Implicit Representations via Operator Learning

Sourav Pal, Harshavardhan Adepu, Clinton Wang, Polina Golland, Vikas Singh

https://github.com/vsingh-group/oinr

ICML International Conference On Machine Learning

The University of Wisconsin–Madison

Massachusetts Institute of Technology

OVERVIEW

O-INR learns parameterized maps between function spaces (operators) providing a generalized view of Implicit Neural Representations (INR).

From INR to O-INR

INR: $m_\theta : \mathbb{R}^D \to \mathbb{R}^R$

Where, θ denotes the parameterization; D and R denote the dimension of the domain and range space. In case of 2D images ; $D = 2$ and $R = 3$.

Positional Encoding:

$$f(x,y) = \left[\sin(2^l \pi x), \cos(2^l \pi x), \sin(2^l \pi y), \cos(2^l \pi y), \ldots\right]$$

$$\Downarrow$$

$$f(x,y) = \left[\sin(\theta x), \cos(\theta x), \sin(\theta y), \cos(\theta y)\right]$$

Embedding Space:

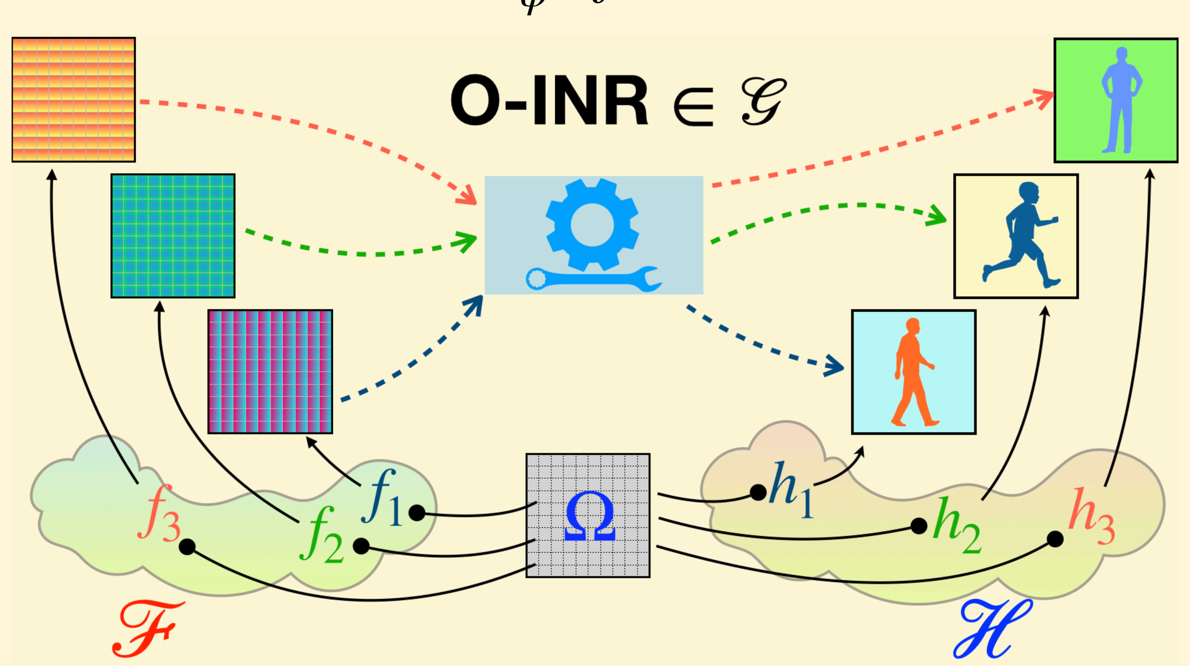
$$\mathcal{F} = \{f_1, f_2, \ldots\} \quad \text{where} \quad f_i : \Omega \to \psi$$

Signal Space:

$$\mathcal{H} = \{h_1, h_2, \ldots\} \quad \text{where} \quad h_i : \Omega \to \phi$$

$h_1, h_2, \ldots$ are different signals, e.g. frames in a video.

