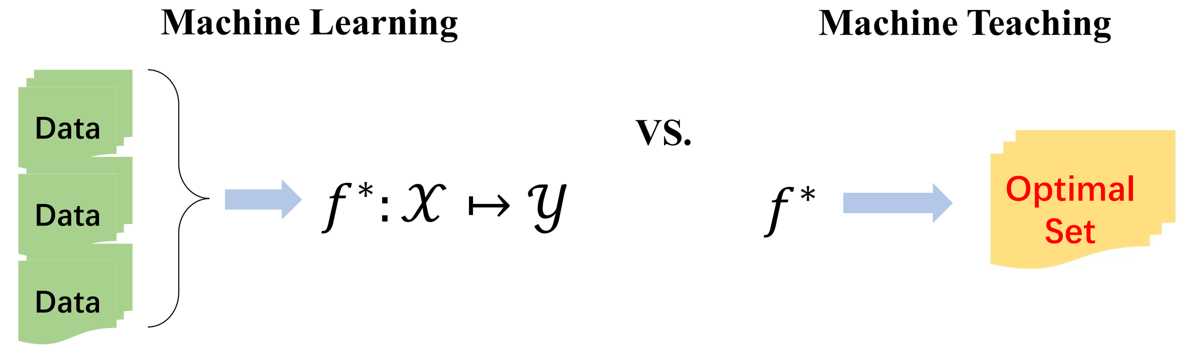
Nonparametric Teaching of Implicit Neural Representations

Chen Zhang[1*], Steven Tin Sui Luo[2*], Jason Chun Lok Li[1], Yik-Chung Wu[1], Ngai Wong[1]
[1]The University of Hong Kong [2]The University of Toronto

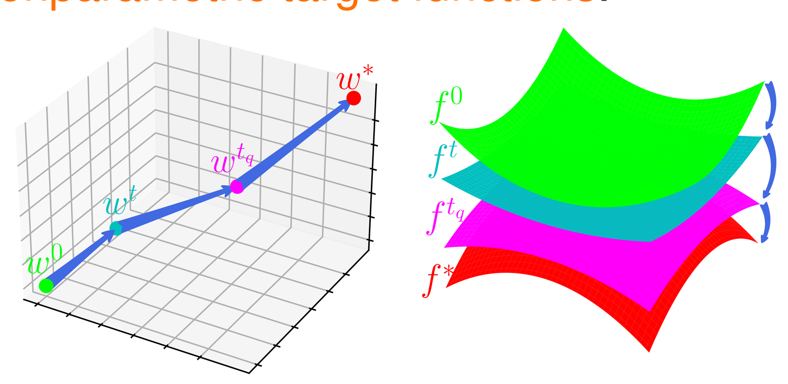
Nonparametric Teaching

Nonparametric teaching (NT) (Zhang et al., 2023b;a) presents a theoretical framework to facilitate efficient example selection when the target function is nonparametric, i.e., implicitly defined.

Specifically, *machine teaching* (Zhu, 2015; Liu et al., 2017; Zhu et al., 2018) considers the design of a training set (dubbed the teaching set) for the learner, with the goal of enabling speedy convergence towards target functions.



NT (Zhang et al., 2023b;a) relaxes the assumption of target functions[†] f being parametric (Liu et al., 2017; 2018), which is f can be represented by a set of parameters $\boldsymbol{w}$, *e.g.*, $f(\boldsymbol{x}) = \langle \boldsymbol{w}, \boldsymbol{x} \rangle$ with input $\boldsymbol{x}$, to encompass the teaching of nonparametric target functions.



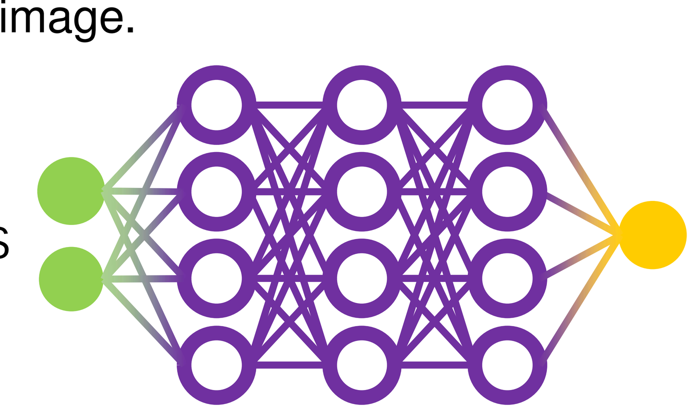
(a) Parametric IMT (b) Nonparametric IMT

[†]The loss $\mathcal{L}$ can be general for different tasks, *e.g.*, square loss for regression and hinge loss for classification.

