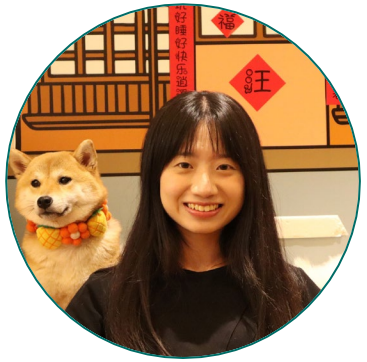
Feat2GS

Probing Visual Foundation Models with Gaussian Splatting


Yue Chen[1]

Xingyu Chen[1]

Anpei Chen[1,3]

Gerard Pons-Moll[3,4]

Yuliang Xiu[1,2]

[1]Westlake University
[2]Max Planck Institute for Intelligent Systems
[3]University of Tübingen, Tübingen AI Center
[4]Max Planck Institute for Informatics

WESTLAKE UNIVERSITY

Tübingen AI Center
tuebingen.ai

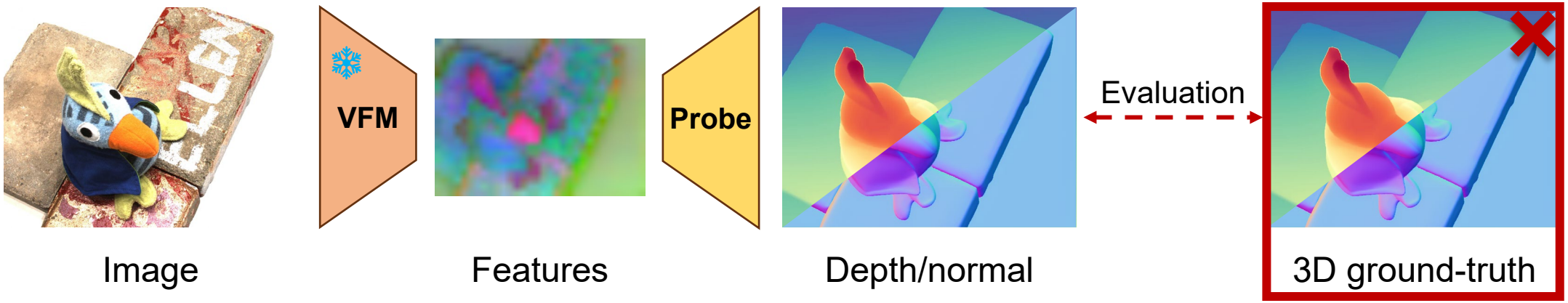
EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Visual Foundation Models

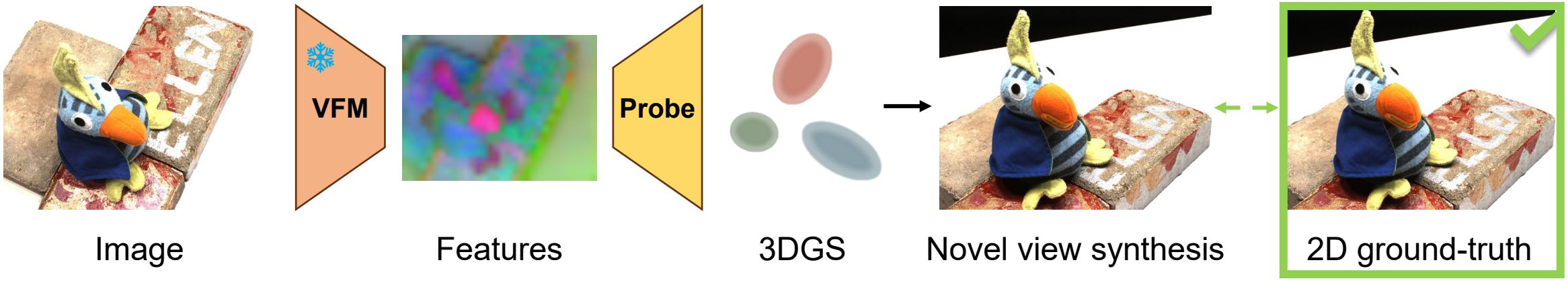
How well do they understand the **3D world**?

