Temperature Estimation using Physics-Informed Neural Networks

A Study on Thermotherapy for Skin Tumors

Alessio Borgi, Eugenio Bugli, Alessandro De Luca

Academic Year 2023/2024

SAPIENZA
UNIVERSITÀ DI ROMA

Table of Contents

Thermotherapy treats diseases by altering skin temperature. Tumors can be weakened or even eradicated by these treatments:

Thermotherapy

- Thermalablation (50-100°C)
- Cryoablation (-20°C to -40°C)
- Hyperthermia (39-45°C)

Applications of **Hyperthermia**:

- Cancer treatment: Enhances the effectiveness of chemotherapy and radiotherapy.
- Tumor targeting: Focuses heat on tumors to weaken or destroy cancer cells.
- Non-invasive methods: Useful for deep tissue and subcutaneous tumors.

Setup

1] Introduction

- **Device Setup**: An antenna, controlled by a robotic arm, is placed on the skin.
- **Safety Measures**: A water bolus is used to prevent skin burns.
- **Non-Invasive Control**:
  - High accuracy in temperature measurement at depth
  - Practical ease of use, minimal bulk, and cost-effectiveness

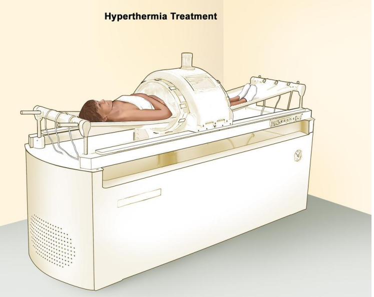


Figure: Hyperthermia device setup

A Partial Differential Equation (PDE) is a mathematical equation that involves multiple independent variables and their partial derivatives.

- Allows real-time temperature monitoring without invasive intervention
- PDEs are used to model various physical phenomena such as heat transfer, wave propagation, and fluid dynamics.
- They describe how a system evolves over time or space in relation to its variables.

Physics-Informed Neural Networks (PINNs)

1] Introduction

The Bio-Heat Equation is complex, but we use Physics-Informed Neural Networks (PINNs) to:

- Solve partial differential equations (PDEs) efficiently by incorporating physical laws directly into the learning process.
- Handle high-dimensional data and complex boundary conditions, which are common in medical applications like hyperthermia treatment.
- Simplify and accelerate 2D temperature control in biological tissues.
- Provide faster calculations compared to traditional numerical methods, making them suitable for real-time applications.

Traditional methods like Finite-Difference Time-Domain focus on accuracy and stability, but are computationally expensive and only suited for pre-treatment analysis. NN-Based Observer Approach:

- Uses deep neural networks (DNNs) to predict temperature in real-time by simulating the system's behavior.
- The DNN model with a perfusion coefficient ω_b, reducing prediction error based on how close the model's ω_b is to the actual system value.

Table of Contents

2] BioHeat Equation

Equation 1

2] BioHeat Equation

Pennes' principal theoretical contribution was the suggestion that the rate of heat transfer between blood and tissue is proportional to the product of the volumetric perfusion rate and the difference between the arterial blood temperature and the local tissue temperature:

$$h_b = \omega_b \rho_b c_b (1 - \kappa)(T_a - T) \tag{1}$$

- h_b: Rate of heat transfer per unit volume of tissue. [W/m^3]
- ω_b: Perfusion rate per volume of tissue. [1/s]
- ρ_b: Density of blood. [kg/m^3]
- c_b: Specific heat of blood. [J/(kg·K)]
- κ: Factor for incomplete thermal equilibrium. [dimensionless]
- T_a: Temperature of arterial blood. [K]
- T: Temperature of local tissue. [K]

Equation 2

2] BioHeat Equation

Bio-Heat Equation:

$$\rho c \frac{\partial T}{\partial t} = k_{\text{eff}} \frac{\partial^2 T}{\partial x^2} - \rho_b c_b \omega_b (T - T_a) + \mathcal{Q} \qquad x \in \Omega,\ t \in [0, t_f] \tag{2}$$

- ρ: Density of the tissue. [kg/m^3]
- c: Specific heat capacity of the tissue. [J/(kg·K)]
- $\frac{\partial T}{\partial t}$: Rate of temperature change in the tissue. [K/s]
- k_{eff}: Effective thermal conductivity of the tissue. [W/(m·K)]
- $\frac{\partial^2 T}{\partial x^2}$: Thermal diffusion within the tissue. [K/m^2]
- $\mathcal{Q}$: Metabolic heat generation and external power in hyperthermia. [W/m^3]

Values

2] BioHeat Equation

Parameter	Value	Parameter	Value
ρ	1050	T_a	37
ρ_b	1043	T_{max}	45
c	3639	Q	0
c_b	3825	w_b	2.22×10^{-3}
k_{eff}	5	w_{min}	0.43×10^{-4}
q_0	16	w_{max}	3.8×10^{-3}
β	15	L_0	0.05
t_f	1800	ΔT	8

Table: Values of parameters used in the model.

Scaling and Normalization

- Numerical instability can occur when solving the Pennes Bio-Heat equation with traditional methods (FDM, FEM, FVM).
- Scaling and normalization are used to mitigate instability:

$$T' = T - T_a, \quad \theta = \frac{T'}{T_M - T_a}, \quad X = \frac{x}{L_o}, \quad \tau = \frac{t}{\tau_f} \tag{3}$$

Equations

2] BioHeat Equation

After some computations:

- **1D Case:**

$$\partial_\tau \theta = a_1 \partial_{XX}\theta - a_2 W\theta + a_3 \mathcal{Q} \tag{4}$$

- **2D Case:**

$$\partial_\tau \theta = a_1 (\partial_{XX}\theta + \partial_{YY}\theta) - a_2 W\theta + a_3 \mathcal{Q} \tag{5}$$

With the coefficient as:

$$a_1 = \frac{(\rho \cdot c \cdot L_0^2)}{(tf \cdot k_{\text{eff}})} \qquad a_2 = \frac{(\rho_b \cdot c_b \cdot w_b \cdot L_0^2)}{k_{\text{eff}}} \qquad a_3 = \frac{L_0^2}{(T_M - T_a)k_{\text{eff}}}$$

Assumption/Considerations

2] BioHeat Equation

There some consideration before approaching the solution system. There could be some discrepancies between its theoretical prediction and the experimental data due to:

- **Anisotropy in Tissue Perfusion**: The heat conduction is not uniform in all direction, the heat conduction and blood perfusion can vary depending on the direction.
- **Heat transfer Blood and Tissue**: The heat transfer is not always proportional to the blood perfusion rate and the temperature difference between the arterial blood and tissue.
- Model the heat generation term $\mathcal{Q}$ as a constant

Table of Contents

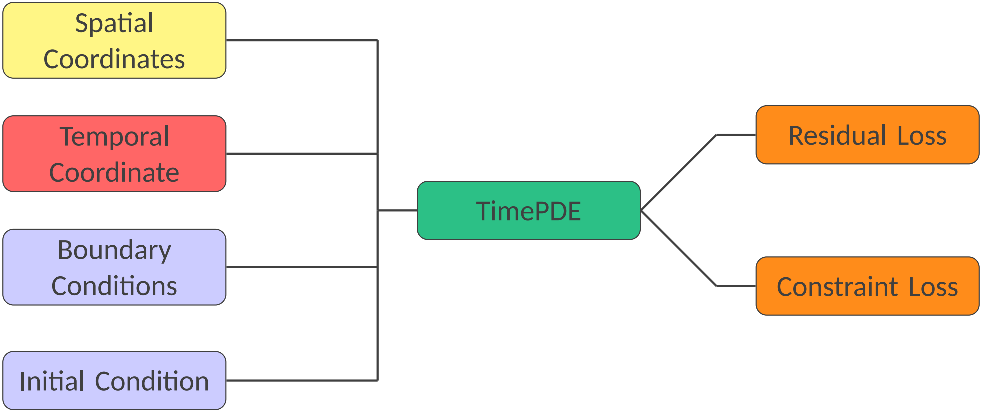
3] Implementation Procedure

Implementation Procedure

- Data generation: Mathematica have been used to generate the solution of the PDE equation related to our problem. This data have been used to test our network performances.
- Training: Our network has been trained by randomly sampling a specific number of points inside the domain. However, we could also have set a specific point locations for training.
- Optimisation procedure: The loss function will be minimized by using gradient-based optimizers like Adam and L-BFGS.

Structure

3] Implementation Procedure



Spatial and Temporal Coordinates

- Spatial Coordinates: The surface of the equation has been implemented in both One and Two dimensional case. For the **One Dimensional** case we have only an X coordinate, while for the **Two Dimensional** case we have both X and Y.
- Temporal Coordinate: Our PDE belongs to the group of **Temporal Dependent PDEs**. This means that we will have a temporal coordinate along with spatial ones, which specifies the behaviour at each instant of time.

Boundary and Initial Conditions

3] Implementation Procedure

- Boundary Conditions: They describe the behaviour of the model along the boundaries of our surface. We have four of them for the Two dimensional case and Two for the One dimensional case.
- Initial Condition: Describes the behaviour of the model at the initial instant of time. Since we are using an Observer to solve the equation, we will use a different initial condition for it.

Neural Bio-Heat Observer (NHBO)

- We would like to have an Observer able to reconstruct the behaviour of the Mathematica Model:

$$\partial_\tau \hat{\theta} = a_1 \nabla^2 \hat{\theta} - a_2 \hat{\theta} + a_3 \tag{6}$$

- Our Observer uses **different Initial conditions** with respect to the one used on Mathematica. By doing this, we are ensuring that the NHBO is learning the general physical principles, rather than a specific secenario. This also leads to better generalization, robustness, and applicability in real-world scenarios where conditions might vary widely.

Loss Function

3] Implementation Procedure

- In order to measure the discrepancy between our network and the constraints imposed by the boundary and initial conditions, we have used the following Loss Function:

$$\mathcal{L}(\theta, \mathcal{T}) = \omega_f \mathcal{L}_f(\theta, \mathcal{T}_f) + \omega_b \mathcal{L}_b(\theta, \mathcal{T}_b) \tag{7}$$

  This function is defined as the **weighted summation** of the $\boldsymbol{L^2}$ **Norm** of two specificic losses.