Implicit Neural Representations

Implicit neural representation (INR) (Sitzmann et al., 2020b; Tancik et al., 2020) focuses on modeling a given signal, which is often discrete, through the use of an overparameterized multilayer perceptron (MLP) such that the signal is accurately fitted by this MLP preserving great details.

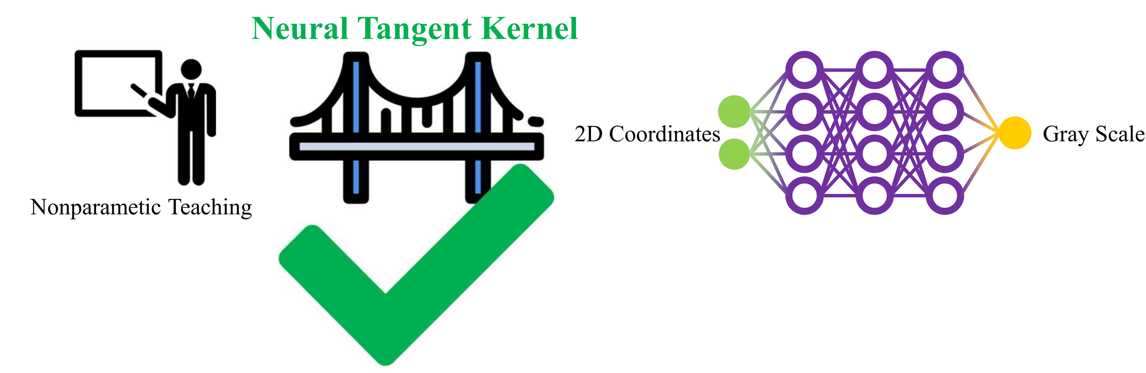
Such an overparameterized MLP inputs low-dimensional coordinates of the given signal and outputs corresponding values for each input location, *e.g.*, the MLP maps 2D input coordinates to their respective 8-bit levels for a grayscale image.



The Bridge Between NT and INRs: Neural Tangent Kernel

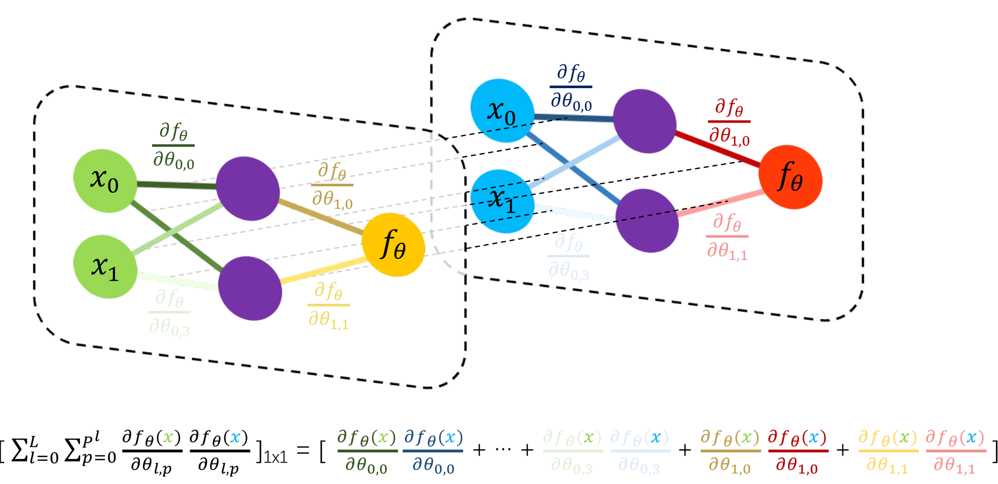
The evolution of an MLP is typically achieved by gradient descent on its parameters, whereas nonparametric teaching involves functional gradient descent as the means of function evolution.

Bridging this (**theoretical** + **practical**) gap is of great value and calls for more examination prior to the application of nonparametric teaching algorithms in the context of INR.



