



Depth Anywhere:

Enhancing 360 Monocular Depth Estimation via Perspective Distillation and Unlabeled Data Augmentation

Ning-Hsu (Albert) Wang¹

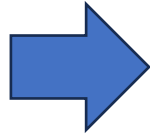
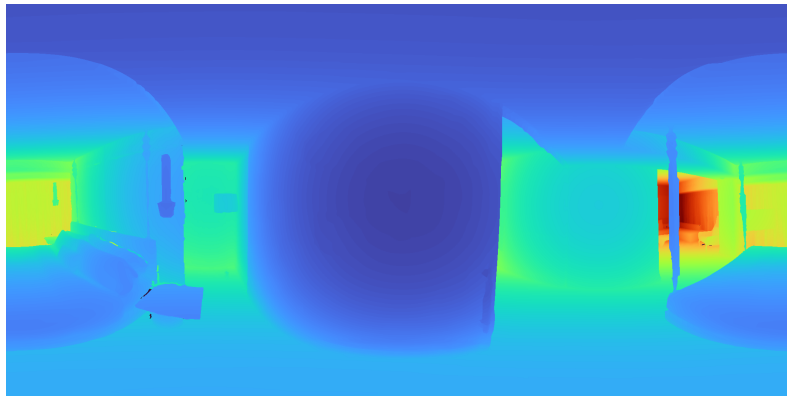
Yu-Lun Liu²

¹Independent Researcher.

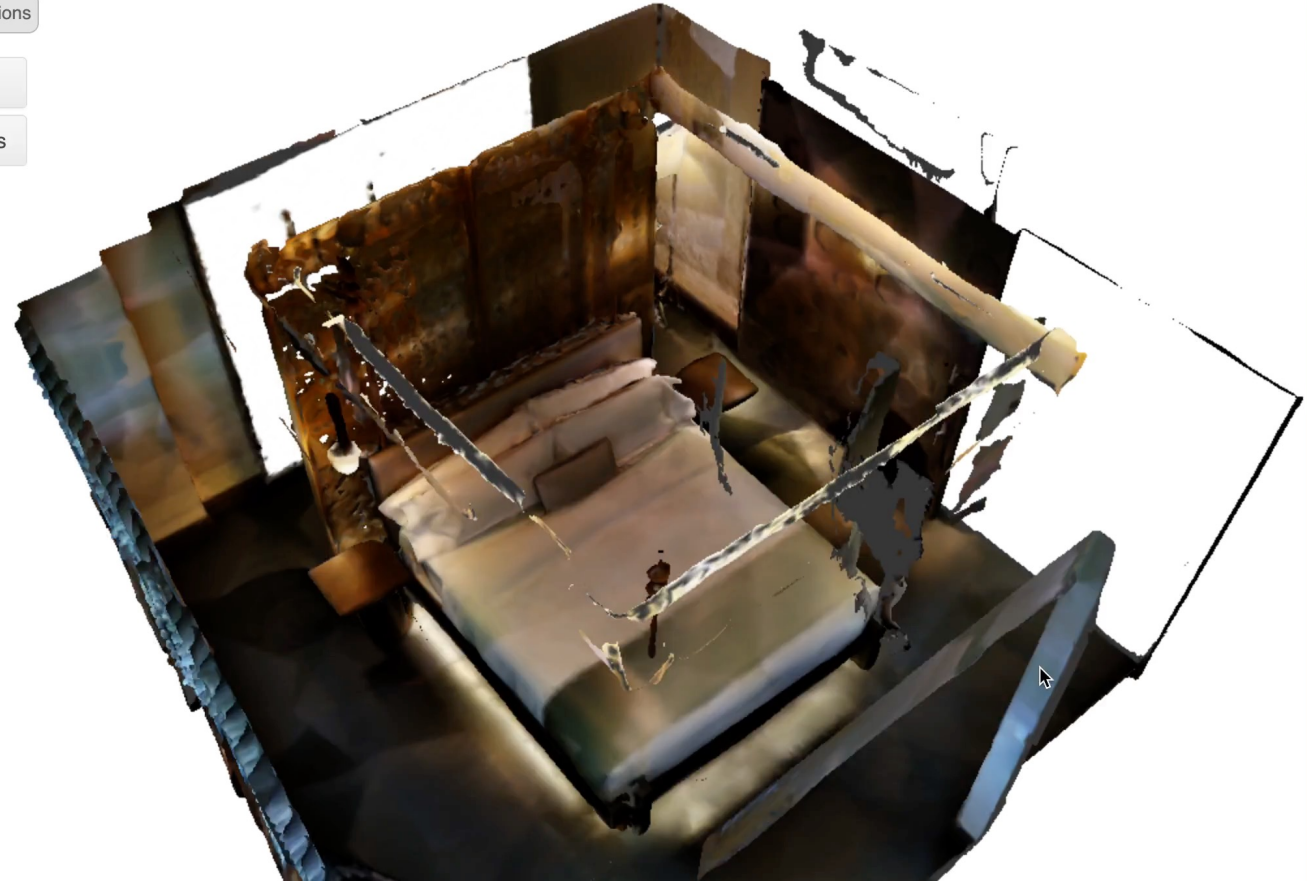
²National Yang Ming Chiao Tung University



