (possibly countable) generating set $\mathcal{G} \subset \text{Diff}^+(M)$ of the mapping-class group $\text{MCG}(M)$.*

For every $f \in \text{Diff}^+(M)$ and $\varepsilon > 0$ small enough, there exist

- (i) $k \in \mathbb{N}$, $\{i_1, i_2, \dots, i_k\} \subset \mathbb{N}$;
- (ii) one-hidden-layer networks $h_1, \dots, h_{g_{i_k}}$;
- (iii) C^1 , time-dependent vector fields $X_k, X_{g_{i_1}}, \dots, X_{g_{i_k}} : [0, 1] \times M \rightarrow TM$

such that

$$f = \Phi_{X_k}^{(1)} \circ (\Phi_{X_{g_{i_1}}}^{(1)} \circ h_{g_{i_1}})^{\varepsilon_1} \circ \dots \circ (\Phi_{X_{g_{i_k}}}^{(1)} \circ h_{g_{i_k}})^{\varepsilon_k} \circ R^\varepsilon,$$

where $\varepsilon_j \in \mathbb{Z}$ ($j = 1, 2, \dots, k$) and $\varepsilon \in \{0, 1\}$.

The sketch of the proof reads as follows.

Step 1: Sobolev universal approximation on manifolds. Extending Hornik's Euclidean theorem [2] with a finite atlas and a partition of unity, we demonstrate that a one-hidden-layer network with any smooth, non-polynomial activation is dense in $C^1(M)$. Hence, for any $f \in \text{Diff}^1(M)$, we obtain a finite network h whose values and Jacobians are uniformly ε -close to those of f .

Step 2: Small- C^1 isotopy. If $\varepsilon > 0$ is small enough, the error map $k = h \circ f^{-1}$ lies in the small C^1 ball around the identity and is isotopic to id_M .

Step 3: Exact correction by flow. The straight-line isotopy yields a time-dependent vector field $X = k - \text{id}$. Its time-1 flow, realized by a single Neural-ODE layer, equals k^{-1} ; thus $\Phi_X^{(1)} \circ h = f$ exactly.

Combining these blocks with Dehn – Lickorish generators gives a finite "NN + NODE" factorization for *every* element of the mapping-class group; this provides the constructive, parameter-bounded neural realization of arbitrary manifold diffeomorphisms.

Acknowledgement This work was supported by the INOUE ENRYO Memorial Grant, TOYO University.

参考文献

- [1] Farb, B and Margalit, D, A Primer on Mapping Class Groups, Princeton Univ Press, 2011.
- [2] Hornik, K. Stinchcombe, M. and White, H., Multilayer feedforward networks are universal approximators, Neural Networks, 2(1989), 359–366.

制御理論を用いた LSTM のレジリエンスの定式化と保証手法

Formulation and Guarantee of LSTM Resilience Based on Control Theory

吉原 爽太 (Yoshihara Sota)¹, 山本 涼介 (Yamamoto Ryosuke)²,
楠元 宏幸 (Kusumoto Hiroyuki)¹, 志村 将也 (Shimura Masanari)¹

¹ 名古屋大学 (Nagoya University),

² アイシン・ソフトウェア株式会社 (AISIN SOFTWARE Co., Ltd.)

e-mail : shimura.masanari.s4@s.mail.nagoya-u.ac.jp

1 はじめに

近年, AI 技術の発展に伴い, 時系列データを対象とする機械学習モデルは, 力学系における状態予測を含む, 多様な用途で活用されている. 特に最近では, Transformer や S4 などの新しいアーキテクチャが登場し, 時系列モデリングの性能は飛躍的に向上しているが, それに伴いモデルの構造は複雑化している.

一方で, AI システムの社会実装が進むにつれ, 品質保証や信頼性評価の重要性が高まっている. 従来の評価指標は主に精度や損失関数に基づいていたが, 実環境では予期せぬ外乱や入力摂動が発生する可能性があり, それに対するモデルのレジリエンスが求められる. レジリエンスとは, 入力の変化に対して出力が過度に変動しないこと, あるいは速やかに安定状態に戻る能力を指す.