- Residual Loss $\boldsymbol{\mathcal{L}_f}$: This component specifies how well the neural network satisfies the governing equation.
- Constraint Loss $\boldsymbol{\mathcal{L}_b}$: This component specifies how well the neural network satisfies the boundary and initial conditions of the problem.

E

Table of Contents

4] Tools & Experimental Settings

Tools

The **Tools** we have been using are:

- **Neural Bio-Heat Observer (NHBO)** implementation using **DeepXDE** library.
  - **Embedding** the **PDE** into the **neural network's loss function** using automatic differentiation.
  - **Time-Dependent PDE** nature $\rightarrow$ **TimePDE** Class
- **Weight & Biases** together with **MatPlotLib** for comparative plotting.
- **Parameters Setting** using **Hydra**.
- **Hyper-parameter Optimization** using **Optuna**.

Hyper-Parameters Optimization

4] Tools & Experimental Settings

We followed a very **Defined Process/Pipeline** in order to select the **Best-Model Settings**:

Activation Function	Initialization
SoftPlus	Glorot Normal
TanH	Glorot Normal
Sin	Glorot Normal
Sigmoid	Glorot Normal
SELU	Glorot Normal
APTx	Glorot Normal
APTx	He Normal
Mish	He Normal
Swish	He Normal
SiLU	He Normal
ReLU	He Normal
GELU	He Normal
ELU	He Normal

Parameters Optimized	Values
number_of_layers	2, 3, 4
number_neurons_per_layer	50, 75, 100
optimizer	Adam, SGD, L-BFGS
max_iters	3, 5, 10, 15, 20, 25, 50, 75, 100, 125, 150, 175, 200, 250, 300, 400, 500, 750, 1000

Parameters Fixed	Values
backend	PyTorch
learning_rate	0.001
output_injection_gain	50
resampling	true
resampler_period	50
initial_weights_regularizer	true

The key performance **Metrics** used are:

- **Mean Absolute Error (MAE)**, **Mean Squared Error (MSE)**, **L2 Relative Error (L2RE)**, **Maximum Absolute Percentage Error (Max-APE)**

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \tag{8}$$

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{9}$$

$$\text{L2RE} = \frac{\sqrt{\sum_{i=1}^{n}(y_i - \hat{y}i)^2}}{\sqrt{\sum i = 1^n y_i^2}} \tag{10}$$

$$\text{Max-APE} = \max\left(\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\right) \ \forall i \in \{1, \ldots, n\} \tag{11}$$

Composite Score

A **Composite Score** has been developed to simplify the interpretation of these metrics.

$$\text{Composite\_Score} = w_{MAE} \cdot norm_{MAE} + w_{MSE} \cdot norm_{MSE} + w_{L2RE} \cdot norm_{L2RE} + w_{MAX\_APE} \cdot norm_{MAX\_APE}$$

where **Min-Max Normalization** is applied:

$$\text{normalized\_metric} = \frac{\text{metric} - \min(\text{metric})}{\max(\text{metric}) + \min(\text{metric})} \quad (12)$$

and the weights are:

$$w_{MAE} = w_{MSE} = w_{L2RE} = w_{MAX\_APE} = 0.25 \quad (13) \qquad \sum_{i \in \{MAE, MSE, L2RE, MAX\_APE\}} w_i = 1 \quad (14)$$

RunTime & Computational Resources

Other two important **Metrics** are:

- **RunTime**: The Model Training time.
- **Hardware Consumes**: We also analyse the resources stress level and in particular:
  - **CPU Utilization (%)**: Percentage of CPU stress level during the Training Process.
  - **Memory Available (MB)**: Quantification of Memory usage. In particular:
    - **Non-Swap**: Data that remains stored in the RAM.
    - **Swap**: Data moved to Disk when RAM is full.
  - **Disk Utilization (GB)**: Quantification of Disk Storage.

Table of Contents

5] (1D) One-dimension

Starting from the previous simplifications, we have used the following equation for the One dimensional case:

$$a_1 \partial_\tau \theta = \partial_{XX}\theta - a_2\theta + a_3 \tag{15}$$

Where the three coefficients are:

$$a_1 = \frac{(\rho \cdot c \cdot L_0^2)}{(tf \cdot k_{\text{eff}})} \qquad a_2 = \frac{(\rho_b \cdot c_b \cdot w_b \cdot L_0^2)}{k_{\text{eff}}} \qquad a_3 = \frac{\mathcal{Q} \cdot L_0^2}{(T_M - T_a)k_{\text{eff}}}$$

For this case of study we have set $\mathcal{Q}$ directly to zero, hence the third coefficient will always be zero.

Boundary Conditions

5] (1D) One-dimension

We have only two boundary conditions, which are specified when $X = 0$ and when $X = 1$, since we have only one spatial dimension. These conditions describe the behaviour of the model along the boundaries of our one dimensional surface.

- The **Left boundary condition** is specified when $X = 0$:

$$\theta(0, \tau) = 0 \qquad \text{for } X = 0, \tau \geq 0 \tag{16}$$

- The **Right boundary condition** is specified when $X = 1$:

$$\partial_x \theta(1, \tau) = K(\tau + \frac{1 - \tau}{1 + e^{-s(\tau - t_c)}} - \theta(1, \tau)) \qquad \text{for } X = 1, \tau \geq 0 \tag{17}$$

Initial Condition

5] (1D) One-dimension

As previously stated, we will have two different initial conditions for the equation model and our observer.

- **Mathematica Model**: we have a simple condition, which sets the initial temperature to zero:

$$\theta(X,0) = 0 \qquad \forall X \in [0,1], \tau = 0 \tag{18}$$

- **NHBO**: we use a condition that depends on the spatial coordinate and on the temperature of the arterial blood:

$$\theta(X,0) = q_0 \frac{X^4}{4(T_{max} - T_a)} \qquad \forall X \in [0,1], \tau = 0 \tag{19}$$

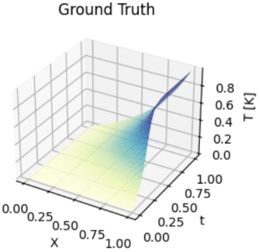
E

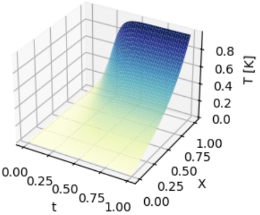
Surface

5] (1D) One-dimension

- Since we have only **One spatial dimension**, the Geometry object from DeepXDE has been defined by using the **Interval** class.
- While the **Temporal dimension** has been implemented using the **TimeDomain** class.

The joint domain of our problem is composed of these two objects combined into a **GeometryXTime** object.





Results [1]

5] (1D) One-dimension

The **1D (Without Q) Results** Optimization Journey starts by varying the **Num_Dense_Layers** hyperparameter. With **2 Dense Layers**:

- **APTx** is the **fastest** Activation function.
- **GELU** is the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
APTx	500	**1072**	0.0324	He normal
APTx	1000	1193	0.2017	Glorot normal
Softplus	1000	1363	0.3666	Glorot normal
Mish	1000	1645	0.0274	He normal
tanh	1000	1982	0.4867	Glorot normal
Swish	1000	2117	0.0280	He normal
sin	1000	2303	0.4794	Glorot normal
SiLU	1000	2440	0.0280	He normal
Sigmoid	1000	2632	0.4294	Glorot normal
SELU	1000	2747	0.3524	Glorot normal
ReLu	50	3081	0.1091	He normal
GELU	500	3137	**0.0264**	He normal
ELU	1000	3255	0.0418	He normal

TABLE II: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
APTx	1000	**1005**	0.0708	He normal
APTx	1000	1234	**0.0619**	Glorot normal
Softplus	1000	1445	0.2858	Glorot normal
Mish	1000	1820	0.0918	He normal
tanh	1000	2224	0.1152	Glorot normal
Swish	1000	2350	0.0848	He normal
sin	1000	2563	0.1025	Glorot normal
SiLU	1000	2698	0.0848	He normal
Sigmoid	25	3038	0.4324	Glorot normal
SELU	1000	3040	0.1323	Glorot normal
ReLu	25	3416	0.0918	He normal
GELU	750	3448	0.0746	He normal
ELU	1000	3575	0.1390	He normal

TABLE III: 3 Dense Layers

AF	Iter.	Runtime	Score	Initialization
APTx	250	**806**	0.0755	He normal
APTx	1000	895	0.4448	Glorot normal
Softplus	1000	1145	0.1643	Glorot normal
Mish	500	1830	0.0946	He normal
tanh	1000	2165	0.1311	Glorot normal
Swish	750	2419	**0.0688**	He normal
sin	1000	2647	0.1135	Glorot normal
SiLU	750	2912	**0.0688**	He normal
Sigmoid	1000	3153	0.1628	Glorot normal
SELU	750	3439	0.3369	Glorot normal
ReLu	1000	3732	0.2362	He normal
GELU	150	4088	0.0690	He normal
ELU	500	4274	0.1389	He normal

TABLE IV: 4 Dense Layers

Results [1]

5] (1D) One-dimension

The **1D (Without Q) Results** Optimization continue by analysing with **3 Dense Layers**:

- **APTx** is the **fastest** Activation function.
- **APTx** is the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
APTx	500	**1072**	0.0324	He normal
APTx	1000	1193	0.2017	Glorot normal
Softplus	1000	1363	0.3666	Glorot normal
Mish	1000	1645	0.0274	He normal
tanh	1000	1982	0.4867	Glorot normal
Swish	1000	2117	0.0280	He normal
sin	1000	2303	0.4794	Glorot normal
SiLU	1000	2440	0.0280	He normal
Sigmoid	1000	2632	0.4294	Glorot normal
SELU	1000	2747	0.3524	Glorot normal
ReLu	50	3081	0.1091	He normal
GELU	500	3137	**0.0264**	He normal
ELU	1000	3255	0.0418	He normal