O-INR: $\mathcal{G}_\phi : f \to h$



INTEGRAL OPERATORS

Integral operators, (functional $\mathscr{C}$). transform between function spaces by integrating over the domain of definition Ω.

$$h(x) = \int_{y \in \Omega} \mathscr{C}(x, f(y))\mathrm{d}y \quad \text{where} \;, f \in \mathcal{F} \text{ and } h \in \mathcal{H}$$

We can learn an integral operator $\mathcal{G}_\phi$ with associated kernel $\mathcal{K}_\phi$ where ϕ denotes the parameterization.

$$h(\omega) = \mathcal{G}\left[f\right](\omega) = \int_{\omega' \in \Omega} \mathcal{K}_\phi(\omega, \omega') f(\omega')\mathrm{d}\omega', \quad \omega \in \Omega$$

The kernel is the only parameterization needed for O-INR.

With the Dirac-Delta function, we can recover standard INR.

GREEN'S THEOREM

For a linear differential operator, $\mathscr{L}$, the Green's function of the operator, $G(\omega, \omega')$ is the fundamental solution to the operator in its Dirac delta inhomogeneous form.

$$\mathscr{L}G(\omega, \omega') = \delta(\omega' - \omega)$$

The property is used to solve the general inhomogeneous form with the forcing function f.

$$\mathscr{L}u(\omega) = f(\omega)$$

via an integral to produce the solution u.

$$u(x) = \int G(\omega, \omega') f(\omega')\mathrm{d}\omega'$$

which, reveals the kernel as the Green's function.

Hence, O-INRs can be viewed as different realizations of the Green's function.

PARAMETERIZATION OF O-INR

The only parameterization of O-INR is via the kernel $\mathcal{K}_\phi$

An efficient choice is to consider the associated kernel to be a convolution kernel.

$$\mathcal{K}_\phi(\omega, \omega') = g_\phi(\omega - \omega')$$

with g_ϕ being standard convolution with parameters ϕ.

O-INR then becomes

$$h(\omega) = \mathcal{G}\left[f\right](\omega) = \int_{\omega' \in \Omega} g_\phi(\omega - \omega') f(\omega')\mathrm{d}\omega'$$

EVALUATION

2D IMAGE

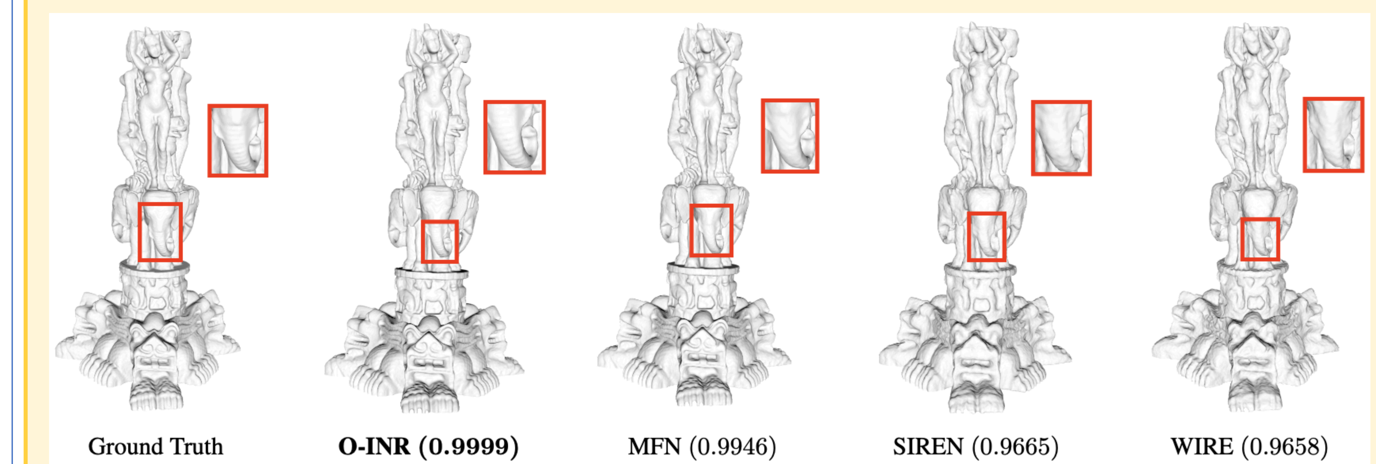
Ground truth | O-INR (32.14) | SIREN (31.84) | WIRE (22.1) | MFN (32.67)