We need 3D probing.

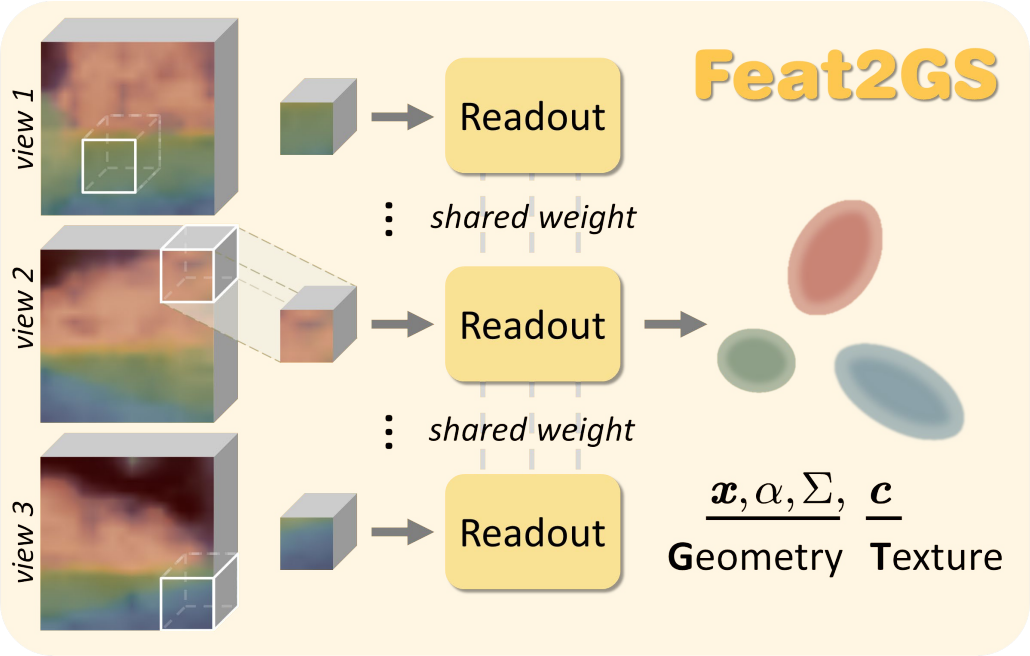
Previous 3D Probing



New Benchmark



Pipeline



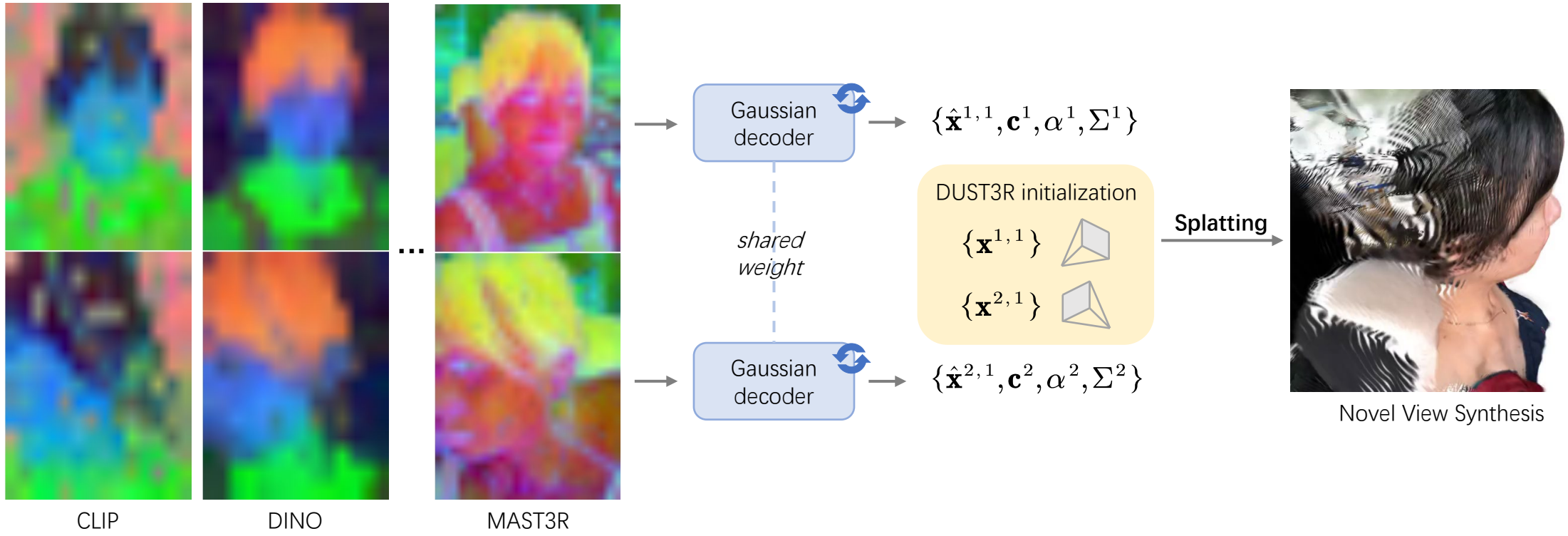
Splatting

Novel View Synthesis

Probing Schemesc

Probing	-Geometry	-Texture	-All
Feature- Readout	$\boldsymbol{x}, \alpha, \Sigma$	$\boldsymbol{c}$	$\boldsymbol{x}, \alpha, \Sigma$, $\boldsymbol{c}$
Free-Optimize	$\boldsymbol{c}$	$\boldsymbol{x}, \alpha, \Sigma$	/

