On a partial differential equation based neural network

Hirotsada Honda^{1, a)}

¹ Faculty of Information and Networking for Innovation and Design, Toyo University,
1–7–11 Akabane-dai, Kita-ku, Tokyo 115–0053, Japan

a) honda.hirotsada@iniad.org

Abstract: A continuous formulation of a neural network based on a partial differential equation (PDE) is proposed. By assuming binary classification with cross-entropy, we formulate the neural network as an optimal control problem of a PDE. The existence of the optimal solution is theoretically proved. By finding the solution and optimal control in Hilbert spaces, we provide the explicit representation of the Fréchet derivative of the cost function at the solution. We also propose a gradient-descent-based algorithm that delivers a (sub-) optimal control.

Keywords: ODENet, neural network, partial differential equation, graphon

Classification: Fundamental Theories for Communications

References

- [1] R.T.Q. Chen, Y. Rubanova, J. Bettencourt, and D.K. Duvenaud, “Neural ordinary differential equations,” *Advances in Neural Information Processing Systems* 31 (NeurIPS 2018), pp. 6571–6583, Dec. 2018.
- [2] W. E, J. Han, and Q. Li, “A mean-field optimal control formulation of deep learning,” *Res. Math. Sci.*, vol.6, 10, 2019. DOI: [10.1007/s40687-018-0172-y](https://doi.org/10.1007/s40687-018-0172-y)
- [3] H. Honda, “On continuous limit of neural network,” *Proc. NOLTA2020*, 2020.
- [4] L. Lovász, *Large Networks and Graph Limits*, American Mathematical Society, Providence, Rhode Island, 2012. DOI: [10.1090/coll/060](https://doi.org/10.1090/coll/060)
- [5] J.L. Lions, *Optimal Control of Systems Governed by Partial Differential Equations*, translated by S.K. Mitter, Springer-Verlag, Berlin, New York, 1971.
- [6] F. Abergel and R. Temam, “On some control problems in fluid mechanics,” *Theoret. Comput. Fluid Dynamics*, vol. 1, pp. 303–325, Nov. 1990. DOI: [10.1007/BF00271794](https://doi.org/10.1007/BF00271794)
- [7] S. Ito, *Partial Differential Equations*, Baihukan, Tokyo, 1970 (in Japanese).
- [8] P. Neittaanmaki, D. Tiba, and J. Sprekels, *Optimization of Elliptic Systems: Theory and Applications*, Springer-Verlag, New York, 2006. DOI: [10.1007/b138797](https://doi.org/10.1007/b138797)
- [9] E.S. Levintin and B.T. Polyak, “Constrained minimization methods,” *USSR Comput. Math. Math. Phys.*, vol. 6, no. 5, pp. 1–50, 1966. DOI: [10.1016/0041-5553\(66\)90114-5](https://doi.org/10.1016/0041-5553(66)90114-5)
- [10] D. Kinderlehrer and G. Stampacchia, *An Introduction to Variational Inequalities and Their Applications*, SIAM, 1987.
- [11] E. Zeidler, *Nonlinear Functional Analysis and Its Applications III: Variational Methods and Applications*, Springer, New York, 1985.

1 Introduction

Since it was first proposed, ODENet [1] has attracted significant interest in the research field, although it can be further improved.

A similar idea can be found in [2], in which they formulated the neural network as a form of an ODE: $\dot{x}_t = f(x_t, \theta_t)$, where x_t is the continuous limit of the states of nodes at time t , $f(\cdot)$ is a certain nonlinear function, and θ_t stands for the parameter to be controlled. Then, they defined the cost function, whose optimal control can be given by solving the Hamilton-Jacobi-Bellman equation. However, this study does not contain the topology of the network in the formulations. In actual neural networks, the weights of edges are trained on the basis of the comparison between the training data and estimated outputs. Furthermore, this study does not consider the variation among nodes in each layer. Motivated by this study, the author proposed the extension of ODENET to the partial differential equation (PDE) in a previous paper [3] by using *graphon* [4] and its extension to Sobolev spaces. Following this, we propose here a PDE-based neural network whose task is binary classification. Because of limited space, we have provided most part of the proofs of the statements in another paper.

