[1] Seoul National University [2] POSTECH [3] Pittburgh University

Generative Neural Fields by Mixtures of Neural Implicit Functions

Tackgeun You[2] Mijeong Kim[1] Jungtaek Kim[3] Bohyung Han[1]

NEURAL INFORMATION PROCESSING SYSTEMS

Introduction

Motivation

- Model averaging on parameter space improves the performance on tasks without additional Inference costs.
- Exploiting model averaging on implicit neural representation as **a compact conditioning mechanism** for generative neural fields.

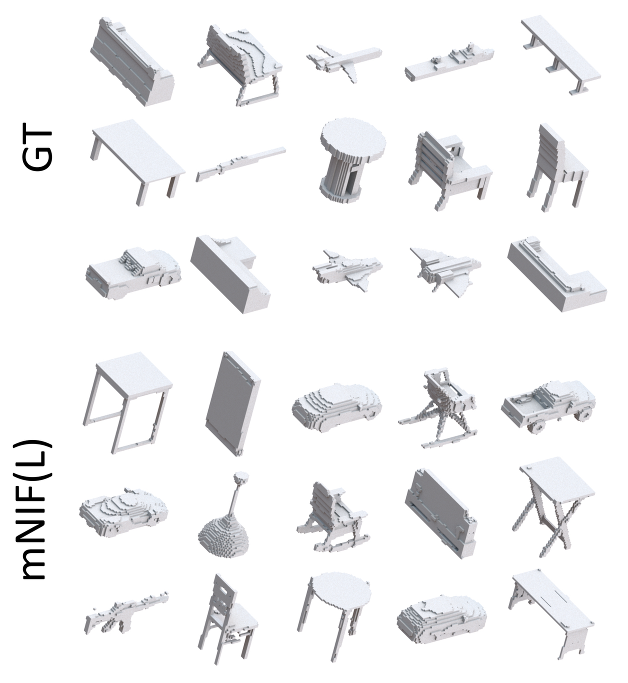
Our Contributions

- We introduce a generative neural field based on mixtures of neural implicit functions, a conditioning mechanism by a linear combination of neural implicit functions.
- Our method allows us to **easily extend the model capacity** by increasing the number of implicit basis M **without affecting the structure of the final model for inference**. Such a property greatly improves the inference efficiency in terms of computation and memory.
- The proposed approach effectively extracts instance-specific information through conducting meta-learning or applying auto-decoding procedures.