Ground Truth | **O-INR (ours)** (32.9dB) | SIREN (30.4dB) | WIRE (31.5dB) | MFN (30.9dB)

DENOISING

Ground Truth | Noisy | O-INR (24.28) | SIREN (20.94) | WIRE (24.48) | **MFN (25.22)**

3D VOLUME



Ground Truth | **O-INR (0.9999)** | MFN (0.9946) | SIREN (0.9665) | WIRE (0.9658)

VIDEO

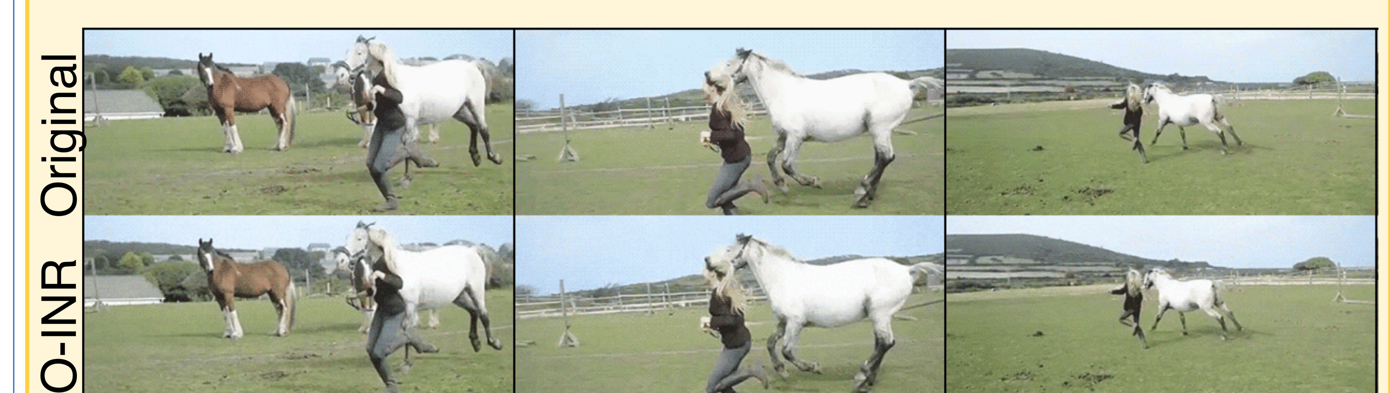
We consider each frame of the video as a different signal defined on the same domain. O-INR is trained to learn the map from encodings (f_n) of this domain Ω to the frames (h_n).

$$\mathcal{G}_\phi : \mathcal{F}_N \to \mathcal{H}_N$$

$$\mathcal{F}_N = \{f_n(\Omega) \,|\, n \in \mathbb{N}\}; \quad \mathcal{H}_N = \{h_n(\Omega) \,|\, n \in \mathbb{N}\}$$