本研究では, 時系列機械学習モデルに対するレジリエンス評価の枠組みを構築する第一歩として, 構造が比較的単純で解析可能な LSTM (Long Short-Term Memory) ネットワークを対象とする. 具体的には, 制御理論における δ ISS (Incremental Input-to-State Stability) の概念を応用し, 入力摂動に対する状態応答の安定性を定量的に表す「復帰時間 (Recovery Time)」を定義する. さらに, 復帰時間をデータ非依存な量で評価する方法を提案する. 本研究の目的は, AI モデルの信頼性向上に向けた新たな設計指針を提供することであり, 今後より複雑なモデルへの展開を視野に入れた基盤的な取り組みである.

2 提案手法

LSTM は状態 $s(t)$, 入力 $x(t)$, 重みパラメータ θ によって, 制御対象の出力 $y(t)$ と次の時刻の状態 $s(t+1)$ を予測する再帰型ニューラルネットワーク (RNN) の一種である. 本稿では, LSTM の詳細なゲート構造は省略して, 次のようにパラメータ θ の関数として記述する:

$$s(t+1) = f_{\theta}(s(t), x(t)), \quad y(t) = g_{\theta}(s(t)).$$

ここでは, 時系列モデルのレジリエンスを定量的に評価・制御する枠組みとして, 制御理論における δ ISS の概念を LSTM に応用する. δ ISS は, 異なる入力系列に対する状態の差が時間とともに収束する性質を保証するものである. 具体的には t に関して単調減少かつ $\|s\|$ に関して単調増加な $\beta(\|s\|, t)$ と $\|x\|$ に対して単調増加な $\gamma(\|x\|)$ を用いて次のように表されることをいう:

$$\|s_1(t) - s_2(t)\| \leq \beta(\|s_1(0) - s_2(0)\|, t) + \gamma(\|x_1(0:t) - x_2(0:t)\|_{2,\infty}).$$

LSTM に対する δ ISS の十分条件は, 近年 [1] や [2] によって提案されているが, 我々は, LSTM の内部状態空間における不変集合 (invariant set) を精緻化することで, より緩和された十分条件を導

出した [3, Theorem 4.2]. これにより, より広範なモデルに対して δ ISS 性を保証可能となった. さらに, 新たなレジリエンス指標として「復帰時間 (Recovery Time)」を定義した [3, Definition 5.1]. これは, 一時的な入力摂動により異常状態に遷移したモデルが, 通常状態に復帰するまでの時間を定量化するものである. 具体的には, ある時刻 t_0 以降で入力が一一致する入力列 $x(t)$, $\hat{x}(t)$ と閾値 e に対して次のように定める:

$$T_R(x, \hat{x}, s(0); e) = \min\{t \geq t_0 : \text{任意の } t' \geq t \text{ に対して } \|y(t', s(0), x) - y(t', s(0), \hat{x})\| \leq e\} - t_0.$$

T_R は入力データ x, \hat{x} に依存する量であり, 各テストデータに対して復帰時間を求めることは難しい. しかし, 外乱が入ってくるような異常データは収集が困難な場合が多く, 実用的な観点からはデータに依存しない評価指標が望まれる. そこで我々は, 不変集合と δ ISS の $\beta(\|s\|, t)$ を LSTM のパラメータ θ を用いて表し,

$$\sup_{x, \hat{x}} T_R(x, \hat{x}) \leq \bar{T}_R(\theta)$$

の様なデータに依存しない上界を与えた [3, Theorem 5.3]. この上界は, モデルの学習パラメータから計算可能であり, 学習済みモデルに対してレジリエンスを定量的に評価する指標として利用できる. さらに, このようなレジリエンスを評価する指標は, 次のようにモデルの学習時に正則化項 Φ として損失関数 \mathcal{L} に組み込むことで, レジリエンスを持つモデルの構築に役立たせることができる:

$$Loss = \mathcal{L} + \Phi(\theta).$$

本研究で提案したレジリエンス評価および制御手法の有効性については, 簡単なモデルを用いた数値実験により検証を行った. 詳細な実験設定および結果については, プレプリント [3] を参照されたい.

