

RHINO: Regularizing the Hash-based Implicit Neural Representation

Hao Zhu¹, Fengyi Liu¹, Qi Zhang², Zhan Ma^{1*} & Xun Cao^{1*}

¹*School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China*

²*Vivo, Xi'an 710055, China*

Appendix A GTK analysis on RHINO

The recent study [1] "Grounding and Enhancing Grid-based Models for Neural Fields," introduces a theoretical framework called Grid Tangent Kernels (GTKs) for hash-based INRs. In [1], the formulation of the grid-based model is given by,

$$g(\mathbf{x}, \mathbf{w}) = \sum_{i \in U(\mathbf{x})} \varphi(\mathbf{x}, \Theta_i) \mathbf{w}_i, \quad (\text{A1})$$

and its corresponding GTK is defined as,

$$[\mathbf{G}_g(t)]_{i,j} = \left\langle \frac{\partial g(\mathbf{X}_i, \mathbf{w}(t))}{\partial \mathbf{w}}, \frac{\partial g(\mathbf{X}_j, \mathbf{w}(t))}{\partial \mathbf{w}} \right\rangle. \quad (\text{A2})$$

Using the GTK, we have tried to further analyse the approximation and generalization properties of the proposed RHINO.

However, it is found that the GTK framework is unavailable for analysing our work due to the unique characteristics of RHINO, represented by $f_\theta(\mathcal{H}_i, \mathcal{T}(x_i))$. Specifically, because the RHINO does not modify the architecture of the original hash-table and the new branch $\mathcal{T}(x_i)$ is concatenated with the hash-table, RHINO could not be expressed by formulation A1. Formally,

$$\begin{aligned} & [\mathbf{G}_{RHINO}(t)]_{i,j} \\ &= \left\langle \frac{\partial \begin{bmatrix} \mathcal{H}(\mathbf{X}_i, \mathbf{w}(t)) \\ \mathcal{T}(\mathbf{X}_i) \end{bmatrix}}{\partial \mathbf{w}}, \frac{\partial \begin{bmatrix} \mathcal{H}(\mathbf{X}_j, \mathbf{w}(t)) \\ \mathcal{T}(\mathbf{X}_j) \end{bmatrix}}{\partial \mathbf{w}} \right\rangle \\ &= \left\langle \frac{\partial \mathcal{H}(\mathbf{X}_i, \mathbf{w}(t))}{\partial \mathbf{w}}, \frac{\partial \mathcal{H}(\mathbf{X}_j, \mathbf{w}(t))}{\partial \mathbf{w}} \right\rangle + \left\langle \frac{\partial \mathcal{T}(\mathbf{X}_i)}{\partial \mathbf{w}}, \frac{\partial \mathcal{T}(\mathbf{X}_j)}{\partial \mathbf{w}} \right\rangle \\ &= [\mathbf{G}_{\mathcal{H}}(t)]_{i,j} + 0 \\ &= [\mathbf{G}_{\mathcal{H}}(t)]_{i,j} \end{aligned} \quad (\text{A3})$$

since $\mathcal{T}(x_i)$ does not contribute to the partial derivative of the hash features $\mathbf{w}$, the GTKs of RHINO are not changed.

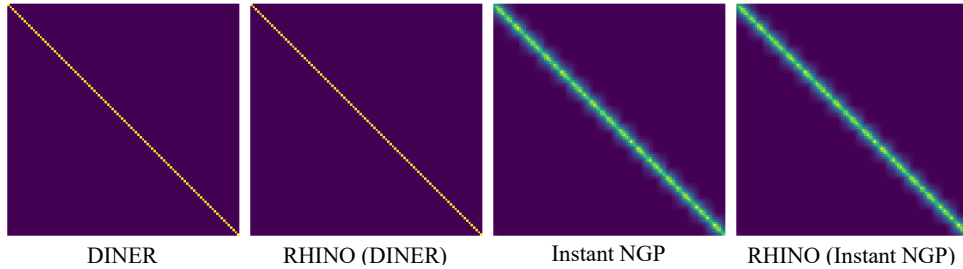


Figure A1 Visualizations of GTKs.

* Corresponding author (email: mazhan@nju.edu.cn, caoxun@nju.edu.cn)

We compare RHINO on two hash-based INRs: DINER [2] and Instant NGP [3]. Following the methodology in [1], each model is trained to fit a 2D image until convergence, after which we sample 100 data points from the training image to compute the GTKs of different methods. As illustrated in Fig. A1, The GTK of RHINO (DINER) is the same as DINER, and RHINO (Instant NGP) shares the same GTK as Instant NGP.

References

- 1 Zhao Z, Fan F, Liao W, Yan J (2024) Grounding and enhancing grid-based models for neural fields. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp 19425–19435
- 2 Zhu H, Xie S, Liu Z, Liu F, Zhang Q, Zhou Y, Lin Y, Ma Z, Cao X (2024) Disorder-invariant implicit neural representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 46(8):5463-5478
- 3 Müller T, Evans A, Schied C, Keller A (2022) Instant neural graphics primitives with a multiresolution hash encoding. *ACM Transactions on Graphics (ToG)* 41(4):1–15