$$f_n(x,y) = [\sin(2^l \pi x) + \gamma, \cos(2^l \pi x) + \gamma, \ldots]$$

$$\gamma_n = \alpha + ((\beta - \alpha)/N)n$$



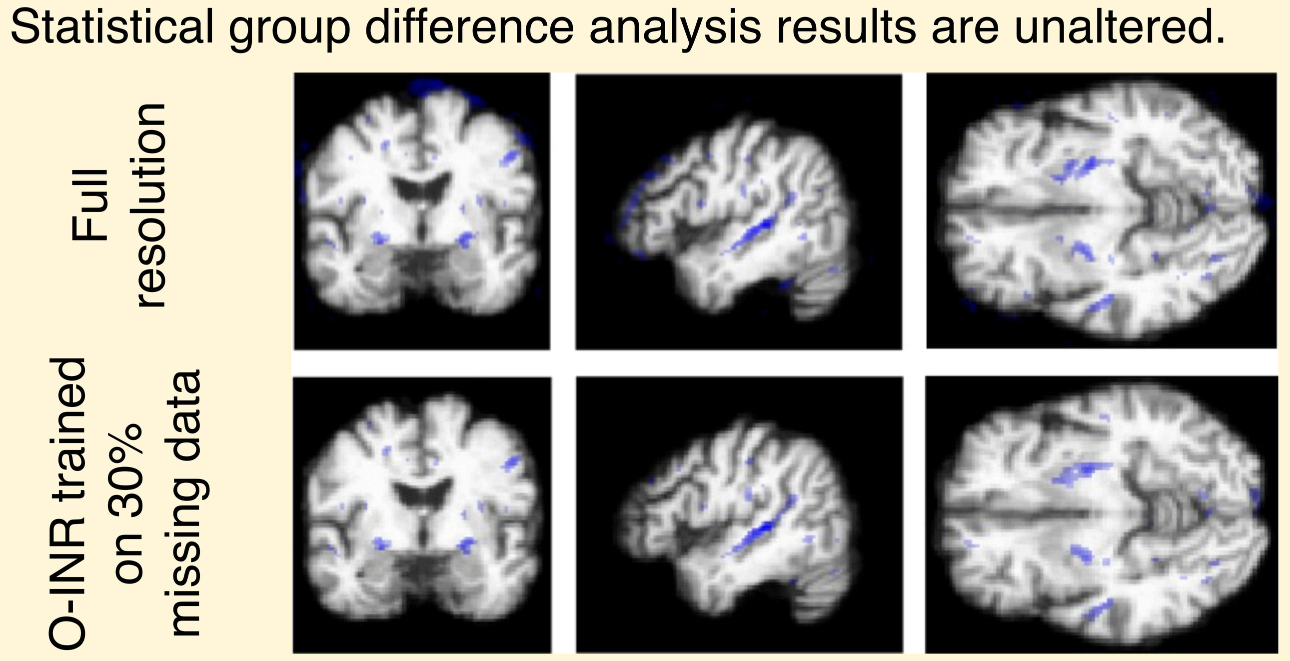
BRAIN IMAGING: SLICE IMPUTATION

Processing clinical scans with lower out of plane resolution can be achieved by O-INR where available scans are treated as related signals defined on the same domain.

O-INR can represent data even when 80% slices are dropped.

Percentage missing	50	66	75	80
Train MSE	2.17 e -5	3.10 e -5	1.39 e -6	1.58 e -5
Test MSE	8.33 e -5	2.38 e -4	4.70 e -4	7.66 e -4

Statistical group difference analysis results are unaltered.



DERIVATIVE

O-INR can exploit convolution theorem to compute derivatives on the signal space seamlessly in one forward pass.

