PNeRV: A Polynomial Neural Representation for Videos (Selected Results)

Snehal Singh Tomar

snehal@smail.iitm.ac.in

Table 1: **Quantitative comparisons** in terms of PSNR (dB) with respect to reconstruction on the Scikit-Bunny video and the UVG datset. PNeRV achieves SOTA performance while maintaining significantly fewer parameters and being up to 4× faster in terms of rate of convergence.

Method	# Params (M)↓	Bunny	Beauty	Bosphorus	Bee	Jockey	SetGo	Shake	Yacht
NeRV-L	12.57	39.63	36.06	37.35	41.23	38.14	31.86	37.22	32.45
HNeRV	11.90	36.23	36.17	30.20	41.58	28.55	29.67	32.44	25.50
E-NeRV	12.49	42.87	36.72	40.06	41.74	39.35	34.68	39.32	35.58
Ours	**11.89**	**44.90**	**39.8**	**41.86**	**43.98**	**39.84**	**35.82**	**41.37**	**36.93**
Gain over E-NeRV	**↓ 0.6**	**↑ 2.03**	**↑ 3.08**	**↑ 1.8**	**↑ 2.24**	**↑ 0.49**	**↑ 1.14**	**↑ 2.05**	**↑ 1.35**

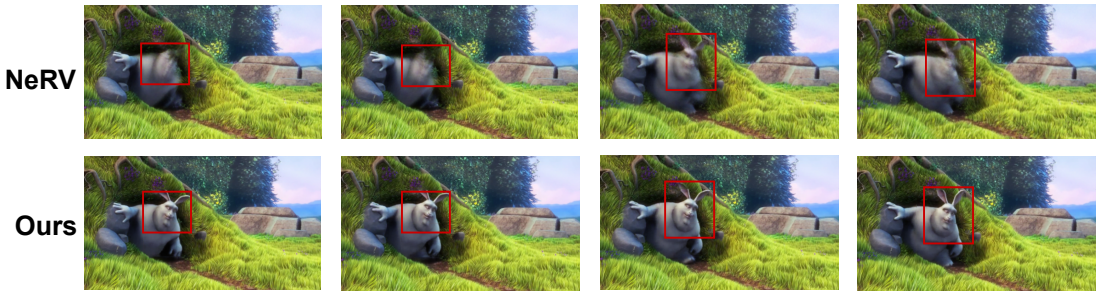

Figure 1: Visualization of reconstructed frames when activation function is not used. We highlight the detailed regions in red boxes which our model is able to reconstruct successfully whereas NeRV [1] fails to reconstruct the same.

References

[1] 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