DiffAssemble
A Unified **Graph-Diffusion** Model for **2D** and **3D Reassembly**

Gianluca Scarpellini[*,1,2], Stefano Fiorini[*,1], Francesco Giuliari[*,1,2], Pietro Morerio[1], Alessio Del Bue[1]

[1]Istituto Italiano di Tecnologia, [2]Università di Genova *Equal Contribution

IIT PAVIS
ISTITUTO ITALIANO DI TECNOLOGIA
Università di Genova

CVPR
JUNE 17-21, 2024
SEATTLE, WA

TL;DR;

Given a set of elements at random positions, can we recover their original pose for reassembly?

Use **Denoising Probabilistic Diffusion Models**, treating the initial poses as random initialization, to iteratively recover the original poses via denoising.

DiffAssemble

TASK

2D	3D	Real-world application

CelebA

WikiArt

everyday

Breaking-bad
(Adapted from [5])

Protein design

Fresco reconstruction

METHOD

KEY IDEAS

- Same approach for 2D and 3D reassembly
- Graph-based Diffusion Model
- Predicts element wise rotation and translation

ARCHITECTURE

Input:
Features are extracted with rotation-equivariant feature encoder (2d or 3d)

Construct graph:
each piece is a node with its features and noisy position and orientation

Output:
DiffAssemble denoises each piece and produces a slightly less noisy pose at output

Pieces   Features h

Equivariant 2D / 3D Encoder

$\oplus \begin{bmatrix} s_t \\ r_t \end{bmatrix}$

Attention based GNN

$\begin{bmatrix} s_{t-1} \\ r_{t-1} \end{bmatrix}$

$t = T$ $t = t$ $t = t - 1$ $t = 0$

Reverse Denoising Process (Inference)

Forward Diffusion Process (Training)

Inference:
At infernece, we initialize the input graph with random poses and iteratively denoise each piece for T steps

RESULTS

2D

METHOD		PuzzleCelebA			PuzzleWikiArts		
	6x6	8x8	10x10	12x12	6x6	8x8	10x10	12x12
Optimization Based	Gallagher [1]	80.21	55.18	71.19	69.81	71.88	61.63	54.15	44.68
	Yu *et al.* [2]	98.63	94.65	98.33	93.33	94.62	92.95	90.99	89.88
	Huroyan *et al.* [3]	98.47	97.45	98.65	97.08	92.69	91.37	89.74	88.28
Learning Based	**DiffAssemble** - No Diff.	99.43	79.84	99.05	91.28	73.07	54.70	22.68	18.27
	DiffAssemble - No Equiv.	96.12	71.62	91.98	64.15	25.31	14.63	8.19	4.96
	DiffAssemble	99.51	87.66	99.30	97.76	90.65	72.79	63.33	53.08

Up to 900 pieces

Robust to missing pieces

3D

METHOD	RMSE (R)↓ degree	RMSE (T)↓ ×10⁻²	PA↑ %
DGL [5]	81.4	14.9	25.4
SE(3)-Equiv [4]	77.9	16.7	8.1
DiffAssemble - No Diffusion Process	83.6	17.1	3.1
DiffAssemble - No Equivariant Enc.	81.7	17.0	18.3
DiffAssemble	73.3	14.8	27.5

- Diffusion process improves overall results.
- **Equivariant features** leads to big improvements.
- Joint diffusion on rotation and translation.