参考文献

- [1] E. Terzi, F. Bonassi, M. Farina and R. Scattolini, Learning model predictive control with long short-term memory networks, *International Journal of Robust and Nonlinear Control*, Vol. 31, No. 18 (2021), 8877–8896.
- [2] F. Bonassi, A. La Bella, G. Panzani, M. Farina and R. Scattolini, Deep Long-Short Term Memory networks: Stability properties and Experimental validation, in: *2023 European Control Conference (ECC)*, pp. 1–6, 2023.
- [3] S. Yoshihara, R. Yamamoto, H. Kusumoto and M. Shimura, Enhancing AI System Resiliency: Formulation and Guarantee for LSTM Resilience Based on Control Theory, *arXiv preprint*, 2025, arXiv: 2505.17696

Parameter identification in elliptic PDEs using coupled complex boundary method with Tikhonov regularization

Sahat Pandapotan Nainggolan¹, Julius Fergy Tiongson Rabago², Hirofumi Notsu³

¹Division of Mathematical and Physical Sciences, Kanazawa University, Kanazawa, Japan

^{2,3}Faculty of Mathematics and Physics, Kanazawa University, Kanazawa 920-1192, Japan

e-mail : ¹sahatnainggolan@stu.kanazawa-u.ac.jp/sahatpandapotannainggolan@gmail.com

1 Inverse diffusion coefficient identification

This work addresses the inverse identification of a spatially varying diffusion coefficient in a second-order elliptic PDE with convection and reaction:

Problem 1. Given f in Ω , g_2 , and g_1 on Γ , find α in Ω satisfying the equation:

$$-\nabla \cdot (\alpha \nabla u) + b \cdot \nabla u + cu = f \quad \text{in } \Omega, \quad u = g_2 \quad \text{on } \Gamma, \quad \text{and} \quad \alpha \frac{\partial u}{\partial \nu} = g_1 \quad \text{on } \Gamma, \quad (1)$$

where $f \in L^2(\Omega)$, $\alpha, c \in L^\infty(\Omega)$, $b \in [L^\infty(\Omega)]^d$, and (g_1, g_2) are given Cauchy data. This formulation generalizes traditional models by incorporating advective and reactive mechanisms, thereby allowing for a more realistic representation of physical systems. The use of a single pair of Dirichlet and Neumann data poses additional challenges but also motivates the development of advanced reconstruction techniques.

2 Recovery of the diffusion coefficient via complex boundary setting

To stabilize the inversion process, we adopt a variational approach based on a complex-valued reformulation inspired by the Coupled Complex Boundary Method (CCBM)[1, 2], where boundary data are encoded as complex values. The inverse problem is stabilized by penalizing the imaginary part of the solution, enabling robust recovery via regularization. To formulate the problem precisely, we consider the following complex boundary value problem:

Problem 2. Determine $\alpha \in \mathcal{A}$ such that $u_i = 0$ in Ω , where u_i represents the imaginary component of the solution $u = u_r + iu_i$ to the boundary value problem :

$$\begin{cases} -\nabla \cdot (\alpha \nabla u) + b \cdot \nabla u + cu = f & \text{in } \Omega, \\ \alpha \nabla u \cdot n + iu = g_2 + ig_1 & \text{on } \Gamma. \end{cases} \quad (2)$$

We proceed under the following assumptions to simplify the analysis:

- $V := H^1(\Omega)$ be the real Sobolev space and its complex counterpart is denoted by \mathbf{V} .
- We define $\mathcal{A} := \{\alpha \in L^\infty(\Omega) \mid \alpha_{\min} \leq \alpha \leq \alpha_{\max}\}$, where $\alpha_{\min}, \alpha_{\max} \in L^\infty(\Omega)$, satisfy $\alpha_{\max} > \alpha_{\min} > 0$.

We define the parameter-to-solution map $\alpha \in \mathcal{A} \mapsto S(\alpha) := u(\alpha) \in \mathbf{V}$, where $u(\alpha)$ is the solution to **Problem 2**. The mapping is continuous and Fréchet differentiable. For convenience, we write u in place of $u(\alpha)$ when no confusion arises.

