Poly-INR: Towards Capturing Multiplicative Interactions for Enhanced Implicit Neural Representation of Videos (Selected Results)

Snehal Singh Tomar

snehal@smail.iitm.ac.in

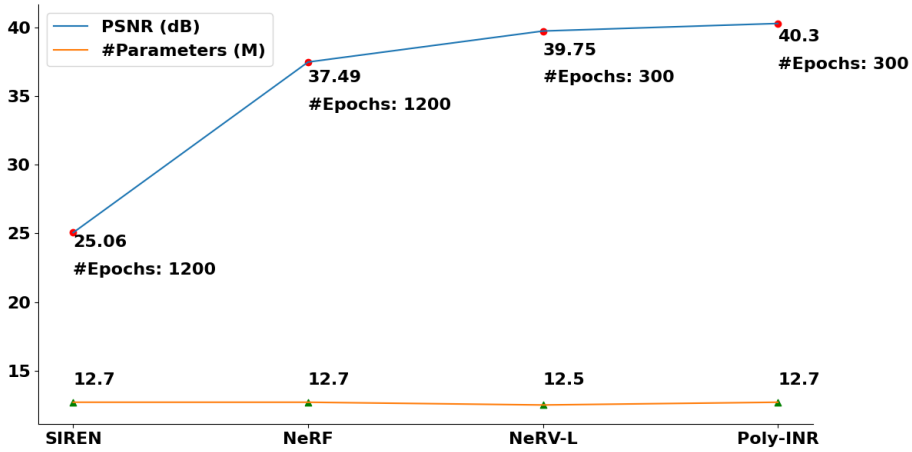

Figure 1: Ploy-INR outperforms SOTA Video INR methods SIREN [1], NeRF [2], and NeRV [3] while having a comparable number of parameters and often being trained for significantly lesser number of epochs. These results are with respect to the "Big Buck Bunny" sequence from scikit-video.

Figure 2: Visualization of reconstructed frames when activation function is not used. We highlight the detailed regions in red boxes which our model is able to reconstruct successfully whereas NeRV [3] fails to reconstruct the same.

References

[1] V. Sitzmann, J. Martel, A. Bergman, D. Lindell, and G. Wetzstein, "Implicit neural representations with periodic activation functions," in *Advances in Neural Information Processing Systems*, H. Larochelle, M. Ranzato, R. Hadsell, M. Balcan, and H. Lin, Eds., vol. 33. Curran Associates, Inc., 2020, pp. 7462–7473. [Online]. Available: https://proceedings.neurips.cc/paper/2020/file/53c04118df112c13a8c34b38343b9c10-Paper.pdf

[2] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "Nerf: Representing scenes as neural radiance fields for view synthesis," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2020. [Online]. Available: http://arxiv.org/abs/2003.08934v2

[3] H. Chen, B. He, H. Wang, Y. Ren, S.-N. Lim, and A. Shrivastava, "NeRV: Neural representations for videos," in *Advances in Neural Information Processing Systems*, A. Beygelzimer, Y. Dauphin, P. Liang, and J. W. Vaughan, Eds., 2021. [Online]. Available: https://openreview.net/forum?id=BbikqBWZTGB