Hereafter, we use the notation $\Omega = (0, 1)$. Let $\mathcal{G}$ be an arbitrary open set in a one- to three-dimensional Euclidean space. With an arbitrary $T > 0$, we denote $\mathcal{G}_T \equiv \mathcal{G} \times (0, T)$, and $\partial\mathcal{G}$ as the boundary of $\mathcal{G}$. The notation $C^l(\mathcal{G})$ ($l \in \mathbb{N}$) denotes a set of functions defined on $\mathcal{G}$ with l -th order continuous derivatives. $L_2(\mathcal{G})$ denotes a set of square-integrable functions defined on $\mathcal{G}$, equipped with the norm $\|f\|_{L_2(\mathcal{G})} \equiv \left(\int_{\mathcal{G}} |f(x)|^2 dx \right)^{\frac{1}{2}}$. For $r \in \mathbb{N}$, we define the Sobolev spaces $H^r(\mathcal{G})$, which are the spaces of functions $f(x)$, $x \in \mathcal{G}$ equipped with the norm $\|f\|_{H^r(\mathcal{G})}^2 \equiv \sum_{|\alpha| \leq r} \|D^\alpha f\|_{L_2(\mathcal{G})}^2$. We also introduce the anisotropic Sobolev space $H^{2,1}(\Omega_T) \equiv L_2(0, T; H^2(\Omega)) \cap L_2(\Omega; H^1(0, T))$. Let $BUC(\mathcal{G})$ be the space of functions that are uniformly bounded and continuous on a set $\mathcal{G}$. Given $T > 0$, we use the notation $\mathcal{H}_T \equiv (0, T) \times \Omega \times \Omega$. Subsequently, we introduce two functions $w_0(x)$ and $w_1(t, x, y)$, which correspond to the weight parameters in the output and hidden layers of the neural network, respectively. They belong to $L_2(\Omega)$ and $L_2(\mathcal{H}_T)$, respectively. In these spaces, we introduce the notations of the inner product as follows:

$$\langle f, g \rangle \equiv \int_{\Omega} f(x)g(x) dx, \quad \langle \langle f, g \rangle \rangle_T \equiv \int_0^T \int_{\Omega} \int_{\Omega} f(t, x, y)g(t, x, y) dt dx dy.$$

For simplicity, we also use the vector notation $\vec{w} \equiv (w_0, w_1)^\top$. The norms of vector and product spaces are defined in the usual manner. We denote the norm of the product space $\mathcal{L} \equiv L_2(\Omega) \times L_2(\mathcal{H}_T)$ merely by $\|\cdot\|$, as long as there is no confusion.

2 Methodology: PDE-based neural network

2.1 Formulation of the neural network

Following our previous paper [3], given $T > 0$, we formulate the continuous version of a neural network as follows:

$$\frac{d}{dt}u(t, x) - \nu \nabla^2 u(t, x) = \phi_1 \left(\int_{\Omega} w_1(t, x, y) u(t, y) dy \right) \text{ in } \Omega_T, \quad (1)$$

where $\nu > 0$, $\phi_1(\cdot)$ denotes the activation function, and $\nabla^2 = \frac{d^2}{dx^2}$. Additionally, we impose the initial and boundary value conditions as follows:

$$u(0, x) = u_0(x) \text{ on } \Omega, \quad u = 0 \text{ on } \partial\Omega \quad \forall t \in (0, T). \quad (2)$$

Hereafter, we consider Eqs. (1) and (2) as the continuous limit of the neural network. Furthermore, to clarify the dependence of $u(t, x)$ on its initial value u_0 and weight parameter $w_1(t, x, y)$, we frequently use the notation $u(t, x; w_1, u_0)$. Next, given a dataset $\mathcal{D} \equiv \{(X_i, Y_i)\}_{i=1}^N$, where $Y_i \in \{0, 1\}$ ($i = 1, 2, \dots, N$) are the labels, and $X_i \in \mathbb{R}^m$ with some $m \in \mathbb{N}$, supervised learning should fit the network so that it minimizes a cost function J . In this study, we limit ourselves to the case of the binary output $Y_i \in \{0, 1\}$, and the case of cross-entropy. Furthermore, we define the function of the output layer in our setting. Let us denote the weight parameter at the output layer by $w_0 = w_0(x)$, $x \in \Omega$. Accordingly, the output of the fan-out layer, given the testing data $X_i = (X_{i1}, X_{i2}, \dots, X_{im})^\top \in \mathbb{R}^m$, will be the following: $\phi_0\left(\int_{\Omega} w_0(x)u(T, x; w_1, u_{0(i)}) dx\right)$, where $u_{0(i)}(x) = \sum_{j_i=1}^m X_{ij_i} \chi_{I_{j_i}}$, with $\chi_{I_{j_i}}$ denoting the indicator functions of $I_{j_i} \equiv ((j_i - 1)/m, j_i/m]$. Accordingly, the cost function is represented as follows:

$$\begin{aligned} J[\vec{w}] = & - \sum_{i=1}^N \left[Y_i \ln \left\{ \phi_0 \left(\int_{\Omega} w_0(y)u(T, y; w_1, u_{0(i)}) dy \right) \right\} \right. \\ & \left. + (1 - Y_i) \ln \left\{ 1 - \phi_0 \left(\int_{\Omega} w_0(y)u(T, y; w_1, u_{0(i)}) dy \right) \right\} \right] \\ & + \frac{c_1}{2} \int_{\Omega} |w_0(x)|^2 dx + \frac{c_2}{2} \int_{\Omega_T} |w_1(t, x, y)|^2 dt dx dy. \end{aligned}$$

To make this cost function coercive, we added the second and third terms via Tikhonov regularization. Additionally, for later convenience, we restrict the space of $\vec{w}$ to $\mathcal{L}^{(R)} \equiv \{\vec{w} \in \mathcal{L} | \|\vec{w}\| \leq R\}$, with some $R > 0$. Next, for given $T > 0$, our problem is written as follows:

$$\begin{aligned} \vec{w}_o = & \operatorname{argmin}_{\vec{w} \in \mathcal{L}^{(R)}} J[\vec{w}], \\ \text{s.t. } \vec{w} \text{ satisfies } & (1)-(2) \text{ with } u_0 \text{ replaced by } u_{0(i)}' \text{'s.} \end{aligned} \quad (P)$$

2.2 Solvability of the equation

Here, we discuss the solvability of problem (1)–(2).

Theorem 2.1. *Let T be an arbitrary positive number, let us assume the following.*

- (i) $u_0 \in L_2(\Omega)$,
- (ii) $\phi_0 \in C^1(\mathbb{R})$, and $0 < \phi_0(x) < 1 \quad \forall x \in \mathbb{R}$,
- (iii) $\phi_1 \in C^2(\mathbb{R}) \cap BUC(\mathbb{R})$, $\phi_1'(x) > 0$ for all $x \in \mathbb{R}$,
- (iv) $|\phi_1'(x)| \vee |\phi_1''(x)| < M$ for all $x \in \mathbb{R}$ for some $M > 0$.

Then, given that $w_1(t, x, y) \in L_2(\mathcal{H}_T)$, problem (1)–(2) has a unique solution $u \in H^{2,1}(\Omega_T)$ satisfying $\|u\|_{H^{2,1}(\Omega_T)} \leq c'(\|u_0\|_{L_2(\Omega)}, T)$ with some positive constant $c'(\|u_0\|_{L_2(\Omega)}, T)$ that depends on $\|u_0\|_{L_2(\Omega)}$ and T .

Next, we state the existence of an optimal control $\vec{w}$ to Problem (P) [5].

Theorem 2.2. *Under the assumptions in Theorem 2.1, there exists at least one element $\vec{w}_o \equiv (\bar{w}_0, \bar{w}_1)^\top \in \mathcal{L}^{(R)}$ such that the following holds:*

$$u(t, x; \bar{w}_1, u_{0(i)}) \in H^{2,1}(\Omega_T) \quad \forall i = 1, 2, \dots, N,$$