TABLE II: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
APTx	1000	**1005**	0.0708	He normal
APTx	1000	1234	**0.0619**	Glorot normal
Softplus	1000	1445	0.2858	Glorot normal
Mish	1000	1820	0.0918	He normal
tanh	1000	2224	0.1152	Glorot normal
Swish	1000	2350	0.0848	He normal
sin	1000	2563	0.1025	Glorot normal
SiLU	1000	2698	0.0848	He normal
Sigmoid	25	3038	0.4324	Glorot normal
SELU	1000	3040	0.1323	Glorot normal
ReLu	25	3416	0.0918	He normal
GELU	750	3448	0.0746	He normal
ELU	1000	3575	0.1390	He normal

TABLE III: 3 Dense Layers

AF	Iter.	Runtime	Score	Initialization
APTx	250	**806**	0.0755	He normal
APTx	1000	895	0.4448	Glorot normal
Softplus	1000	1145	0.1643	Glorot normal
Mish	500	1830	0.0946	He normal
tanh	1000	2165	0.1311	Glorot normal
Swish	750	2419	**0.0688**	He normal
sin	1000	2647	0.1135	Glorot normal
SiLU	750	2912	**0.0688**	He normal
Sigmoid	1000	3153	0.1628	Glorot normal
SELU	750	3439	0.3369	Glorot normal
ReLu	1000	3732	0.2362	He normal
GELU	150	4088	0.0690	He normal
ELU	500	4274	0.1389	He normal

TABLE IV: 4 Dense Layers

Results [1]

5] (1D) One-dimension

The **1D (Without Q) Results** Optimization continue by analysing with **4 Dense Layers**:

- **APTx** is the **fastest** Activation function.
- **Swish** and **SiLU** are the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
APTx	500	**1072**	0.0324	He normal
APTx	1000	1193	0.2017	Glorot normal
Softplus	1000	1363	0.3666	Glorot normal
Mish	1000	1645	0.0274	He normal
tanh	1000	1982	0.4867	Glorot normal
Swish	1000	2117	0.0280	He normal
sin	1000	2303	0.4794	Glorot normal
SiLU	1000	2440	0.0280	He normal
Sigmoid	1000	2632	0.4294	Glorot normal
SELU	1000	2747	0.3524	Glorot normal
ReLu	50	3081	0.1091	He normal
GELU	500	3137	**0.0264**	He normal
ELU	1000	3255	0.0418	He normal

TABLE II: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
APTx	1000	**1005**	0.0708	He normal
APTx	1000	1234	**0.0619**	Glorot normal
Softplus	1000	1445	0.2858	Glorot normal
Mish	1000	1820	0.0918	He normal
tanh	1000	2224	0.1152	Glorot normal
Swish	1000	2350	0.0848	He normal
sin	1000	2563	0.1025	Glorot normal
SiLU	1000	2698	0.0848	He normal
Sigmoid	25	3038	0.4324	Glorot normal
SELU	1000	3040	0.1323	Glorot normal
ReLu	25	3416	0.0918	He normal
GELU	750	3448	0.0746	He normal
ELU	1000	3575	0.1390	He normal

TABLE III: 3 Dense Layers

AF	Iter.	Runtime	Score	Initialization
APTx	250	**806**	0.0755	He normal
APTx	1000	895	0.4448	Glorot normal
Softplus	1000	1145	0.1643	Glorot normal
Mish	500	1830	0.0946	He normal
tanh	1000	2165	0.1311	Glorot normal
Swish	750	2419	**0.0688**	He normal
sin	1000	2647	0.1135	Glorot normal
SiLU	750	2912	**0.0688**	He normal
Sigmoid	1000	3153	0.1628	Glorot normal
SELU	750	3439	0.3369	Glorot normal
ReLu	1000	3732	0.2362	He normal
GELU	150	4088	0.0690	He normal
ELU	500	4274	0.1389	He normal

TABLE IV: 4 Dense Layers

Results [2]

5] (1D) One-dimension

We continue the **1D (Without Q) Results** Optimization Journey by varying the **Num_Dense_Nodes** hyperparameter. With **50 Dense Nodes**:

- **APTx** is the **fastest** Activation function.
- **GELU** is the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
APTx	500	**1072**	0.0324	He normal
APTx	1000	1193	0.2017	Glorot normal
Softplus	1000	1363	0.3666	Glorot normal
Mish	1000	1645	0.0274	He normal
tanh	1000	1982	0.4867	Glorot normal
Swish	1000	2117	0.0280	He normal
sin	1000	2303	0.4794	Glorot normal
SiLU	1000	2440	0.0280	He normal
Sigmoid	1000	2632	0.4294	Glorot normal
SELU	1000	2747	0.3524	Glorot normal
ReLu	50	3081	0.1091	He normal
GELU	500	3137	**0.0264**	He normal
ELU	1000	3255	0.0418	He normal

TABLE II: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
APTx	500	**14**	0.9999	He normal
GELU	500	38	**0.0000**	He normal
Mish	1000	19	0.5000	He normal
Swish	1000	110	0.5000	He normal
SiLU	1000	159	0.5000	He normal

TABLE V: Dense Nodes 75

AF	Iter.	Runtime	Score	Initialization
Mish	1000	17	0.9999	He normal
Swish	1000	147	**0.0000**	He normal
SiLU	1000	210	**0.0000**	He normal
APTx	500	**10**	0.9998	He normal
GELU	500	41	**0.0000**	He normal

TABLE VI: Dense Nodes 100

Results [2]

5] (1D) One-dimension

The **1D (Without Q) Results** Optimization Journey continue by analysing **75 Dense Nodes**:

- **APTx** is the **fastest** Activation function.
- **GELU** is the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
APTx	500	**1072**	0.0324	He normal
APTx	1000	1193	0.2017	Glorot normal
Softplus	1000	1363	0.3666	Glorot normal
Mish	1000	1645	0.0274	He normal
tanh	1000	1982	0.4867	Glorot normal
Swish	1000	2117	0.0280	He normal
sin	1000	2303	0.4794	Glorot normal
SiLU	1000	2440	0.0280	He normal
Sigmoid	1000	2632	0.4294	Glorot normal
SELU	1000	2747	0.3524	Glorot normal
ReLu	50	3081	0.1091	He normal
GELU	500	3137	**0.0264**	He normal
ELU	1000	3255	0.0418	He normal

TABLE II: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
APTx	500	**14**	0.9999	He normal
GELU	500	38	**0.0000**	He normal
Mish	1000	19	0.5000	He normal
Swish	1000	110	0.5000	He normal
SiLU	1000	159	0.5000	He normal

TABLE V: Dense Nodes 75

AF	Iter.	Runtime	Score	Initialization
Mish	1000	17	0.9999	He normal
Swish	1000	147	**0.0000**	He normal
SiLU	1000	210	**0.0000**	He normal
APTx	500	**10**	0.9998	He normal
GELU	500	41	**0.0000**	He normal

TABLE VI: Dense Nodes 100

Results [2]

5] (1D) One-dimension

The **1D (Without Q) Results** Optimization Journey continue by analysing **100 Dense Nodes**:

- **APTx** is the **fastest** Activation function.
- **Swish**, **SiLU** and **GELU** are the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
APTx	500	**1072**	0.0324	He normal
APTx	1000	1193	0.2017	Glorot normal
Softplus	1000	1363	0.3666	Glorot normal
Mish	1000	1645	0.0274	He normal
tanh	1000	1982	0.4867	Glorot normal
Swish	1000	2117	0.0280	He normal
sin	1000	2303	0.4794	Glorot normal
SiLU	1000	2440	0.0280	He normal
Sigmoid	1000	2632	0.4294	Glorot normal
SELU	1000	2747	0.3524	Glorot normal
ReLu	50	3081	0.1091	He normal
GELU	500	3137	**0.0264**	He normal
ELU	1000	3255	0.0418	He normal

TABLE II: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
APTx	500	**14**	0.9999	He normal
GELU	500	38	**0.0000**	He normal
Mish	1000	19	0.5000	He normal
Swish	1000	110	0.5000	He normal
SiLU	1000	159	0.5000	He normal

TABLE V: Dense Nodes 75

AF	Iter.	Runtime	Score	Initialization
Mish	1000	17	0.9999	He normal
Swish	1000	147	**0.0000**	He normal
SiLU	1000	210	**0.0000**	He normal
APTx	500	**10**	0.9998	He normal
GELU	500	41	**0.0000**	He normal

TABLE VI: Dense Nodes 100

Results [3]

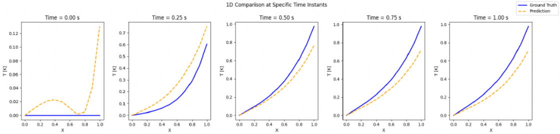
5] (1D) One-dimension

Since we have a **tie** on the **Best Composite Score Model**, we can decide the best model between the **SiLU** and **GELU** depending on what we want to weight more:

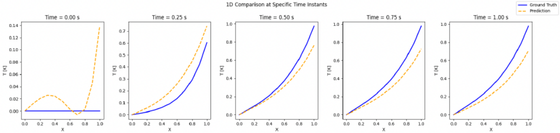
- **GELU** & **Swish** have better **RunTime**.
- **SiLU** has slightly better **Scores**.

Network Activation	L2RE	MSE
SiLU	0.304959534	0.01476774416
GELU	0.3080251058	0.01506613866

TABLE VII: Comparison of L2RE and MSE for SiLU and GELU



(a) Gelu



(b) SiLU

Fig. 1: Comparison of GELU, SiLU in 2D

Results [3]

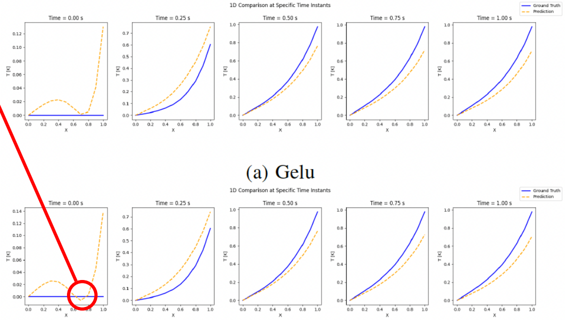
5] (1D) One-dimension

By looking at the **GELU** and **SiLU performance 2D graphs**, we note that:

- They are **very similar**.
- For **SiLU**, the value falls **below zero**, i.e., not a viable option. However, no subsequent result is affected.