360 Monocular Depth Estimation

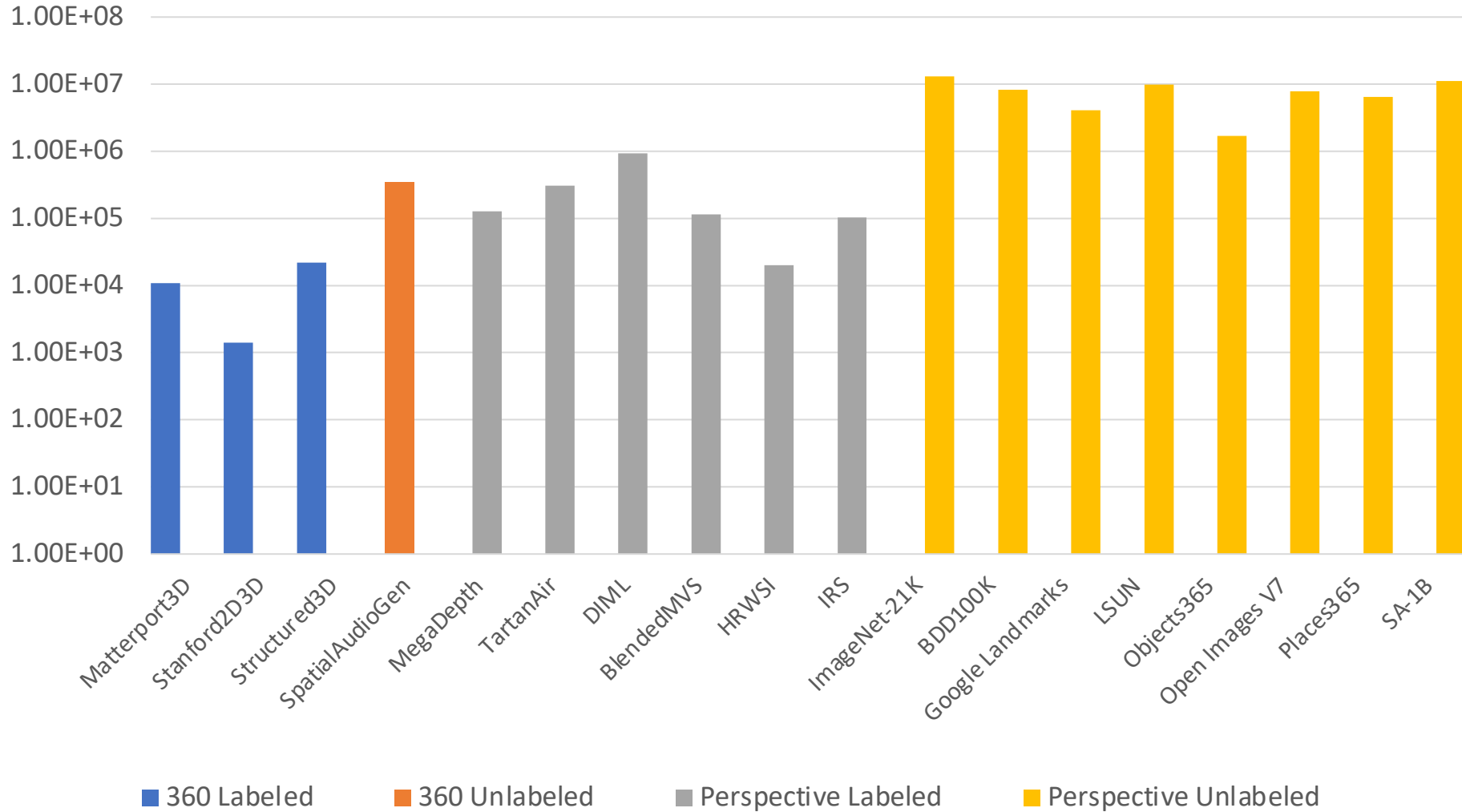


- Instructions
- Parts
- Images



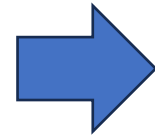
360 Monocular Depth Estimation

Dataset Statistic



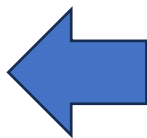
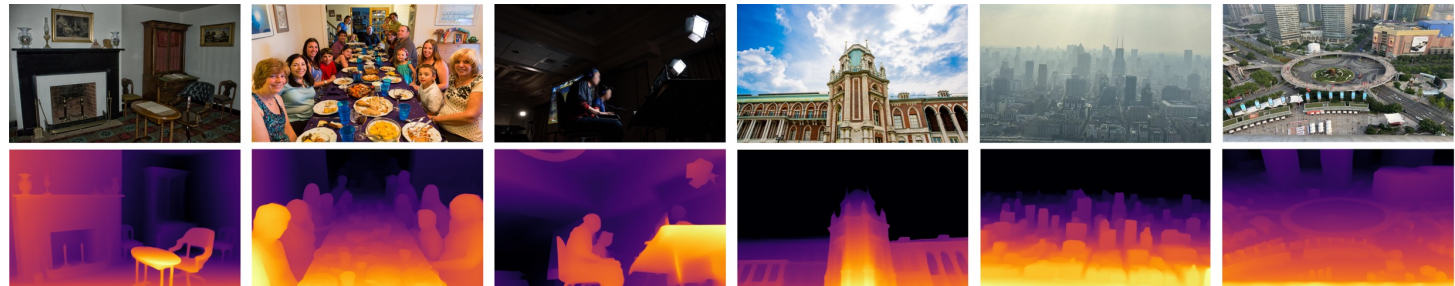
Cross Camera Model Knowledge Distillation

Marigold: Repurposing Diffusion-Based Image Generators for Monocular Depth Estimation