$$J[\vec{w}_o] \leq J[\vec{w}] \quad \forall \vec{w} \in \mathcal{L}^{(R)}.$$

Proof. Because $J[\vec{w}]$ is bounded below, we can take a minimizing sequence $\vec{w}^{(n)} \equiv (w_0^{(n)}, w_1^{(n)})^\top$ for Problem (P), and let us define $u_i^{(n)} \equiv u(t, x; w_1^{(n)}, u_{0(i)})$. The sequence $\{(w_0^{(n)}, w_1^{(n)})\}_{n=1}^\infty$ is bounded in $L_2(\Omega) \times L_2(\mathcal{H}_T)$. Thereby, for each $i = 1, 2, \dots, N$, $\{u_i^{(n)}\}_{n=1}^\infty$ is bounded in $H^{2,1}(\Omega_T)$. We can then find $\vec{w}_o = (\bar{w}_0, \bar{w}_1)^\top$, $\bar{u}_i$, the subsequences $\vec{w}^{(n')} \equiv (w_0^{(n')}, w_1^{(n')})^\top$ of $(w_0^{(n)}, w_1^{(n)})$, and $\{u_i^{(n')}\}$ of $\{u_i^{(n)}\}$, satisfying

$$(w_0^{(n')}, w_1^{(n')}) \rightarrow (\bar{w}_0, \bar{w}_1) \text{ weakly in } L_2(\Omega) \times L_2(\mathcal{H}_T),$$

$$u_i^{(n')} \rightarrow \bar{u}_i \text{ weakly in } L_2(0, T; H^2(\Omega)),$$

as n tends to infinity. Because $J[\vec{w}]$ is weakly lower semi-continuous, we have [5, 6] $J[\vec{w}_o] \leq \lim_{n' \rightarrow \infty} J[\vec{w}^{(n')}]$. Additionally, we can show that $\vec{w}_o$ and the corresponding solutions $u(t, x; \bar{w}_1, u_{0(i)}) = \bar{u}_i$ ($i = 1, 2, \dots, N$) construct an admissible solution to Problem (P). $\square$

2.3 Auxiliary lemma

Next, we prepare an auxiliary lemma.

Lemma 2.1. *In addition to the assumptions in Theorem 2.1, let us assume that $\vec{w} \in \mathcal{L}^{(R)}$ is given, and we introduce $\Delta_i u(t, x) \equiv u(t, x; w_1 + \Delta w_1, u_{0(i)}) - u(t, x; w_1, u_{0(i)})$. Additionally, let us assume $c_0 MR < \varepsilon'$, where ε' is a certain small number, and $c_0 > 0$ denotes the minimum positive constant that guarantees the estimate $\|\mathcal{Z}f\|_{H^{2,1}(\Omega_T)} \leq c_0 \|f\|_{L_2(\Omega_T)}$, where $\mathcal{Z}f(t, x) \equiv \int_0^t d\tau \int_\Omega Z(t - \tau, x, y) f(\tau, y) dy$. Here, $Z(t, x, y) = \sum_{n=-\infty}^\infty K(t, x - y + 2n) + K(t, x + y + 2n)$, where $K(t, x) = \frac{1}{\sqrt{4\pi t}} e^{-\frac{x^2}{4t}}$. Accordingly, $\Delta_i u(t, x) \in H^{2,1}(\Omega_T)$ and is represented as*

$$\Delta_i u(t, x) = \int_0^t d\tau \int_{\Omega \times \Omega} \check{\Gamma}_{w_1}^{(i)}(t, \tau, x, y, z) \Delta w_1(\tau, y, z) dy dz + o(\|\vec{w}\|),$$

where $\check{\Gamma}_{w_1}^{(i)}(t, \cdot, x, \cdot, \cdot) \in L_2(\mathcal{H}_T)$ for each $t \in (0, T)$, $x \in \Omega$, and $w_1 \in L_2^{(R)}(\mathcal{H}_T)$.

Proof. We briefly write the essence of the proof here. Notably, $\Delta_i u$ solves

$$\begin{cases} \frac{\partial(\Delta_i u)}{\partial t} - \nu \nabla^2(\Delta_i u) = \phi \left(\int_\Omega (w + \Delta w)(t, x, y) u(t, y; w + \Delta w, u_{0(i)}) dy \right) \\ \quad - \phi \left(\int_\Omega w(t, x, y) u(t, y; w, u_{0(i)}) dy \right) \text{ in } \Omega_T, \\ \Delta_i u|_{\partial\Omega} = 0 \quad t \in (0, T), \\ \Delta_i u(0, x) = 0 \text{ on } \Omega. \end{cases} \quad (3)$$

As for the form of the fundamental solution $Z(t, x, y)$, one can refer to [7]. After applying the mean value theorem, we can obtain the explicit representation of the solution by using the Neumann series. $\square$

2.4 Necessary condition

We state the necessary condition for obtaining the optimal control to Problem (P) [6].