We consider the following regularized cost functional

$$J_\rho(\alpha) := J(\alpha) + R(\alpha, \rho) = \frac{1}{2}w_0\|u_1(\alpha)\|_{0,\Omega}^2 + \frac{1}{2}w_1\|\nabla u_i(\alpha)\|_{0,\Omega}^2 + \frac{\rho}{2}\|\alpha\|_{0,\Omega}^2, \quad (3)$$

where $w_0, w_1 \in (0, 1)$, and ρ is the Tikhonov regularization parameter. In order to address the inverse problem, we constrain the class of admissible diffusion coefficients and formulate the following regularized optimization problem.

Problem 3. Find $\alpha_\rho \in B \subset \mathcal{A}$ such that $\alpha_\rho = \inf_{\alpha \in B} J_\rho(\alpha)$.

Theorem 4. Let B is a finite-dimensional closed convex subset of \mathcal{A} , and let $\rho > 0$ be the Tikhonov regularization parameter such that J_ρ is strictly convex. Then, Problem 3 has a unique solution $\alpha_\rho \in B$, which depends continuously on all data. Moreover, α_ρ is characterized by

$$(\nabla u_{\rho,r} \nabla w_{\rho,i} - \nabla u_{\rho,i} \nabla w_{\rho,r} + \rho \alpha_\rho, \mu - \alpha_\rho)_{0,\Omega} \geq 0, \quad \forall \mu \in B, \quad (4)$$

where $w_\rho = w_{\rho,r} + iw_{\rho,i} \in \mathbf{V}$ is the unique solution of the adjoint problem:

$$\begin{cases} -\nabla \cdot (\alpha \nabla w) + \mathbf{b} \cdot \nabla w + cw = w_0 \operatorname{Im}(u) + w_1 \Delta \operatorname{Im}(u) & \text{in } \Omega, \\ \alpha \frac{\partial w}{\partial n} - iw = 0 & \text{on } \Gamma, \end{cases} \quad (5)$$

and $u_\rho = S(\alpha_\rho) = u_{\rho,r} + iu_{\rho,i}, v \in \mathbf{V}$ solves Problem 2 in which α is replaced by α_ρ .

3 Numerical Approximation

Using the classical gradient descent method with Tikhonov regularization, the reconstruction result in Figure 1 shows the feasibility of the proposed approach with a single measurement. The method achieves stable convergence and performs well despite the presence of measurement noise, thanks to the regularization strategy.

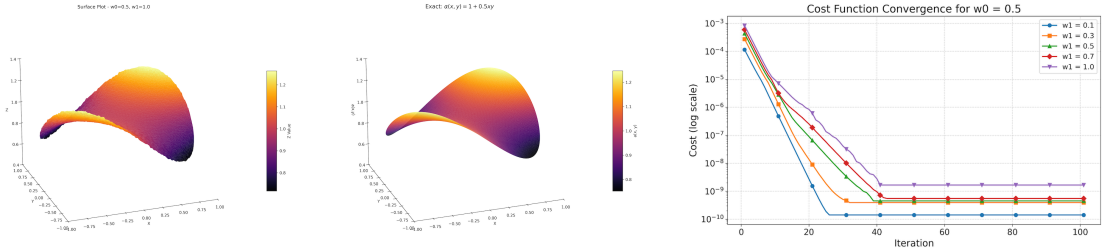


Figure 1: **Left:** Reconstruction of a smooth diffusion coefficient given by $\alpha(x, y) = 1 + 0.5xy$, with initial guess $\alpha_0 = 1.0$, $\rho = 10^{-9}$, and weights $w_0 = 0.5$, $w_1 \in \{0.1, 0.3, 0.5, 0.7, 1.0\}$. **Middle:** The exact diffusion coefficient α . **Right:** Cost history for the smooth-function case.

References

- [1] Gong, R., Xu, Y., and Cheng, X., Inverse parameter identification using coupled complex boundary method, *Inverse Problems in Science and Engineering*, **24**(9), 2016, 1594–1612.
- [2] Zheng, Y., and Zou, J., The CCBM method for inverse coefficient problems using complex Robin-type boundary conditions, *Inverse Problems*, **35**(8), 2019, 085006.