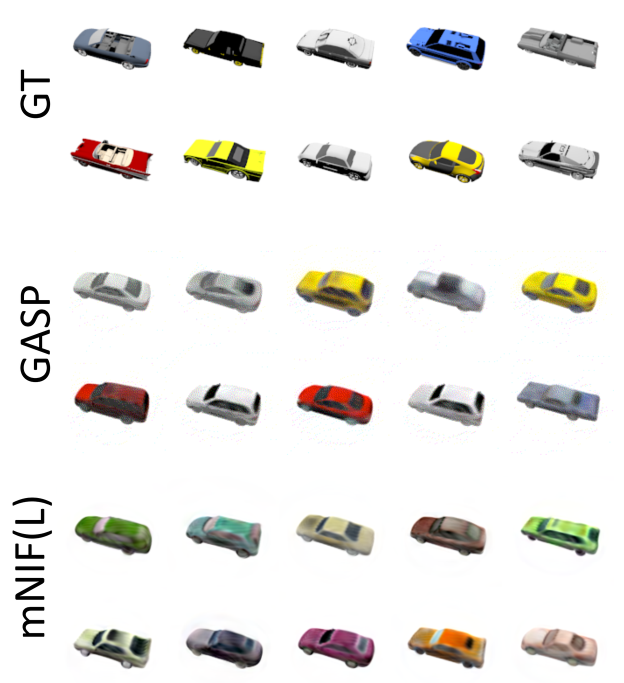
Generation and Interpolation

Image Generation

Voxel Generation



Radiance Field Generation

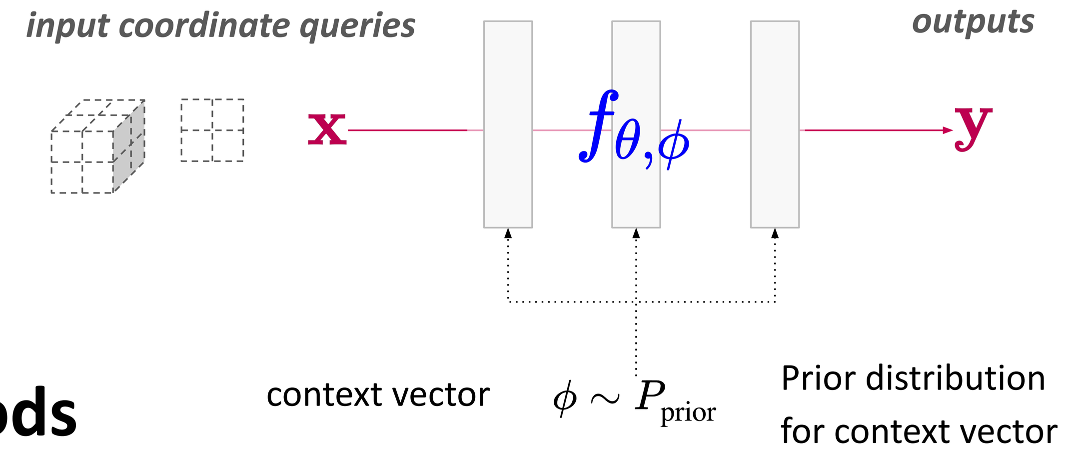


Interpolation results on samples from image, voxel and radiance field

Generative Neural Fields

A Generative Mechanism for Neural Fields

- **Architecture** conditioning mechanism for implicit neural representation
- **Algorithm** training and sampling strategy for generative neural fields



Comparison to the other methods

Model	Modality	Conditioning Mechanism	Training Conditioned Implicit Neural Representation	Sampling Context Vector
GASP (AISTATS 22)	Image, Voxel, Spherical map	HyperNetwork with block-wise weights	Generative Adversarial Network	Pre-defined random noise
GEM (NeurIPS 21)	Image, Voxel, Audio-video	HyperNetwork with low-rank decomposition	Auto-decoding	Non-parametric sampling
Functa (ICML 22)	Image, Voxel, NeRF scene, Spherical map	Bias modulation	Fast-context adaptation via meta learning (CAVIA)	MLP-based Diffusion model
HyperDiffusion (ICCV 23)	Voxel, Voxel video	No conditioning	Fitting each sample	Transformer-based weight estimator
Ours	Image, Voxel, NeRF scene	Mixtures of neural implicit functions	CAVIA and auto-decoding	Diffusion model

Experiments

Image, Voxel and Radiance Fields Reconstruction & Generation

Table 1. image generation performance on CelebA-HQ 64^2

Model	# param		Reconstruction		Generation				Inference Efficiency		
	Learnable	Inference	PSNR	rFID	FID	Precision	Recall	F1	GFLOPS	fps	Memory
Functa	3.3 M	2,629.6 K	26.6	28.4	40.4	0.577	0.397	0.470	8.602	332.9	144.1
GEM	99.0 M	921.3 K	-	-	30.4	0.642	0.502	0.563	3.299	559.6	70.3
GASP	34.2 M	83.1 K	-	-	13.5	0.836	0.312	0.454	0.305	1949.3	16.4
DPF	62.4 M	-	-	-	13.2	0.866	0.347	0.495	-	-	-
mNIF (S)	4.6 M	17.2 K	31.5	10.9	21.0	0.787	0.324	0.459	0.069	2958.6	10.2
mNIF (L)	33.4 M	83.3 K	34.5	5.8	13.2	0.902	0.679	0.679	0.340	891.3	24.4