Neural Tangent Kernel (Jacot et al., 2018; Lee et al., 2019) is a symmetric and positive definite kernel function, which is derived from the analysis of the evolution of a neural network (the MLP is considered).

$$K_{\theta^t}(\boldsymbol{x}_i, \cdot) = \left\langle \frac{\partial f_\theta}{\partial \theta}\bigg|_{\cdot,\theta^t}, \frac{\partial f_\theta}{\partial \theta}\bigg|_{\boldsymbol{x}_i,\theta^t} \right\rangle \quad (1)$$

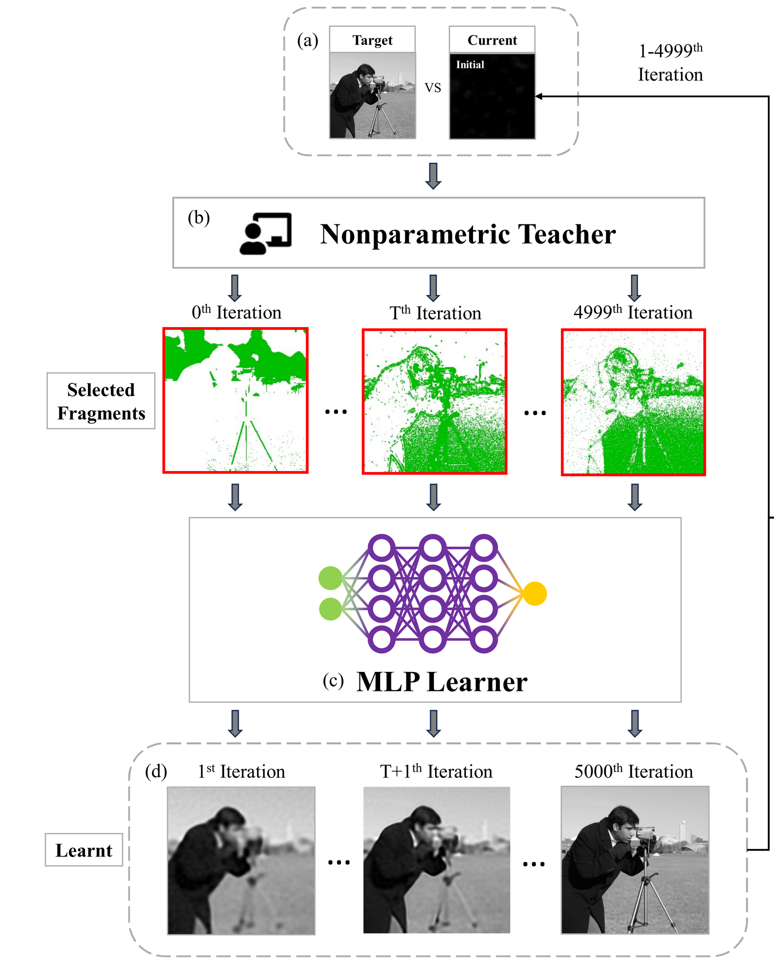


Main Contribution

Our key contributions are:

- We propose Implicit Neural Teaching (INT) that novelly interprets implicit neural representation (INR) via the theoretical lens of nonparametric teaching, which in turn enables the utilization of greedy algorithms from the latter to effectively bolster the training efficiency of INRs.
- We unveil a strong link between the evolution of a multilayer perceptron (MLP) using gradient descent on its parameters and that of a function using functional gradient descent in nonparametric teaching. This connects nonparametric teaching to MLP training, thus expanding the applicability of nonparametric teaching towards deep learning. We further show that the dynamic NTK, derived from gradient descent on the parameters, converges to the canonical kernel of functional gradient descent.
- We showcase the effectiveness of INT through extensive experiments in INR training across multiple modalities. Specifically, INT saves training time for 1D audio (-31.63%), 2D images (-38.88%) and 3D shapes (-35.54%), while upkeeping its reconstruction quality.

INT Workflow and Algorithm



Algorithm 1 Implicit Neural Teaching

Input: Target signal f^*, initial MLP f_{θ^0}, the size of selected training size $k \leq N$, small constant $\epsilon > 0$ and maximal iteration number T.