Network Activation	L2RE	MSE
SiLU	0.304959534	0.01476774416
GELU	0.3080251058	0.01506613866

TABLE VII: Comparison of L2RE and MSE for SiLU and GELU



(a) Gelu

(b) SiLU

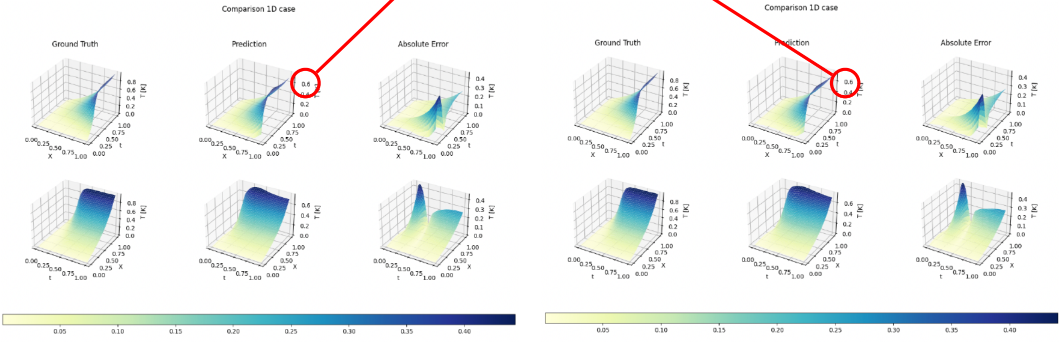
Fig. 1: Comparison of GELU, SiLU in 2D

Results [4]

5] (1D) One-dimension

By looking at the **GELU** and **SiLU performance 3D graphs**, we note that:

They only reach 0.6!



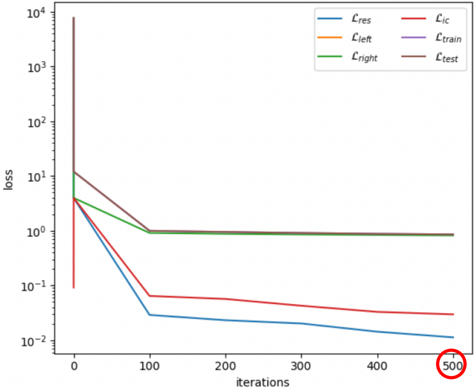
(a) Gelu

(b) SiLU

Results [5]

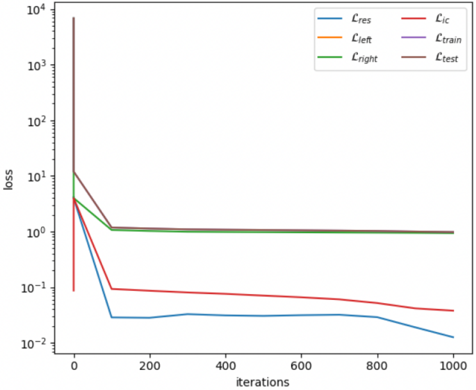
5] (1D) One-dimension

By looking at the **GELU** and **SiLU Losses graphs**, we note that:



(a) Gelu

It converges faster!



(b) SiLU

Results [6]

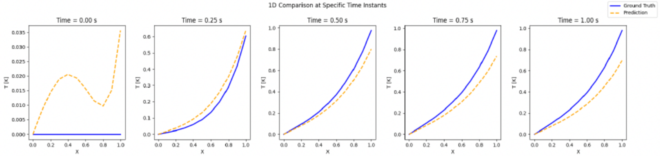
5] (1D) One-dimension

By employing also the **L-BFGS** optimizer, we obtain:

- **ReLU** has better **RunTime**.
- **Sin** has better **Scores**.

AF	Iter.	Runtime	Score	Initialization
APTx	25	59	0.1024	He normal
APTx	25	51	0.0379	Glorot normal
Softplus	25	22	0.5489	Glorot normal
Mish	25	147	0.0841	He normal
tanh	25	37	0.0205	Glorot normal
Swish	25	60	0.0750	He normal
sin	25	42	**0.0199**	Glorot normal
SiLU	25	54	0.0750	He normal
Sigmoid	25	27	0.2810	Glorot normal
SELU	25	9	0.5151	Glorot normal
ReLu	25	**4**	0.1762	He normal
GELU	25	52	0.0960	He normal
ELU	25	11	0.0369	He normal

TABLE VIII: LBFGS with 100 Dense Nodes



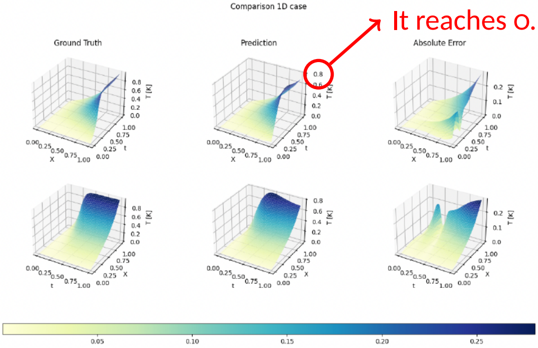
(a) Comparison 2D between prediction and ground truth for the sin model

Results [7]

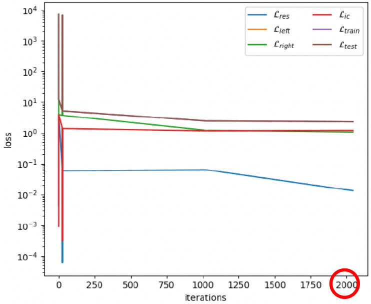
5] (1D) One-dimension

The **Sin** model, being less accurate w.r.t. the previous result, it **reaches ground truth values**, using **L-BFGS**, in more iterations:

It reaches 0.8!



(b) Comparison 3D between prediction and ground truth for the sin model



(c) Loss of the sin model

Results [8]

5] (1D) One-dimension

Focusing on the **HW Complexities**, by analysing the **CPU** and **Disk Utilization** results, we observe:

- **APTx** and **Mish** require **more CPU Utilization(%)**.
- **APTx** and **SoftPlus** require more **Disk Utilization (GB)**.
- **APTx** was the fastest model. $\rightarrow$ **Time-HW Trade-Off**

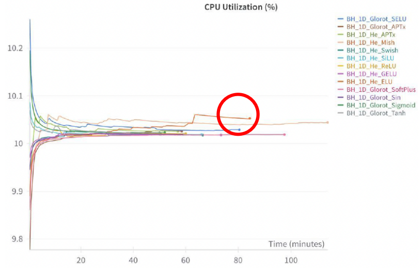


Fig. 12: CPU Utilization (%)

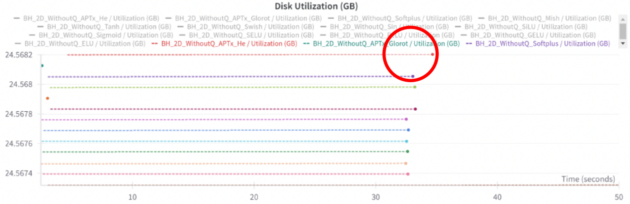


Fig. 16: Disk Utilization (GB)

Results [9]

5] (1D) One-dimension

By analysing instead the **Memory Available (Swap)** and **(No-Swap)** results, we obtain:

- **SiLU** consumes less **Memory Swap(MB)**.
- **APTx** consumes less **Memory No-Swap(MB)**

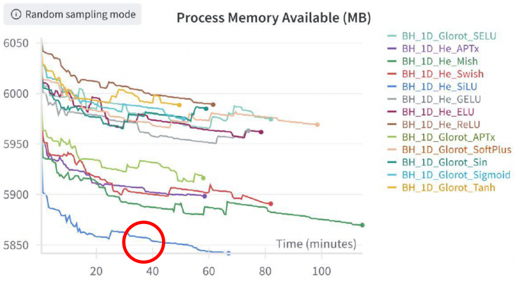


Fig. 14: Memory with Swap (%)

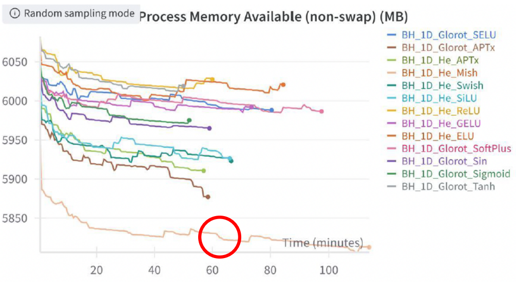


Fig. 13: Memory with no Swap (%)

Table of Contents

6] (2D WithoutQ) Two-dimensions

In the **Two Dimensional case** we have used this version of the equation:

$$a_1 \partial_\tau \theta = \nabla^2 \theta - a_2 \theta + a_3 \tag{20}$$

Which can be formulated as follows:

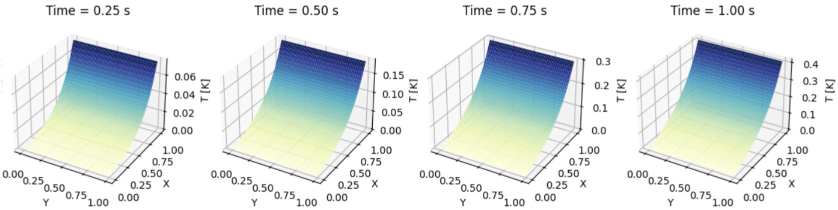
$$a_1 \partial_\tau \theta = \partial_{XX} \theta + \partial_{YY} \theta - a_2 \theta + a_3 \tag{21}$$

In our Analysis we have focused on both the cases when we have $\mathcal{Q} = 0$ and when $\mathcal{Q} \neq 0$. Both problems have been formulated with the same initial and boundary conditions.