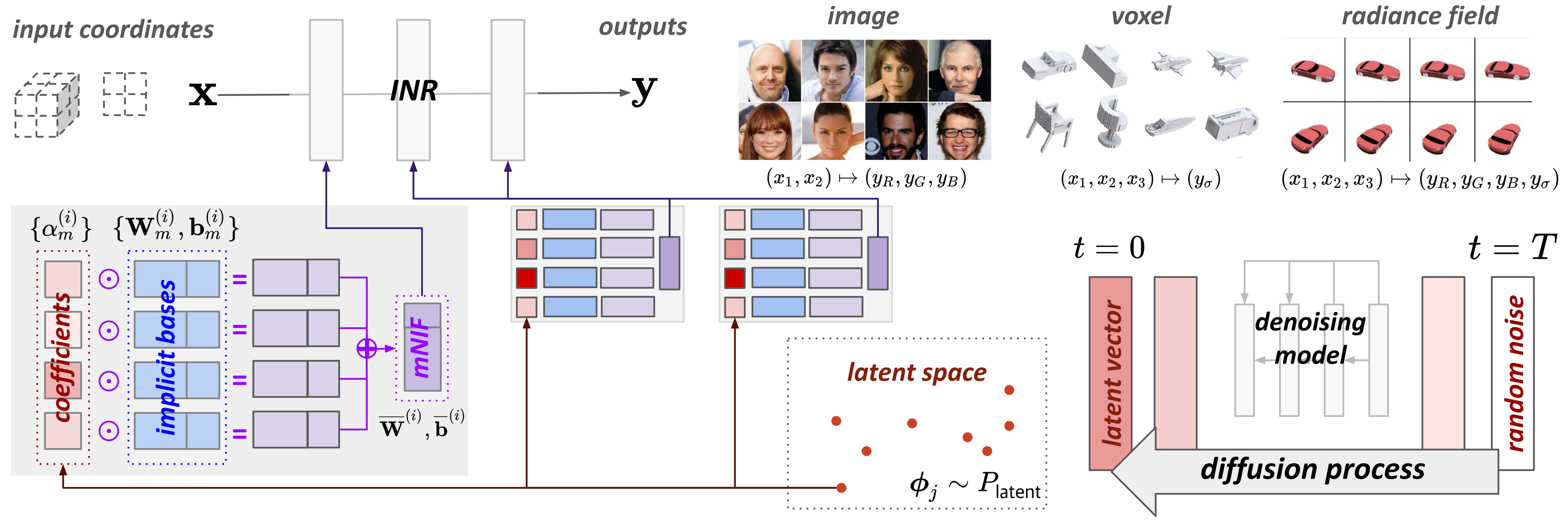
Table 2. voxel generation performance on ShapeNet 64^3

Model	# param		Reconstruction		Generation		Inference Efficiency		
	Learnable	Inference	MSE	PSNR	Coverage	MMD	GFLOPS	fps	Memory
GASP	34.2 M	83.1 K	0.0296	16.5	0.341	0.0021	8.7	180.9	763.1
GEM	99.0 M	921.3 K	0.0153	21.3	0.409	0.0014	207.0	16.7	4010.0
DPF	62.4 M	-	-	-	0.419	0.0016	-	-	-
mNIF (S)	4.6 M	17.2 K	0.0161	20.9	0.430	0.0014	4.4	191.5	642.1
mNIF (L)	33.4 M	83.3 K	0.0166	21.4	0.437	0.0013	21.6	69.6	1513.3

Table 3. radiance field scene generation performance on SRN Cars

Model	# param		Reconstruction	Generation	Inference Efficiency		
	Learnable	Inference	MSE	FID	GFLOPS	fps	Memory
Functa	3.9 M	3,418.6 K	24.2	80.3	1.789	2.0	28.0
mNIF (S)	4.6 M	17.2 K	25.9	79.5	0.009	97.7	1.3