Framework

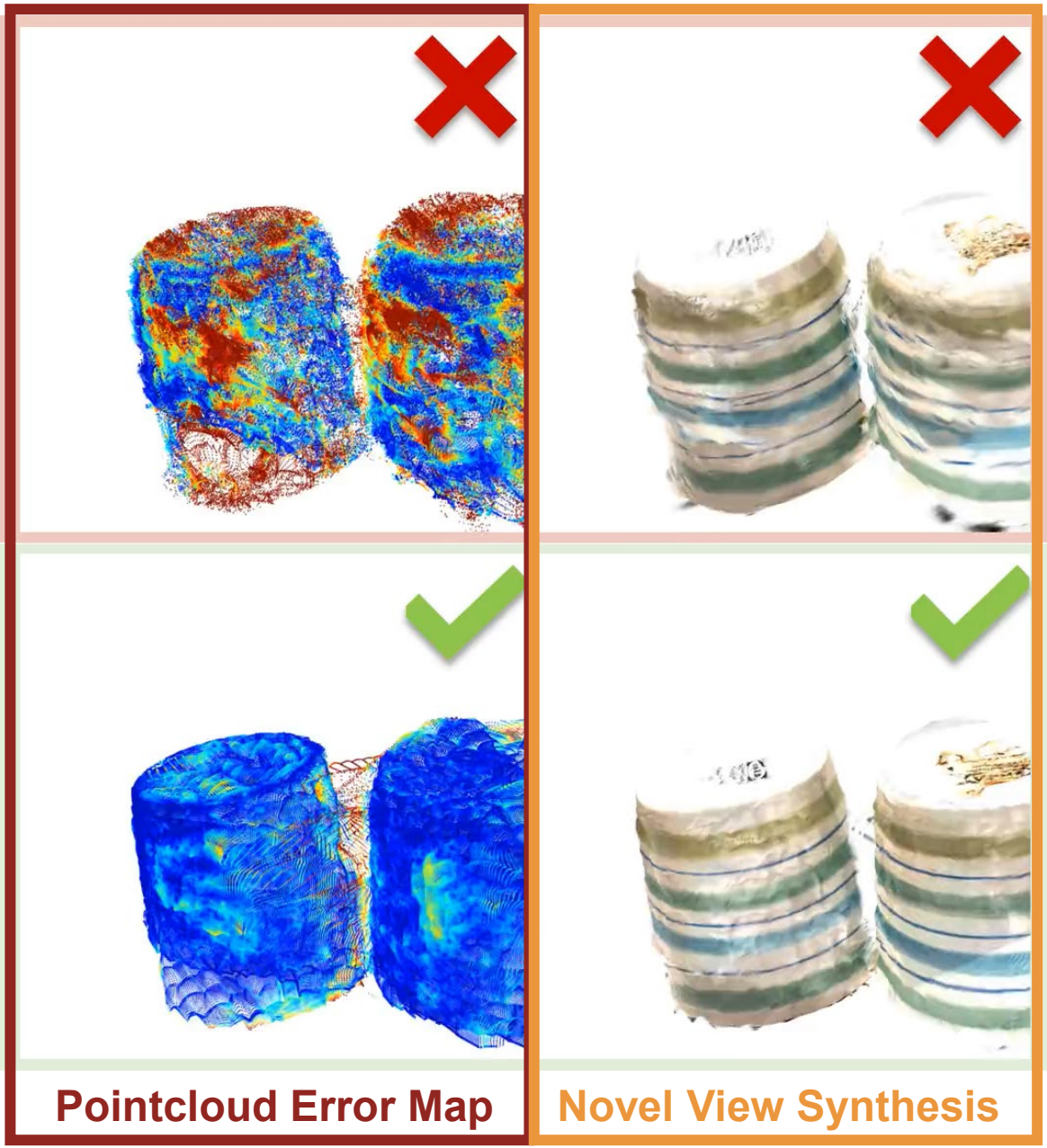

CLIP
DINO
MAST3R
Gaussian decoder
shared weight
Gaussian decoder
$\{\hat{\mathbf{x}}^{1,1}, \mathbf{c}^1, \alpha^1, \Sigma^1\}$
DUST3R initialization
$\{\mathbf{x}^{1,1}\}$
$\{\mathbf{x}^{2,1}\}$
$\{\hat{\mathbf{x}}^{2,1}, \mathbf{c}^2, \alpha^2, \Sigma^2\}$
Splatting
Novel View Synthesis

Our Findings: 3D Metrics and 2D Metrics are well-aligned.

DINOv2

vs.

DUSt3R



Pointcloud Error Map

Novel View Synthesis

Accuracy↓: -2.448
Completeness ↓: -.277
Distance↓: -6.163

PSNR↑: +1.02
SSIM↑: +.0233
LPIPS↓: -.0301

Geometry Probing

RADIO

MASt3R

Texture Probing

Findings:

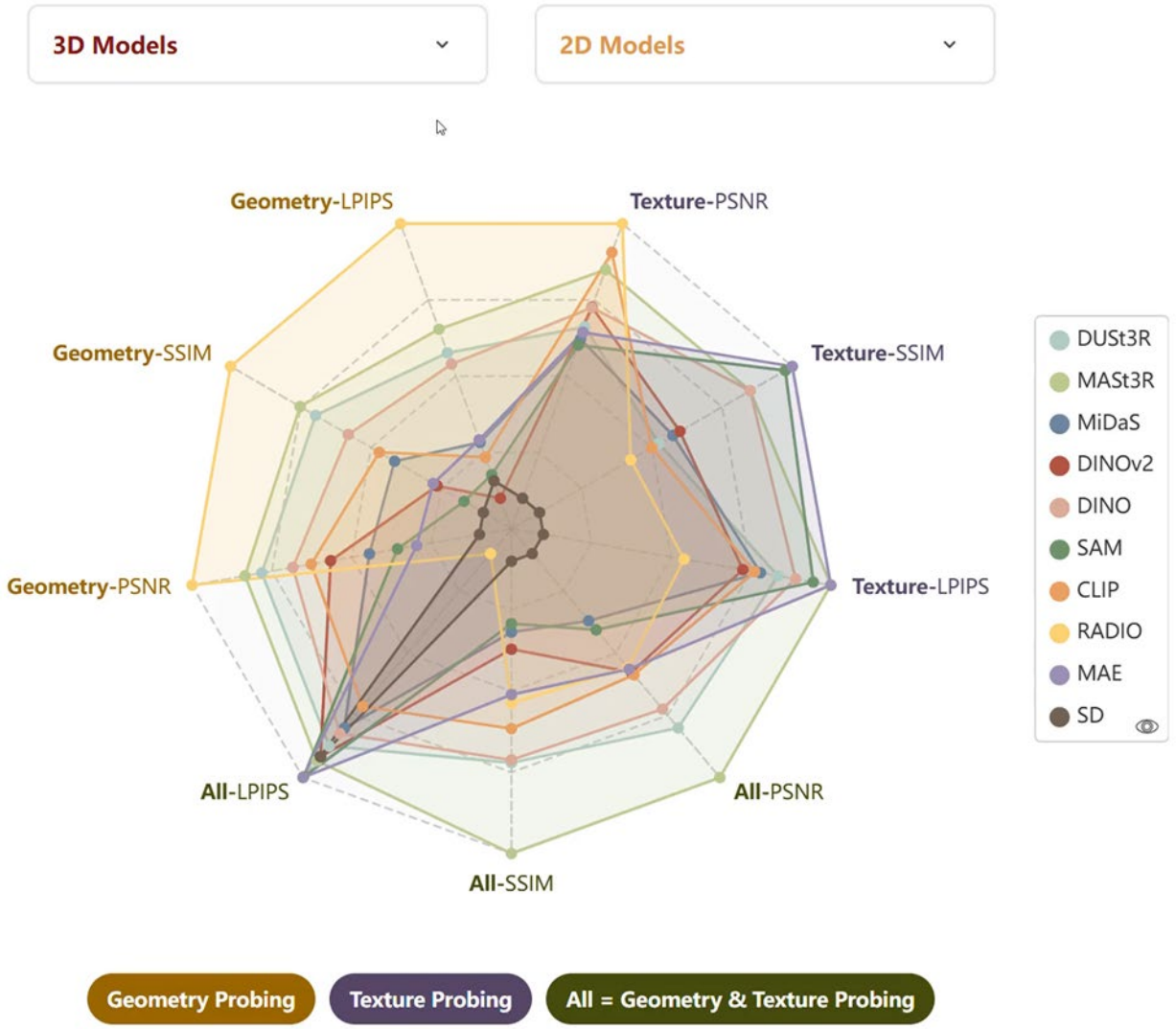
Foundation models capture geometry well, **but struggle with texture.**

Application

Input images

Novel View Synthesis / Normal Results

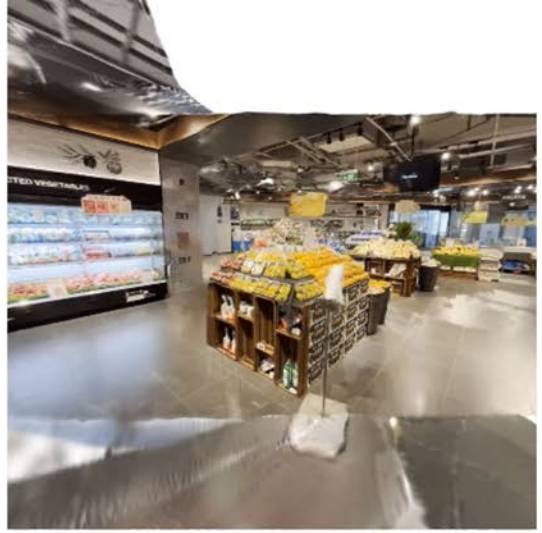
Interactive Probing Results



Welcome to our Website!

🔗 fanegg.github.io/Feat2GS

Play online and interact now!



Thanks!