Set $f_{\theta^t} \leftarrow f_{\theta^0}$, $t = 0$.

while $t \leq T$ and $\|[f_{\theta^t}(\boldsymbol{x}_i) - f^*(\boldsymbol{x}_i)]_N\|_2 \geq \epsilon$ **do**

  The teacher selects k teaching examples:

  `/* Examples corresponding to the` k `largest` $|f_{\theta^t}(\boldsymbol{x}_i) - f^*(\boldsymbol{x}_i)|$. `*/`

  $\{\boldsymbol{x}_i\}_k{}^* = \underset{\{\boldsymbol{x}_i\}_k \subseteq \{\boldsymbol{x}_i\}_N}{\arg\max} \|[f_{\theta^t}(\boldsymbol{x}_i) - f^*(\boldsymbol{x}_i)]_k\|_2$.

  Provide $\{\boldsymbol{x}_i\}_k{}^*$ to the MLP learner.

  The learner updates f_{θ^t} based on received $\{\boldsymbol{x}_i\}_k{}^*$:

  `// Parameter-based gradient descent.`

  $\theta^t \leftarrow \theta^t - \frac{\eta}{k} \sum_{\boldsymbol{x}_i \in \{\boldsymbol{x}_i\}_k{}^*} \nabla_\theta \mathcal{L}(f_{\theta^t}(\boldsymbol{x}_i), f^*(\boldsymbol{x}_i))$.

  Set $t \leftarrow t + 1$.

end

Experiments and Results

- **Toy 2D Cameraman fitting.**

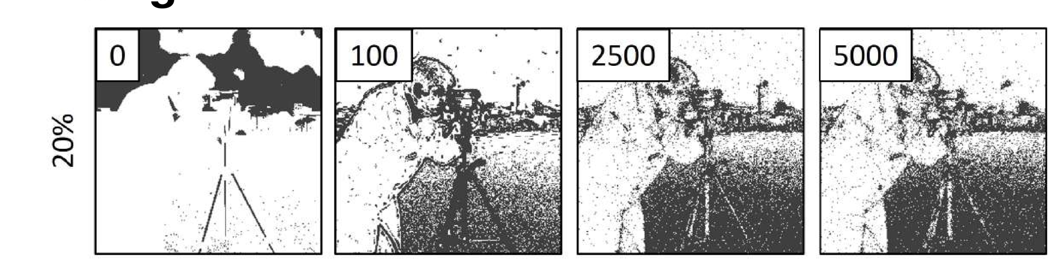


Figure: Progression of INT selected pixels (marked as black) at corresponding iterations when training with INT 20%.

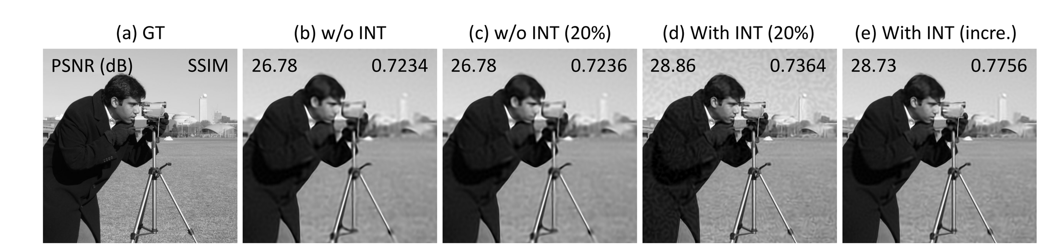


Figure: Reconstruction quality of SIREN. (b) trains SIREN without (w/o) INT using all pixels. (c) trains it w/o INT using 20% randomly selected pixels. (d) trains it using INT of 20% selection rate. (e) trains it using progressive INT (*i.e.*, increasing selection rate progressively from 20% to 100%).

- **INT on multiple real-world modalities.**

The encoding time is measured excluding data I/O latency.

INT	Modality	Time (s)	PSNR(dB) / IoU(%) ↑
✗	Audio	23.05	48.38±3.50
	Image	345.22	36.09±2.51
	Megapixel	16.78K	31.82
	3D Shape	144.58	97.07±0.84
✓	Audio	15.76 (-31.63%)	48.15±3.39
	Image	211.04 (-38.88%)	36.97±3.59
	Megapixel	11.87K (-29.26%)	33.01
	3D Shape	93.19 (-35.54%)	96.68±0.83