Theorem 2.3. *Under the assumptions of Lemma 2.1, we have:*

$$J[\vec{w} + \Delta \vec{w}] - J[\vec{w}] = c_1 \langle w_0, \Delta w_0 \rangle + c_2 \langle w_1, \Delta w_1 \rangle + \sum_{i=1}^N \phi'_0 \left(\int_{\Omega} w_0(x) u(T, x; w_1, u_{0(i)}) dx \right) f_{1(i)}(T; \vec{w}) \times \left\{ \langle u(T, \cdot; w_1, u_{0(i)}), \Delta w_0 \rangle + \langle \tilde{\Gamma}_{w_1}^{(i)}, \Delta w_1 \rangle \right\} + o(\|\Delta \vec{w}\|),$$

where $f_{1(i)}(t; \vec{w}) \equiv \frac{Y_i}{\phi_0 \left(\int_{\Omega} w_0(x) u(t, x; w_1, u_{0(i)}) dx \right)} - \frac{1-Y_i}{1-\phi_0 \left(\int_{\Omega} w_0(x) u(t, x; w_1, u_{0(i)}) dx \right)}$, and $\tilde{\Gamma}_{w_1}^{(i)} \in L_2(\mathcal{H}_T)$.

From Theorem 2.3, we have the representation of the Fréchet gradient of J at $\vec{w}$.

$$J'_{\vec{w}} = \left(\sum_{i=1}^N \phi'_0 \left(\int_{\Omega} w_0(x') u(T, x'; w_1, u_{0(i)}) dx' \right) f_{1(i)}(T; \vec{w}) u(T, x; w_1, u_{0(i)}) + c_1 w_0(x), \right. \\ \left. \sum_{i=1}^N \phi'_0 \left(\int_{\Omega} w_0(x') u(T, x'; w_1, u_{0(i)}) dx' \right) f_{1(i)}(T; \vec{w}) \tilde{\Gamma}_{w_1}^{(i)}(t, x, y) + c_2 w_1(t, x, y) \right)^{\top}.$$

We also assert

Lemma 2.2. *Under the assumptions of Lemma 2.1, the Fréchet gradient $J'_{\vec{w}}$ is Lipschitz continuous with respect to $\vec{w}$:*

$$\|J'_{\vec{w}} - J'_{\vec{w}'}\| \leq L_0 \|\vec{w} - \vec{w}'\|,$$

where $L_0 > 0$ is the Lipschitz constant.

2.5 Numerical algorithm

Let us propose an algorithm that provides a sub-optimal solution. Mathematically, the algorithm can be obtained by regarding the weight parameter functions as elements in Hilbert spaces, and then identifying the spaces with their dual spaces [8]. Notably, the magnitude of $\|\vec{w}\|$ is restricted to be less than R . Therefore, we shall employ gradient projection method. We update $\vec{w}$ according to the following rule:

$$\vec{w}^{(j+1)} = \mathbf{P}_R [\vec{w}^{(j)} - \alpha_j J'_{\vec{w}^{(j)}}] \quad (j = 0, 1, 2, \dots), \quad (4)$$

where $\{\alpha_j\}_j$ denotes a sequence of small positive numbers, satisfying $\varepsilon_1 < \alpha_j < \frac{2}{L_0 + 2\varepsilon_2}$ ($j = 0, 1, 2, \dots$) with some positive constants ε_i ($i = 1, 2$) [9], and $\mathbf{P}_R$ denotes a map defined on $\mathcal{L}$ that replaces its arguments with their projection on a ball of radius R in $\mathcal{L}$:

$$\mathbf{P}_R[\vec{w}] \equiv \begin{cases} \vec{w} & \text{if } \|\vec{w}\| \leq R, \\ \frac{R\vec{w}}{\|\vec{w}\|} & \text{otherwise.} \end{cases}$$