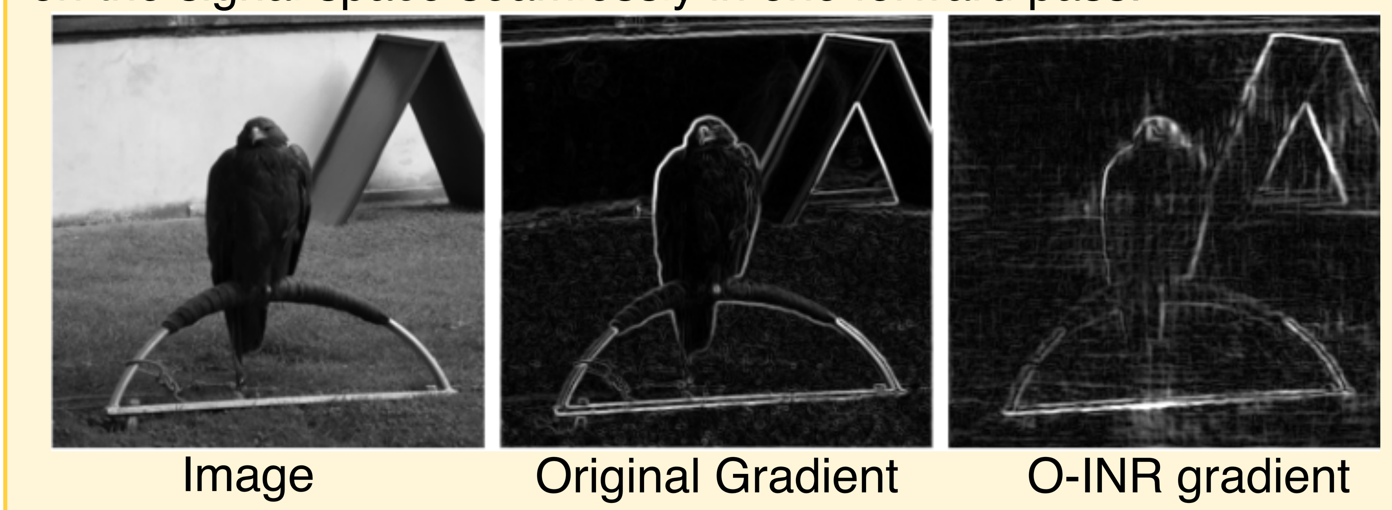


Image | Original Gradient | O-INR gradient

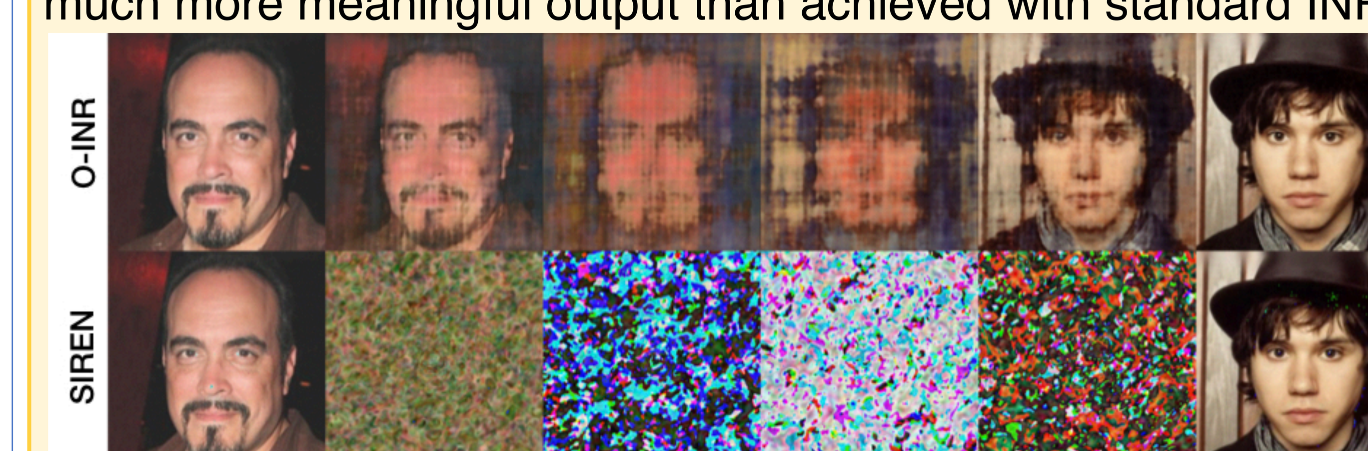
CALDERON-ZYGMUND OPERATOR

O-INR can be realized with operators other than convolution such as pseudo-differential operators including CZ operator.

Original | **CZ-operator** (36.47) | **SIREN** (25.64)

WEIGHT SPACE INTERPOLATION

Interpolating the trained weights of 2 different O-INR produces much more meaningful output than achieved with standard INR.



WEIGHT SPACE CLASSIFICATION

Downstream tasks such as classification on the weight space of trained O-INR can achieve state of the art results.

	MNIST	Fashion-MNIST
DWSNets	85.71	67.06
O-INR	96.57	80.18

CONCLUSION

O-INR provides a generalization of INR via transformation between function spaces with a learnable operator.

It does not need MLPs yet can achieve comparable performance.

Downstream signal processing tasks are much easier.

Weight space of O-INR is more structured.

International Conference on Machine Learning (ICML), 2024

Research supported by NIH RF1 AG059312, Vilas Board of Trustees, MIT-CSAIL-Wistron Program