Surface

6] (2D WithoutQ) Two-dimensions

We have **Two Spatial Coordinates** and One Temporal Coordinate in our joint domain. Our Surface is now modeled as a Square. However, we have excluded all vertices due to **Reduce Overfitting** and **Avoid Convergence Issues**.



Boundary and Initial Conditions

6] (2D WithoutQ) Two-dimensions

We have used the same initial condition for both the Observer and the equation Model. In this case we have a total of four boundary conditions, since now they are also define for the upper and lower boundary. The **Upper** and **Lower** boundary conditions are defined with respect the Y coordinate.

$$\partial_Y \theta(X, 0, \tau) = 0 \qquad \forall X \in [0, 1] \wedge \tau \geq 0 \tag{22}$$

$$\partial_Y \theta(X, 1, \tau) = 0 \qquad \forall X \in [0, 1] \wedge \tau \geq 0 \tag{23}$$

While the **Left** and **Right** boundary conditions are defined with respect the X coordinate.

$$\theta(0, Y, \tau) = 0 \qquad \forall Y \in [0, 1] \wedge \tau \geq 0 \tag{24}$$

$$\partial_X \theta(1, Y, \tau) = \tau \qquad \forall Y \in [0, 1] \wedge \tau \geq 0 \tag{25}$$

What do we expect from our Network

6] (2D WithoutQ) Two-dimensions

- We have used the same Initial Condition with additional Boundary Conditions with respect to the One dimensional case.
- We are dealing with a more complex problem that has another dimension, hence we will have a total of 3 dimensions (two spatial and 1 temporal).
- As a consequence, we have to evaluate the performances of our network at **Specific Instants of time**.

Results [1]

6] (2D WithoutQ) Two-dimensions

The **2D (Without Q) Results** Optimization Journey starts by varying the **Num_Dense_Layers** hyperparameter. With **2 Dense Layers**:

- **APTx** is the **fastest** Activation function.
- **ELU** is the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
ELU	500	7632	**0.0000**	He normal
GELU	500	5487	0.2470	He normal
ReLu	400	5084	0.3804	He normal
SELU	500	4586	0.1726	Glorot normal
Sigmoid	125	4441	0.2492	Glorot normal
SiLU	150	3973	0.1061	He normal
sin	50	3670	0.2329	Glorot normal
Swish	150	3110	0.1061	He normal
tanh	400	2477	0.0714	Glorot normal
Mish	150	2264	0.1506	He normal
Softplus	125	1859	0.7726	Glorot normal
APTx	300	1271	0.2762	Glorot normal
APTx	300	**834**	0.276	He normal

TABLE X: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
ELU	500	20731	**0.0193**	He normal
GELU	500	20271	0.1421	He normal
ReLu	500	19836	0.4709	He normal
SELU	500	19374	0.3973	Glorot normal
Sigmoid	75	19305	0.0814	Glorot normal
SiLU	500	18511	0.0681	He normal
sin	50	18473	0.3534	Glorot normal
Swish	500	17620	0.0681	He normal
tanh	50	17585	0.2139	Glorot normal
Mish	500	16703	0.1073	He normal
Softplus	500	16231	0.8569	Glorot normal
APTx	500	15803	0.1858	Glorot normal
APTx	500	**700**	0.1858	He normal

TABLE XI: 3 Dense Layers

AF	Iter.	Runtime	Score	Initialization
ELU	1000	2310	0.0236	He normal
GELU	1000	2129	0.0352	He normal
ReLU	1000	1964	0.6648	He normal
SELU	1000	1789	0.5267	Glorot normal
Sigmoid	1000	1635	0.7915	Glorot normal
SiLU	1000	1464	**0.0225**	He normal
sin	250	1423	0.1659	Glorot normal
Swish	1000	1119	**0.0225**	He normal
tanh	250	1072	0.0390	Glorot normal
Mish	1000	746	0.0375	He normal
Softplus	750	610	0.0508	Glorot normal
APTx	1000	405	0.0917	Glorot normal
APTx	1000	**249**	0.0917	He normal

TABLE XII: 4 Dense Layers

Results [1]

6] (2D WithoutQ) Two-dimensions

The **2D (Without Q) Results** Optimization continue by analysing with **3 Dense Layers**:

- **APTx** is the **fastest** Activation function.
- **ELU** is the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
ELU	500	7632	**0.0000**	He normal
GELU	500	5487	0.2470	He normal
ReLu	400	5084	0.3804	He normal
SELU	500	4586	0.1726	Glorot normal
Sigmoid	125	4441	0.2492	Glorot normal
SiLU	150	3973	0.1061	He normal
sin	50	3670	0.2329	Glorot normal
Swish	150	3110	0.1061	He normal
tanh	400	2477	0.0714	Glorot normal
Mish	150	2264	0.1506	He normal
Softplus	125	1859	0.7726	Glorot normal
APTx	300	1271	0.2762	Glorot normal
APTx	300	**834**	0.276	He normal

TABLE X: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
ELU	500	20731	**0.0193**	He normal
GELU	500	20271	0.1421	He normal
ReLu	500	19836	0.4709	He normal
SELU	500	19374	0.3973	Glorot normal
Sigmoid	75	19305	0.0814	Glorot normal
SiLU	500	18511	0.0681	He normal
sin	50	18473	0.3534	Glorot normal
Swish	500	17620	0.0681	He normal
tanh	50	17585	0.2139	Glorot normal
Mish	500	16703	0.1073	He normal
Softplus	500	16231	0.8569	Glorot normal
APTx	500	15803	0.1858	Glorot normal
APTx	500	**700**	0.1858	He normal

TABLE XI: 3 Dense Layers

AF	Iter.	Runtime	Score	Initialization
ELU	1000	2310	0.0236	He normal
GELU	1000	2129	0.0352	He normal
ReLU	1000	1964	0.6648	He normal
SELU	1000	1789	0.5267	Glorot normal
Sigmoid	1000	1635	0.7915	Glorot normal
SiLU	1000	1464	**0.0225**	He normal
sin	250	1423	0.1659	Glorot normal
Swish	1000	1119	**0.0225**	He normal
tanh	250	1072	0.0390	Glorot normal
Mish	1000	746	0.0375	He normal
Softplus	750	610	0.0508	Glorot normal
APTx	1000	405	0.0917	Glorot normal
APTx	1000	**249**	0.0917	He normal

TABLE XII: 4 Dense Layers

Results [1]

6] (2D WithoutQ) Two-dimensions

The **2D (Without Q) Results** Optimization continue by analysing with **4 Dense Layers**:

- **APTx** is the **fastest** Activation function.
- **Swish** and **SiLU** are the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
ELU	500	7632	**0.0000**	He normal
GELU	500	5487	0.2470	He normal
ReLu	400	5084	0.3804	He normal
SELU	500	4586	0.1726	Glorot normal
Sigmoid	125	4441	0.2492	Glorot normal
SiLU	150	3973	0.1061	He normal
sin	50	3670	0.2329	Glorot normal
Swish	150	3110	0.1061	He normal
tanh	400	2477	0.0714	Glorot normal
Mish	150	2264	0.1506	He normal
Softplus	125	1859	0.7726	Glorot normal
APTx	300	1271	0.2762	Glorot normal
APTx	300	**834**	0.276	He normal

TABLE X: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
ELU	500	20731	**0.0193**	He normal
GELU	500	20271	0.1421	He normal
ReLu	500	19836	0.4709	He normal
SELU	500	19374	0.3973	Glorot normal
Sigmoid	75	19305	0.0814	Glorot normal
SiLU	500	18511	0.0681	He normal
sin	50	18473	0.3534	Glorot normal
Swish	500	17620	0.0681	He normal
tanh	50	17585	0.2139	Glorot normal
Mish	500	16703	0.1073	He normal
Softplus	500	16231	0.8569	Glorot normal
APTx	500	15803	0.1858	Glorot normal
APTx	500	**700**	0.1858	He normal

TABLE XI: 3 Dense Layers

AF	Iter.	Runtime	Score	Initialization
ELU	1000	2310	0.0236	He normal
GELU	1000	2129	0.0352	He normal
ReLU	1000	1964	0.6648	He normal
SELU	1000	1789	0.5267	Glorot normal
Sigmoid	1000	1635	0.7915	Glorot normal
SiLU	1000	1464	**0.0225**	He normal
sin	250	1423	0.1659	Glorot normal
Swish	1000	1119	**0.0225**	He normal
tanh	250	1072	0.0390	Glorot normal
Mish	1000	746	0.0375	He normal
Softplus	750	610	0.0508	Glorot normal
APTx	1000	405	0.0917	Glorot normal
APTx	1000	**249**	0.0917	He normal

TABLE XII: 4 Dense Layers

Results [2]

6] (2D WithoutQ) Two-dimensions

We continue the **2D (Without Q) Results** Optimization Journey by varying the **Num_Dense_Nodes** hyperparameter. With **50 Dense Nodes**:

- **APTx** is the **fastest** Activation function.
- **ELU** is the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
ELU	500	7632	**0.0000**	He normal
GELU	500	5487	0.2470	He normal
ReLu	400	5084	0.3804	He normal
SELU	500	4586	0.1726	Glorot normal
Sigmoid	125	4441	0.2492	Glorot normal
SiLU	150	3973	0.1061	He normal
sin	50	3670	0.2329	Glorot normal
Swish	150	3110	0.1061	He normal
tanh	400	2477	0.0714	Glorot normal
Mish	150	2264	0.1506	He normal
Softplus	125	1859	0.7726	Glorot normal
APTx	300	1271	0.2762	Glorot normal
APTx	300	**834**	0.276	He normal

TABLE X: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
ELU	250	13913	**0.0042**	He normal
tanh	1000	571	0.2514	Glorot normal
SiLU	750	453	0.9069	He normal
Swish	750	306	0.9069	He normal
Mish	750	**136**	0.7563	He normal