Algorithm 1 Compute the optimal control

Require: Dataset $\mathcal{D} = \{(X_i, Y_i)\}_{i=1}^N, \{\alpha_j\}$, the initial weight $\vec{w}^{(0)} = (w_0^{(0)}, w_1^{(0)})^\top$ and iteration time $S \in \mathbb{N}$

Ensure: The (sub-) optimal control $\vec{w}_o = (\bar{w}_0, \bar{w}_1)^\top$

while $j < S$ **do**

$j \leftarrow j + 1$

$$w_0^{(j)} \leftarrow -\alpha_j \left(c_1 w_0^{(j-1)} + \sum_{i=1}^N f_{1(i)}(T; \vec{w}^{(j-1)}) \phi'_0 \left(\int_{\Omega} w_0(x') u(T, x'; w_1, u_{0(i)}) dx' \right) \times u(T, \cdot; w_1^{(j-1)}, u_{0(i)}) \right),$$

$$w_1^{(j)} \leftarrow -\alpha_j \left(c_2 w_1^{(j-1)} + \sum_{i=1}^N f_{1(i)}(T; \vec{w}^{(j-1)}) \phi'_0 \left(\int_{\Omega} w_0(x') u(T, x'; w_1, u_{0(i)}) dx' \right) \times \tilde{\Gamma}_{w_1^{(j-1)}}^{(i)} \right),$$

if $\|\vec{w}^{(j-1)} - \alpha_{j-1} J'_{\vec{w}^{(j-1)}}\| > R$ **then**

$$\vec{w}^{(j)} \leftarrow \frac{R(\vec{w}^{(j-1)} - \alpha_{j-1} J'_{\vec{w}^{(j-1)}})}{\|\vec{w}^{(j-1)} - \alpha_{j-1} J'_{\vec{w}^{(j-1)}}\|}$$

end if

end while

2.6 Convergence theorems

Theorem 2.4. Under the assumptions of Lemma 2.1, $\vec{w}^{(j)}$ in (4) satisfies $\|\vec{w}^{(j+1)} - \vec{w}^{(j)}\| \rightarrow 0$ as $j \rightarrow \infty$ and $J[\vec{w}^{(j)}]$ monotonically decreases.

Proof. It is known that every closed subspace of a uniformly convex Banach space is uniformly convex. According to Levinton and Polyak [9], Theorem 5.1, we have the statement above. $\square$

Next, we formulate a variational inequality that corresponds to (P) [10]:

$$\langle J'_{\vec{w}}, \vec{w}' - \vec{w} \rangle_T \geq 0 \quad \forall \vec{w}' \in \mathcal{L}^{(R)}. \quad (5)$$

Following [11], we have

Theorem 2.5. In addition to the assumptions in Theorem 2.4, we assume that c_i ($i = 1, 2$) are sufficiently large, and $\alpha_j = \alpha_0$ ($j = 1, 2, \dots$) with a sufficiently small positive number $\alpha > 0$. Then, the iteration of (4) gives the unique solution $\vec{w}^*$ to inequality (5). In addition, if $\vec{w}^{(0)} - \alpha_0 J'_{\vec{w}^{(0)}} \in \mathcal{L}^{(R)}$ holds, then the following claims hold with some $K_0 \in [0, 1)$.

$$(i) \quad \|\vec{w}^{(j)} - \vec{w}^*\| \rightarrow 0 \text{ as } j \rightarrow +\infty,$$

$$(ii) \quad \|\vec{w}^{(j)} - \vec{w}^*\| \leq \frac{\alpha_0 K_0^j}{(K_0 - 1)} \|J'_{\vec{w}^{(0)}}\| \text{ for } j = 1, 2, \dots$$

3 Conclusion

In this paper, we extended the concept of ODENet to the PDE version. We theoretically proved the existence of the optimal solution to the proposed framework. We also proposed a numerical scheme to obtain the (sub-) optimal solution based on gradient descent.

However, the drawbacks of our method are some restrictions introduced to theoretically guarantee the existence of an optimal control and the explicit representation of the Fréchet derivative of the cost function. Our future study will be performed with the relaxation of these restrictions, and it will also concern the numerical implementation of the proposed method.

Acknowledgments

We thank the anonymous reviewer whose comments and suggestions helped improve and clarify this manuscript. This work was supported by the INOUE ENRYO Memorial Grant, TOYO University.