Mixtures of Neural Implicit Functions



Weighted averaging of neural implicit functions *g* with latent mixture coefficients

$$\mathbf{h}^{(i+1)} = \sin\left(w_0\left(\textstyle\sum_{m=1}^{M} \alpha_m^{(i)} \mathbf{g}_m(\mathbf{h}^{(i)})\right)\right)$$
$$= \sin\left(w_0\left(\left(\textstyle\sum_{m=1}^{M} \alpha_m^{(i)} \mathbf{W}_m^{(i)}\right)\mathbf{h}^{(i)} + \textstyle\sum_{m=1}^{M} \alpha_m^{(i)} \mathbf{b}_m^{(i)}\right)\right)$$
$$= \sin\left(w_0\left(\overline{\mathbf{W}}^{(i)} \mathbf{h}^{(i)} + \overline{\mathbf{b}}^{(i)}\right)\right)$$
$$\overline{\mathbf{W}}^{(i)} =: \textstyle\sum_{m=1}^{M} \alpha_m^{(i)} \mathbf{W}_m^{(i)} \qquad \overline{\mathbf{b}}^{(i)} =: \textstyle\sum_{m=1}^{M} \alpha_m^{(i)} \mathbf{b}_m^{(i)}$$

$$\left[\alpha_1^{(0)}, \alpha_2^{(0)}, \cdots, \alpha_{M-1}^{(L+1)}, \alpha_M^{(L+1)}\right] = \mathbf{T}\boldsymbol{\phi},$$

- ***Efficiency*** *is mainly proportional to* *the number of queries* (modality-dependent) × *an inference cost per queries* (our target).

Ablation Study on Mixtures of Neural Implicit Functions

Table 4. Ablation study on mixture coefficient configuration of mNIF on CelebA-HQ 642

Exp	Mixture	(L, W, M, H)	# Params		Reconstruction (Train)				Reconstruction (Test)	
			Learnable	Inference	PSNR	rFID	rPrecision	rRecall	PSNR	rFID
(a)	Shared	(2,64,64,64)	557.2 K	8.7 K	22.20	50.70	0.497	0.003	22.11	55.87
	Layer-specific	(2,64,64,256)	557.2 K		24.45	38.23	0.461	0.013	24.35	42.65
	Latent	(2,64,64,256)	623.0 K		25.27	31.67	0.534	0.040	25.09	36.85
(b)	Latent	(2,64,16,256)	155.8 K	8.7 K	22.02	53.01	0.433	0.001	21.84	57.23
		(2,64,64,256)	623.0 K		25.27	31.67	0.534	0.040	25.09	36.85
		(2,64,256,256)	2.5 M		26.64	24.62	0.640	0.134	25.74	32.17
		(2,64,1024,256)	10.0 M		26.84	23.71	0.642	0.155	25.85	32.14
(c)	Latent	(5,128,256,256)	21.8 M	83.3 K	31.17	10.65	0.890	0.750	25.35	31.58
		(5,128.256,512)	22.3 M		32.11	8.96	0.918	0.845	27.92	24.79
		(5,128,256,1024)	23.2 M		32.71	8.09	0.935	0.893	29.45	22.05

Table 5. Study on context adaptation strategy

Strategy	Second-order	Ninner	PSNR	rFID
Auto-decoding	-	-	24.68	34.31
Meta-learning (CAVIA)	First-order	3	22.68	57.27
	Use	3	26.64	24.64
	Use	5	26.77	24.37

Table 6. Study on longer training of mNIF

Epoch	Reconstruction		Generation
	PSNR	rFID	FID
400	32.11	8.96	14.96
800	33.10	7.31	14.30
1200	33.61	6.91	15.25

[Hypernetwork] HyperNetworks. *ICLR 2017*
[CAVIA] Fast Context Adaptation via Meta-Learning. *ICML 2019*
[GEM] Learning Signal-Agnostic Manifolds of Neural Fields. *NeurIPS 2021*
[Functa] From data to functa: Your data point is a function and you can treat it like one. *ICML 2022*
[GASP] Generative Models as Distributions of Functions. *AISTATS 2022*
[DPF] Diffusion Probabilistic Fields. *ICLR 2023*

Github Repository
http://github.com/tackgeun/mNIF