TABLE XIII: Dense Nodes 75

AF	Iter.	Runtime	Score	Initialization
ELU	1000	814	**0.04576**	He normal
tanh	1000	653	0.2769	Glorot normal
SiLU	1000	472	0.1583	He normal
Swish	1000	299	0.1583	He normal
Mish	1000	**100**	0.2721	He normal

TABLE XIV: Dense Nodes 100

Results [2]

6] (2D WithoutQ) Two-dimensions

The **2D (Without Q) Results** Optimization Journey continue by analysing **75 Dense Nodes**:

- **Mish** is the **fastest** Activation function.
- **ELU** is the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
ELU	500	7632	**0.0000**	He normal
GELU	500	5487	0.2470	He normal
ReLu	400	5084	0.3804	He normal
SELU	500	4586	0.1726	Glorot normal
Sigmoid	125	4441	0.2492	Glorot normal
SiLU	150	3973	0.1061	He normal
sin	50	3670	0.2329	Glorot normal
Swish	150	3110	0.1061	He normal
tanh	400	2477	0.0714	Glorot normal
Mish	150	2264	0.1506	He normal
Softplus	125	1859	0.7726	Glorot normal
APTx	300	1271	0.2762	Glorot normal
APTx	300	**834**	0.276	He normal

TABLE X: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
ELU	250	13913	**0.0042**	He normal
tanh	1000	571	0.2514	Glorot normal
SiLU	750	453	0.9069	He normal
Swish	750	306	0.9069	He normal
Mish	750	**136**	0.7563	He normal

TABLE XIII: Dense Nodes 75

AF	Iter.	Runtime	Score	Initialization
ELU	1000	814	**0.04576**	He normal
tanh	1000	653	0.2769	Glorot normal
SiLU	1000	472	0.1583	He normal
Swish	1000	299	0.1583	He normal
Mish	1000	**100**	0.2721	He normal

TABLE XIV: Dense Nodes 100

Results [2]

6] (2D WithoutQ) Two-dimensions

The **2D (Without Q) Results** Optimization Journey continue by analysing **100 Dense Nodes**:

- **Mish** is the **fastest** Activation function.
- **ELU** is the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
ELU	500	7632	**0.0000**	He normal
GELU	500	5487	0.2470	He normal
ReLu	400	5084	0.3804	He normal
SELU	500	4586	0.1726	Glorot normal
Sigmoid	125	4441	0.2492	Glorot normal
SiLU	150	3973	0.1061	He normal
sin	50	3670	0.2329	Glorot normal
Swish	150	3110	0.1061	He normal
tanh	400	2477	0.0714	Glorot normal
Mish	150	2264	0.1506	He normal
Softplus	125	1859	0.7726	Glorot normal
APTx	300	1271	0.2762	Glorot normal
APTx	300	**834**	0.276	He normal

TABLE X: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
ELU	250	13913	**0.0042**	He normal
tanh	1000	571	0.2514	Glorot normal
SiLU	750	453	0.9069	He normal
Swish	750	306	0.9069	He normal
Mish	750	**136**	0.7563	He normal

TABLE XIII: Dense Nodes 75

AF	Iter.	Runtime	Score	Initialization
ELU	1000	814	**0.04576**	He normal
tanh	1000	653	0.2769	Glorot normal
SiLU	1000	472	0.1583	He normal
Swish	1000	299	0.1583	He normal
Mish	1000	**100**	0.2721	He normal

TABLE XIV: Dense Nodes 100

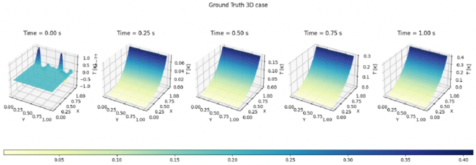
A

Results [3]

6] (2D WithoutQ) Two-dimensions

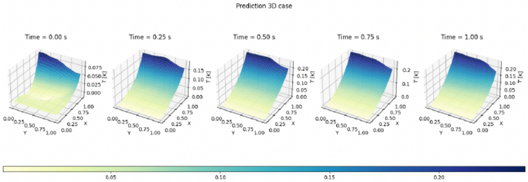
By looking at the **SiLU** and **ELU performance 3D graphs**, we note that:

- **APTx** has better **RunTime.**
- **ELU** has better **Scores.**

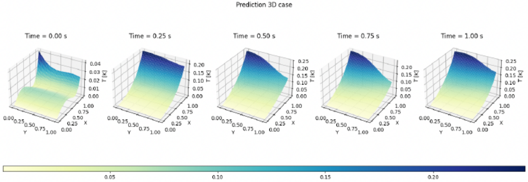


(a) Ground Truth

Prediction 3D case



(b) ELU

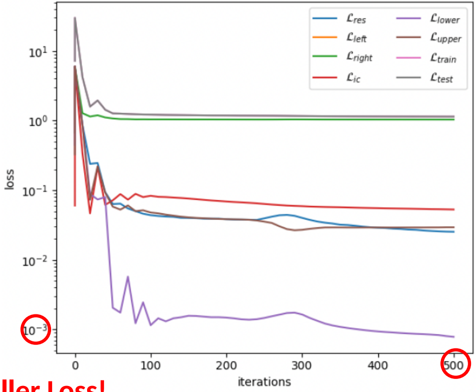


(c) SiLU

Results [4]

6] (2D WithoutQ) Two-dimensions

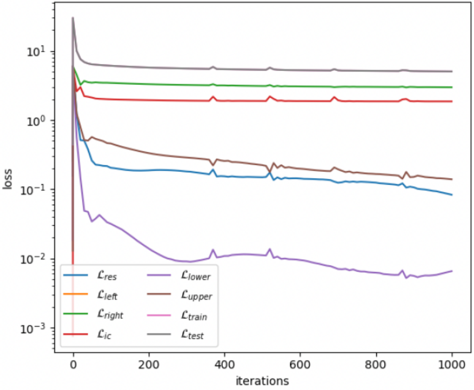
By looking at the **ELU** and **SiLU Losses graphs**, we note that:



Smaller Loss!

Faster Convergence!

(a) ELU

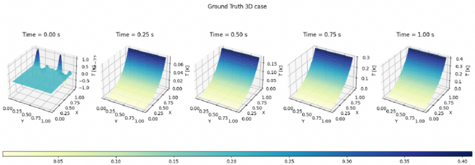


(b) SiLU

Results [5]

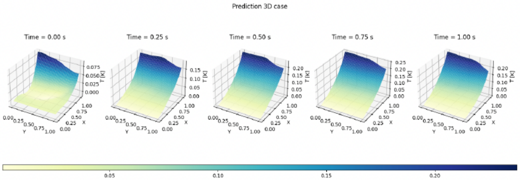
6] (2D WithoutQ) Two-dimensions

By employing also the **L-BFGS** optimizer, we obtain even more precise predictions for both **ELU** and **SiLU**:

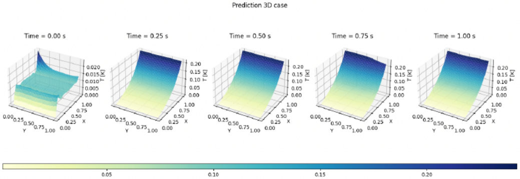


(a) Ground Truth

Prediction 3D case



(a) ELU

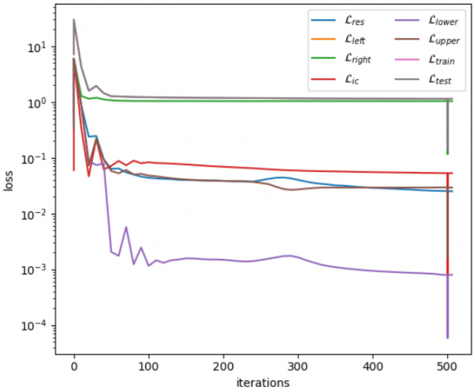


(b) SiLU

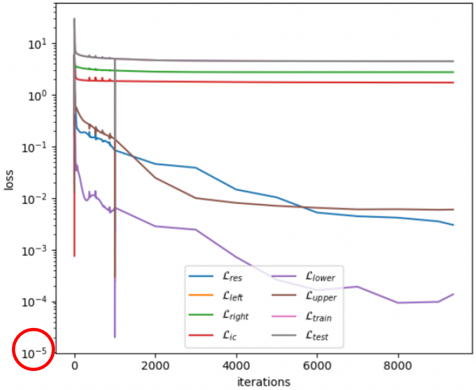
Results [6]

6] (2D WithoutQ) Two-dimensions

We can also make the SiLU VS GELU Comparison using **LBFGS** using their **Losses**:



(a) ELU



Smaller Loss!

(b) SiLU

Results [7]

6] (2D WithoutQ) Two-dimensions

Focusing on the **HW Complexities**, by analysing the **CPU**, **Disk** and **Memory Utilization** results, we observe:

- **ELU** and **GELU** require **more CPU Utilization(%)**.
- **APTx**, **SoftPlus** require **more Disk Utilization (GB)** and **less Memory Utilization (MB)**

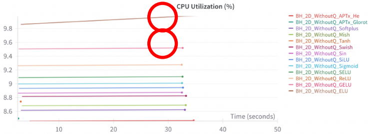


Fig. 15: CPU Utilization (%)

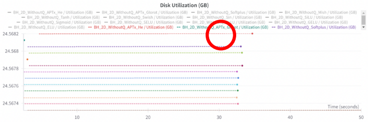


Fig. 16: Disk Utilization (GB)

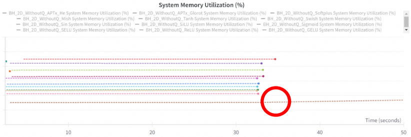


Fig. 17: System Memory Utilization (%)

Table of Contents

7] (2D WithQ) Two-dimensions

Assumptions Made on $\mathcal{Q}$ (Tissue)

As previously stated, Hyperthermia treatments have been successfully applied to a set of different tumors. Each of them, is characterized by different properties which have key relevance to our problem.