Input   SE(3)-Equivariant   Ours   Ground Truth

REFERENCES

[1] Andrew C Gallagher. Jigsaw puzzles with pieces of unknown orientation. CVPR 2012
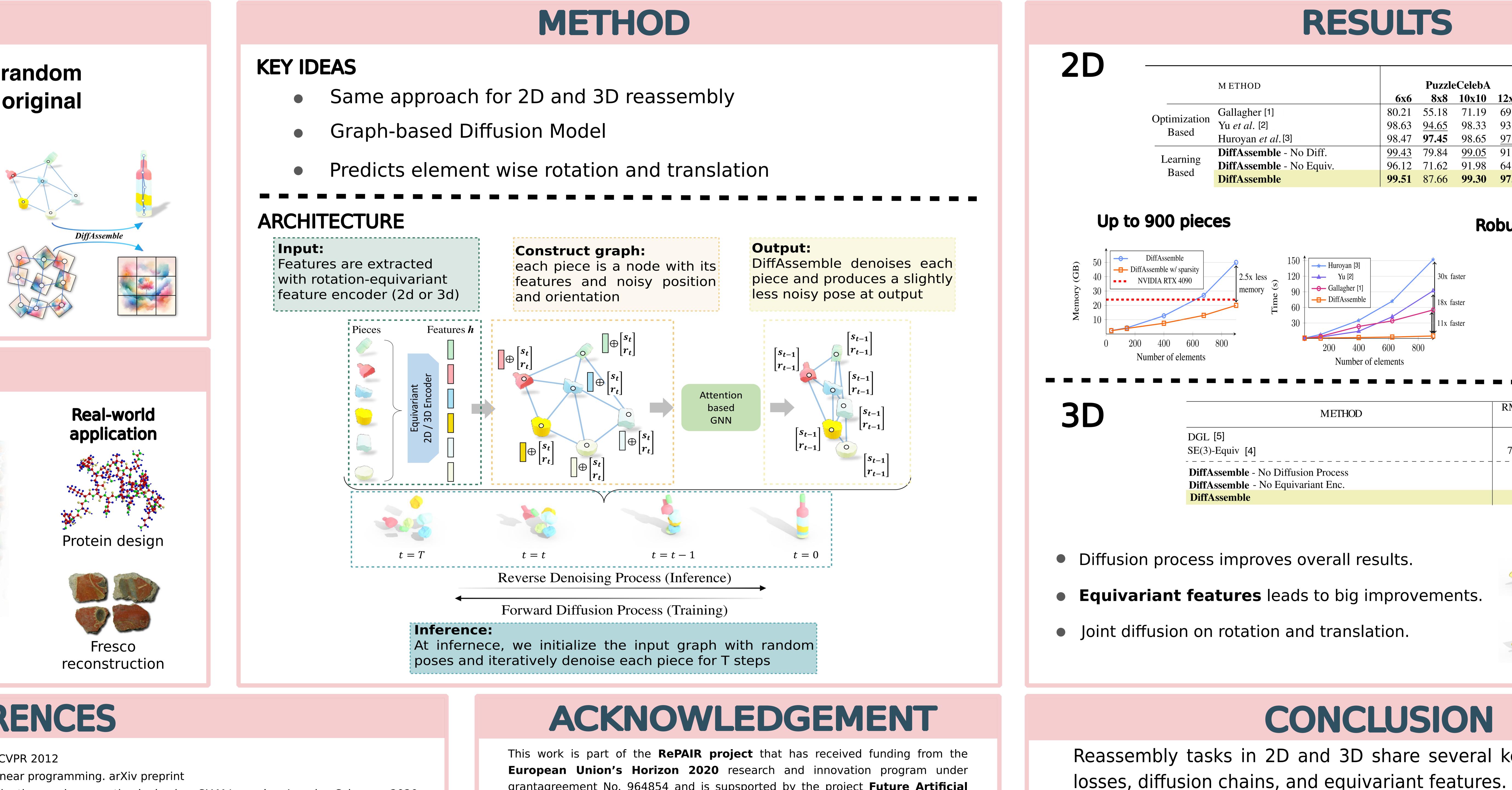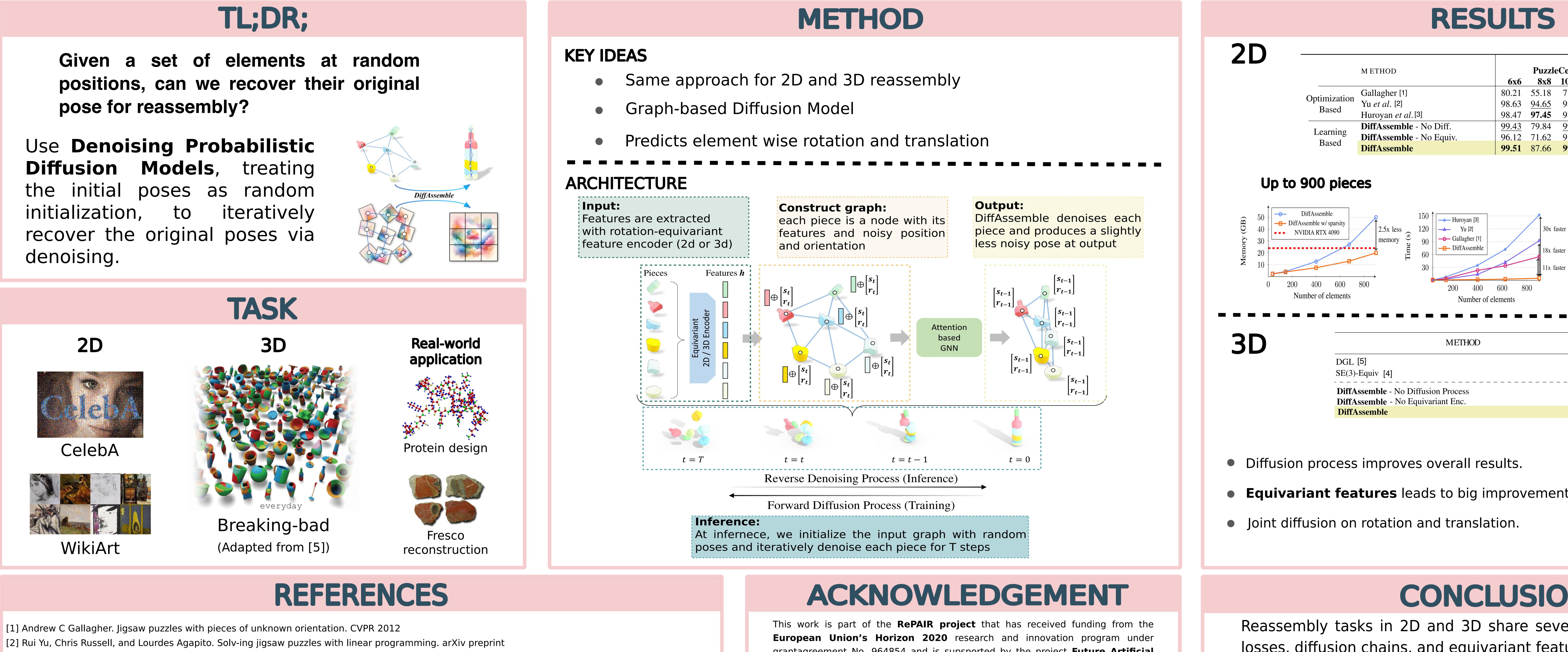[2] Rui Yu, Chris Russell, and Lourdes Agapito. Solv-ing jigsaw puzzles with linear programming. arXiv preprint
[3] Vahan Huroyan, Gilad Lerman, and Hau-Tieng Wu.Solving jigsaw puzzles by the graph connection lapla-cian. SIAM Journal on Imaging Sciences, 2020
[4] Ruihai Wu, Chenrui Tie, Yushi Du, Yan Zhao, and Hao Dong. Leveraging se(3) equivariance for learn-ing 3d geometric shape assembly. ICCV 2023
[5] Silvia Sell´an, Yun-Chun Chen, Ziyi Wu, Animesh Garg,and Alec Jacobson. Breaking bad: A dataset for geo-metric fracture and reassembly. NeurIPS 2022

ACKNOWLEDGEMENT

CONCLUSION

Reassembly tasks in 2D and 3D share several key properties, e.g., losses, diffusion chains, and equivariant features.
DiffAssemble reaches SOTA performance in most 2D and 3D scenarios, revealing a common ground between these tasks.