- Head and Neck tumors
- Breast tumors
- Melanoma
- Soft tissue Sarcoma
- Chest wall recurrences
- Cutaneous Lymphoma

E

Assumptions Made on $\mathcal{Q}$ (Tissue)

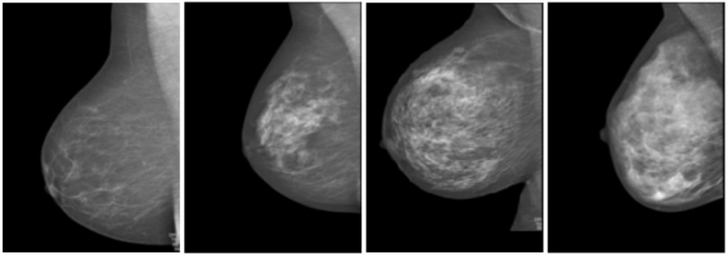
In our work, we have chosen among all, to work with the **Breast tissue**, since it's the most common type of cancer in the world. Its main characteristics are:

- It is an **heterogeneous tissue** encased in a layer of skin
- It is composed of **Fat** and Fibroglandular tissue
- It has a **variable density**, which is why it is difficult to generate standard models
- There exist different types of cancer, each with its own characteristics
- Healthy tissue has different properties with respect to the cancerous one

Assumptions Made on $\mathcal{Q}$ (Tissue)

This type of tissue is charachterized by a **variable density**. This property changes accordingly to different factors like:

- Age (denser when younger)
- Quantity of **Fibroglandular** tissue (the more you have it, the denser the tissue will be)
- Hormone Therapies (denser)
- Genetics



(a) (b) (c) (d)

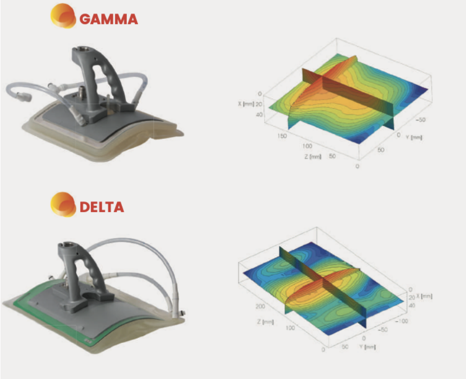
Assumptions Made on $\mathcal{Q}$ (Machine)

We have assumed to use the **ALBA ON 4000D** Platform as our reference Hyperthermia system.

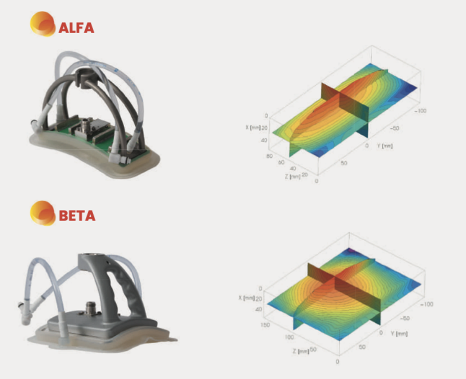
- This Machine works at a fixed frequency of 434 MHz (**Microwave Level**)
- It has an integrated Water Bolus (not taken into account in our work)
- Effectively heats targets, in the range of **41-43°C** for 60 minutes, as required by the ESHO guidelines
- It has **four different sized applicators**, each with an effective heating depth of 4 ± 0.5cm and different radiative areas

Antenna Applicators

7] (2D WithQ) Two-dimensions



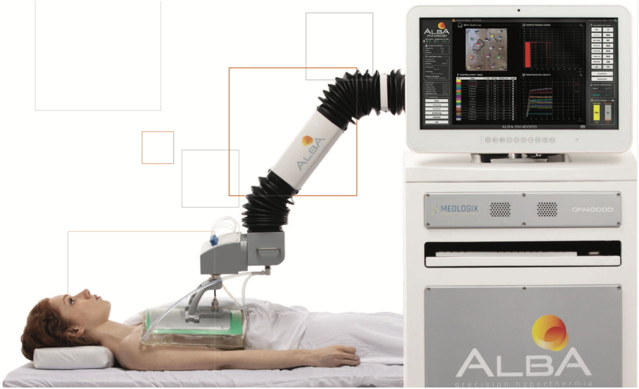
Head, Neck, Breast



Breast and Chest Wall

ALBA ON 4000D

7] (2D WithQ) Two-dimensions



In all previous cases, we have not taken into account the presence of the Metabolic Heat Source. This quantity is modeled as follows:

$$\mathcal{Q} = q_m + q_p \qquad [W/m^3] \tag{26}$$

- $\mathbf{q_m}$ represent the Metabolic Heat Generation of the tissue. This quantity varies with respect to the health of the tissue.
- $\mathbf{q_p}$ represent the Energy Deposition Rate due to the heating procedure. This quantity has been modeled as follows:

$$q_p = \frac{(\text{Absorption Efficiency})(\text{Power Erogated})}{\text{Cancer Volume}} \tag{27}$$

Problems faced

7] (2D WithQ) Two-dimensions

From the previous assumptions and approximations, we can summarize all the inaccuracies here:

- We did not have access to machine specifications like the power erogated
- The absorption rate of the tissue is variable with respect to the its heterogeneous structure
- It is difficult to generalize since the tissue has huge differences from one patient to another
- The positioning of the cancerous tissue has not been taken into account

Results [1]

7] (2D WithQ) Two-dimensions

The **2D (With Q) Results** Optimization Journey starts by varying the **Num_Dense_Layers** hyperparameter. With **2 Dense Layers**:

- **APTx** is the **fastest** Activation function.
- **SiLU** and Swish are the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
ELU	25	41830	0,0295	He normal
GELU	25	26908	0,5346	He normal
ReLU	25	26476	0,1213	He normal
SELU	25	26060	0,1029	Glorot normal
Sigmoid	1000	25363	0,0856	Glorot normal
SiLU	25	25335	**0,0211**	He normal
sin	25	24939	0,1020	Glorot normal
Swish	25	24546	**0,0211**	He normal
Tanh	25	3641	0,0679	Glorot normal
Mish	25	3090	0,0315	He normal
SoftPlus	1000	1943	0,0350	Glorot normal
APTx	25	1917	0,0587	Glorot normal
APTx	25	**1496**	0,0587	He normal

TABLE XVII: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
SiLU	50	17843	**0,0099**	He normal
Swish	50	1056	0,0099	He normal
ELU	50	831	0,1199	He normal
Mish	50	600	0,0237	He normal
SoftPlus	75	**306**	0,2110	Glorot normal

TABLE XVIII: 3 Dense Layers

AF	Iter.	Runtime	Score	Initialization
SiLU	50	63985	0,0205	He normal
Swish	50	983	0,0205	He normal
ELU	50	774	0,0841	He normal
Mish	50	561	**0,0186**	He normal
SoftPlus	50	**355**	0,0245	Glorot normal

TABLE XIX: 4 Dense Layers

Results [1]

7] (2D WithQ) Two-dimensions

The **2D (With Q) Results** Optimization continue by analysing with **3 Dense Layers**:

- **SoftPlus** is the **fastest** Activation function.
- **SiLU** is the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
ELU	25	41830	0,0295	He normal
GELU	25	26908	0,5346	He normal
ReLU	25	26476	0,1213	He normal
SELU	25	26060	0,1029	Glorot normal
Sigmoid	1000	25363	0,0856	Glorot normal
SiLU	25	25335	**0,0211**	He normal
sin	25	24939	0,1020	Glorot normal
Swish	25	24546	**0,0211**	He normal
Tanh	25	3641	0,0679	Glorot normal
Mish	25	3090	0,0315	He normal
SoftPlus	1000	1943	0,0350	Glorot normal
APTx	25	1917	0,0587	Glorot normal
APTx	25	**1496**	0,0587	He normal

TABLE XVII: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
SiLU	50	17843	**0,0099**	He normal
Swish	50	1056	0,0099	He normal
ELU	50	831	0,1199	He normal
Mish	50	600	0,0237	He normal
SoftPlus	75	**306**	0,2110	Glorot normal

TABLE XVIII: 3 Dense Layers

AF	Iter.	Runtime	Score	Initialization
SiLU	50	63985	0,0205	He normal
Swish	50	983	0,0205	He normal
ELU	50	774	0,0841	He normal
Mish	50	561	**0,0186**	He normal
SoftPlus	50	**355**	0,0245	Glorot normal

TABLE XIX: 4 Dense Layers

Results [1]

7] (2D WithQ) Two-dimensions

The **2D (With Q) Results** Optimization continue by analysing with **4 Dense Layers**:

- **SoftPlus** is the **fastest** Activation function.
- **Mish** is the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
ELU	25	41830	0,0295	He normal
GELU	25	26908	0,5346	He normal
ReLU	25	26476	0,1213	He normal
SELU	25	26060	0,1029	Glorot normal
Sigmoid	1000	25363	0,0856	Glorot normal
SiLU	25	25335	**0,0211**	He normal
sin	25	24939	0,1020	Glorot normal
Swish	25	24546	**0,0211**	He normal
Tanh	25	3641	0,0679	Glorot normal
Mish	25	3090	0,0315	He normal
SoftPlus	1000	1943	0,0350	Glorot normal
APTx	25	1917	0,0587	Glorot normal
APTx	25	**1496**	0,0587	He normal

TABLE XVII: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
SiLU	50	17843	**0,0099**	He normal
Swish	50	1056	0,0099	He normal
ELU	50	831	0,1199	He normal
Mish	50	600	0,0237	He normal
SoftPlus	75	**306**	0,2110	Glorot normal

TABLE XVIII: 3 Dense Layers

AF	Iter.	Runtime	Score	Initialization
SiLU	50	63985	0,0205	He normal
Swish	50	983	0,0205	He normal
ELU	50	774	0,0841	He normal
Mish	50	561	**0,0186**	He normal
SoftPlus	50	**355**	0,0245	Glorot normal

TABLE XIX: 4 Dense Layers

Results [2]

7] (2D WithQ) Two-dimensions

We continue the **2D (Without Q) Results** Optimization Journey by varying the **Num_Dense_Nodes** hyperparameter. With **50 Dense Nodes**:

- **APTx** is the **fastest** Activation function.
- **SiLU** and **Swish** are the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
ELU	25	41830	0,0295	He normal
GELU	25	26908	0,5346	He normal
ReLU	25	26476	0,1213	He normal
SELU	25	26060	0,1029	Glorot normal
Sigmoid	1000	25363	0,0856	Glorot normal
SiLU	25	25335	**0,0211**	He normal
sin	25	24939	0,1020	Glorot normal
Swish	25	24546	**0,0211**	He normal
Tanh	25	3641	0,0679	Glorot normal
Mish	25	3090	0,0315	He normal
SoftPlus	1000	1943	0,0350	Glorot normal
APTx	25	1917	0,0587	Glorot normal
APTx	25	**1496**	0,0587	He normal

TABLE XVII: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
SiLU	1000	1023	**0,0394**	He normal
Swish	1000	820	**0,0394**	He normal
ELU	50	783	0,3171	He normal
Mish	50	580	0,0954	He normal
SoftPlus	500	**196**	0,1695	Glorot normal

TABLE XX: Dense Layers 75

AF	Iter.	Runtime	Score	Initialization
SiLU	50	16306	**0,0099**	He normal
Swish	50	1085	**0,0099**	He normal
ELU	50	864	0,1199	He normal
Mish	50	623	0,0237	He normal
SoftPlus	100	**277**	0,2059	Glorot normal

TABLE XXI: Dense Layers 100

Results [2]

7] (2D WithQ) Two-dimensions

The **2D (Without Q) Results** Optimization Journey continue by analysing **75 Dense Nodes**:

- **SoftPlus** is the **fastest** Activation function.
- **SiLU** and **Swish** are the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
ELU	25	41830	0,0295	He normal
GELU	25	26908	0,5346	He normal
ReLU	25	26476	0,1213	He normal
SELU	25	26060	0,1029	Glorot normal
Sigmoid	1000	25363	0,0856	Glorot normal
SiLU	25	25335	**0,0211**	He normal
sin	25	24939	0,1020	Glorot normal
Swish	25	24546	**0,0211**	He normal
Tanh	25	3641	0,0679	Glorot normal
Mish	25	3090	0,0315	He normal
SoftPlus	1000	1943	0,0350	Glorot normal
APTx	25	1917	0,0587	Glorot normal
APTx	25	**1496**	0,0587	He normal

TABLE XVII: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
SiLU	1000	1023	**0,0394**	He normal
Swish	1000	820	**0,0394**	He normal
ELU	50	783	0,3171	He normal
Mish	50	580	0,0954	He normal
SoftPlus	500	**196**	0,1695	Glorot normal

TABLE XX: Dense Layers 75

AF	Iter.	Runtime	Score	Initialization
SiLU	50	16306	**0,0099**	He normal
Swish	50	1085	**0,0099**	He normal
ELU	50	864	0,1199	He normal
Mish	50	623	0,0237	He normal
SoftPlus	100	**277**	0,2059	Glorot normal

TABLE XXI: Dense Layers 100

Results [2]

7] (2D WithQ) Two-dimensions

The **2D (Without Q) Results** Optimization Journey continue by analysing **100 Dense Nodes**:

- **SoftPlus** is the **fastest** Activation function.
- **SiLU** and **Swish** are the **best** in terms of Composite Score.

AF	Iter.	Runtime	Score	Initialization
ELU	25	41830	0,0295	He normal
GELU	25	26908	0,5346	He normal
ReLU	25	26476	0,1213	He normal
SELU	25	26060	0,1029	Glorot normal
Sigmoid	1000	25363	0,0856	Glorot normal
SiLU	25	25335	**0,0211**	He normal
sin	25	24939	0,1020	Glorot normal
Swish	25	24546	**0,0211**	He normal
Tanh	25	3641	0,0679	Glorot normal
Mish	25	3090	0,0315	He normal
SoftPlus	1000	1943	0,0350	Glorot normal
APTx	25	1917	0,0587	Glorot normal
APTx	25	**1496**	0,0587	He normal

TABLE XVII: 2 Dense Layers

AF	Iter.	Runtime	Score	Initialization
SiLU	1000	1023	**0,0394**	He normal
Swish	1000	820	**0,0394**	He normal
ELU	50	783	0,3171	He normal
Mish	50	580	0,0954	He normal
SoftPlus	500	**196**	0,1695	Glorot normal

TABLE XX: Dense Layers 75

AF	Iter.	Runtime	Score	Initialization
SiLU	50	16306	**0,0099**	He normal
Swish	50	1085	**0,0099**	He normal
ELU	50	864	0,1199	He normal
Mish	50	623	0,0237	He normal
SoftPlus	100	**277**	0,2059	Glorot normal

TABLE XXI: Dense Layers 100

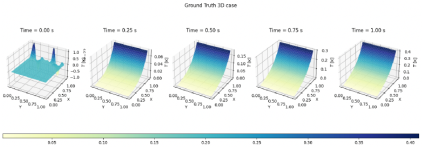
A

Results [3]

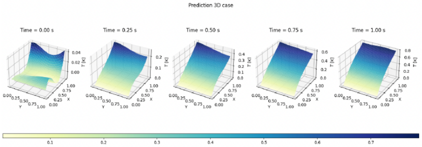
7] (2D WithQ) Two-dimensions

By looking at the **Mish** and **SiLU performance 3D graphs**, we note that:

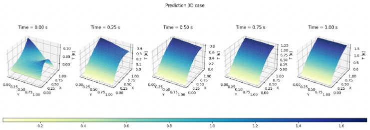
- **Swish** has better **RunTime.**
- **SiLU** has smoother profiles and better **Scores.**



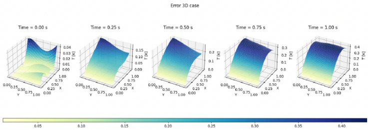
(a) Ground Truth



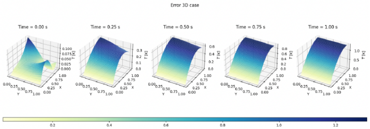
(b) Mish prediction



(c) SiLU prediction



(d) Mish error

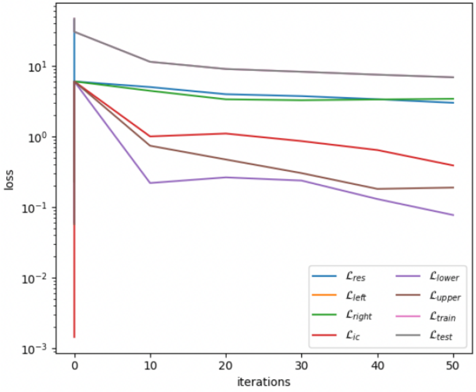


(e) SiLU error

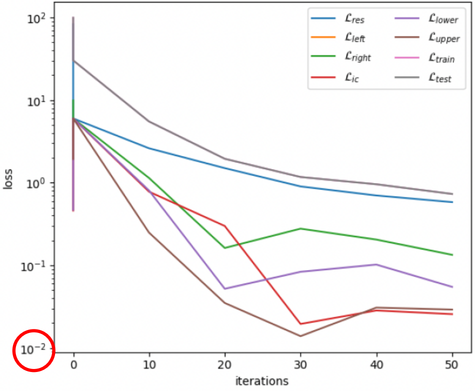
Results [4]

7] (2D WithQ) Two-dimensions

By looking at the **Mish** and **SiLU Losses graphs**, we note that:



(a) Mish



Smaller Loss !

(b) SiLU

Results [5]

7] (2D WithQ) Two-dimensions

Focusing on the **HW Complexities**, by analysing the **CPU**, **Disk** and **Memory Utilization** results, we observe:

- **ELU** and **APTx** require **less CPU Utilization(%)**, **TanH more Memory Utilization (MB)**.
- **APTx**, **SoftPlus** require **more Disk Utilization (GB)**.

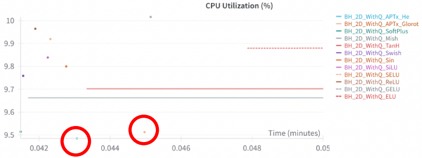


Fig. 18: CPU Utilization (%)

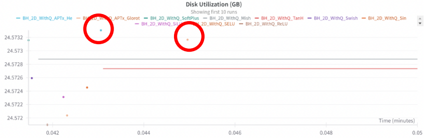


Fig. 19: Disk Utilization (GB)

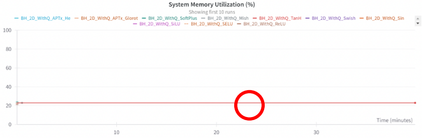


Fig. 20: System Memory Utilization (%)

Table of Contents

8] Conclusion

Future Formula Adjustments

To address the assumptions and improve future models:

- **Anisotropy in Tissue Perfusion**: Use an anisotropic thermal conductivity tensor to account for directional differences in heat conduction:

$$\nabla \cdot (k_{\text{eff}} \nabla T) \tag{28}$$

- **Non-linear Perfusion Model**: Introduce a temperature-dependent perfusion term to capture the non-proportional relationship between blood perfusion and temperature difference:

$$h_b = \omega_b(T)\rho_b c_b(T_a - T) \tag{29}$$

Future Work

8] Conclusion

Future work can improve some assumptions like those in the previous slide, not only in terms of formulas but also in methodology:

- Apply the model to new types of tissues for broader medical applications
- Model the surface with multiple coordinates to better represent complex geometries
- Develop new architectures for the neural network and explore new configurations
- Find more accurate values for the simulations
- Combine the strengths of the Adam and L-BFGS optimization algorithms to improve convergence speed and accuracy
- Model the heat generation term $\mathcal{Q}$ with a temporal dependency

Objective Completed

8] Conclusion

- PINNs are effective for real-time medical thermal treatments, balancing computational efficiency and prediction accuracy.
- Optimization techniques, such as Adam and L-BFGS, reveal that ELU and SiLU activation functions, combined with He normal initialization, yield the best results.
- Costs and scores for each scenario demonstrate flexibility for future experiments.
- Results were obtained in both 1D and 2D cases, with and without heat generation (Q).

Thank You for Your Attention!