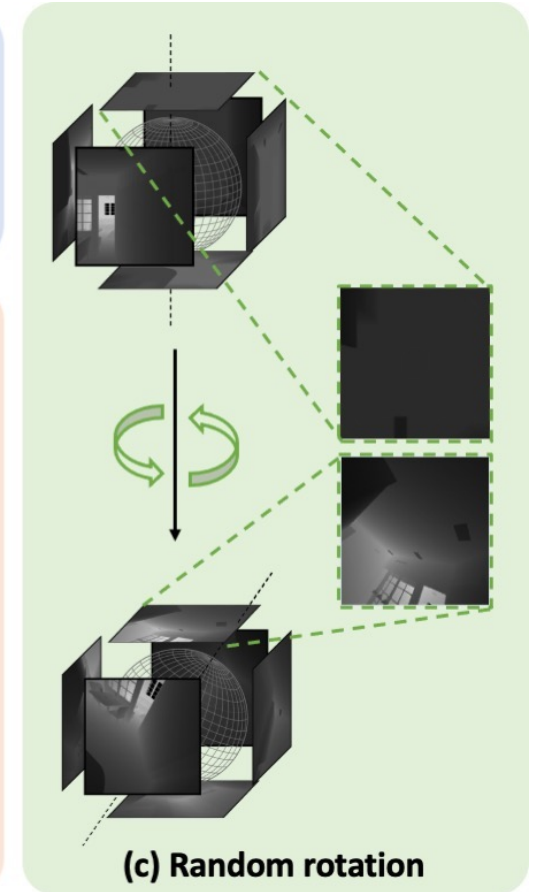
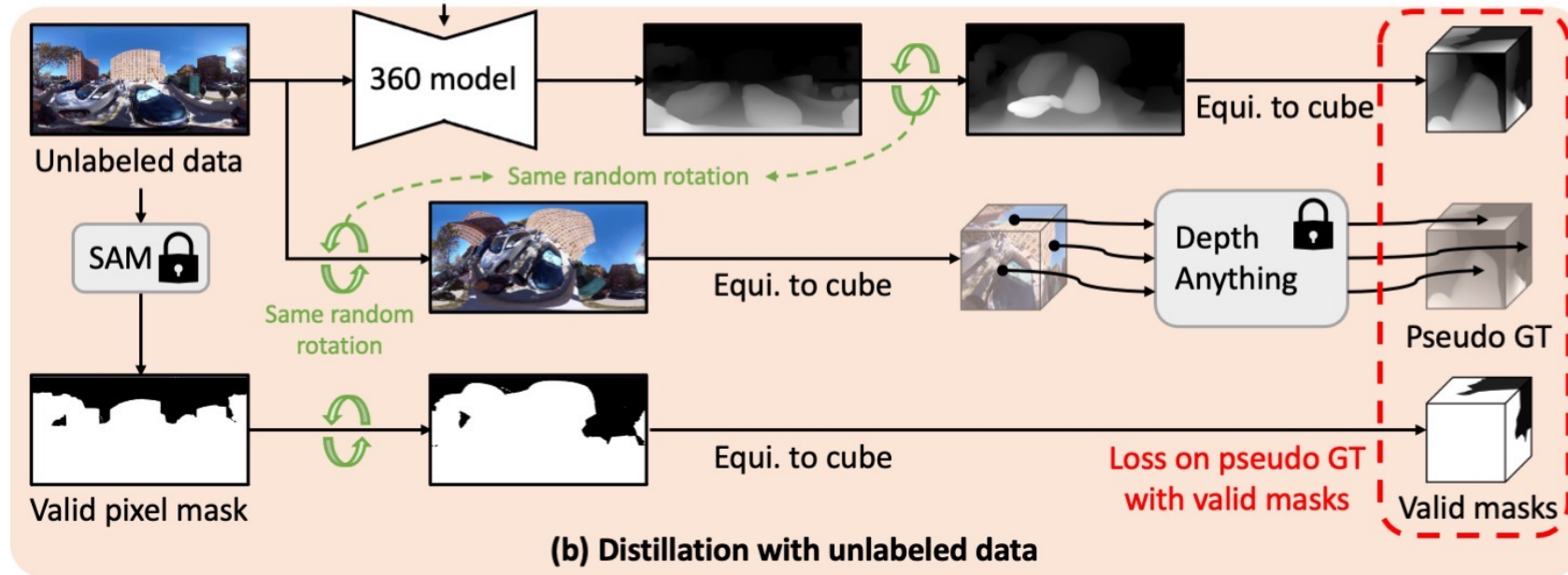
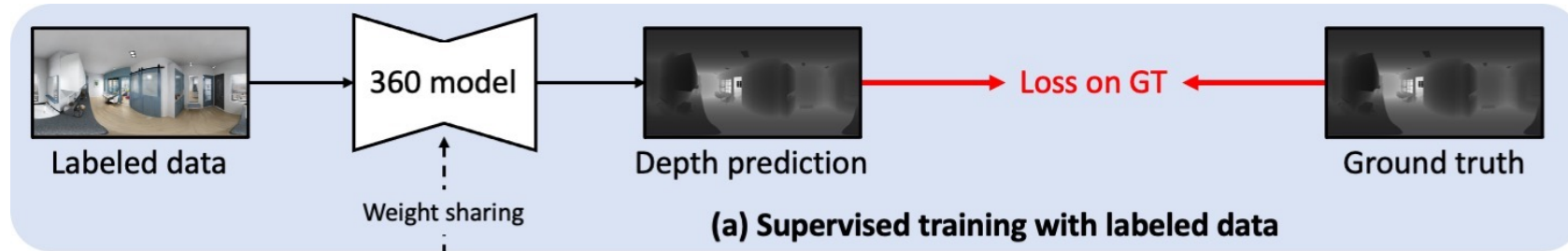
2 labeled datasets (74K images)

Depth Anything Unleashing the Power of Large-Scale Unlabeled Data

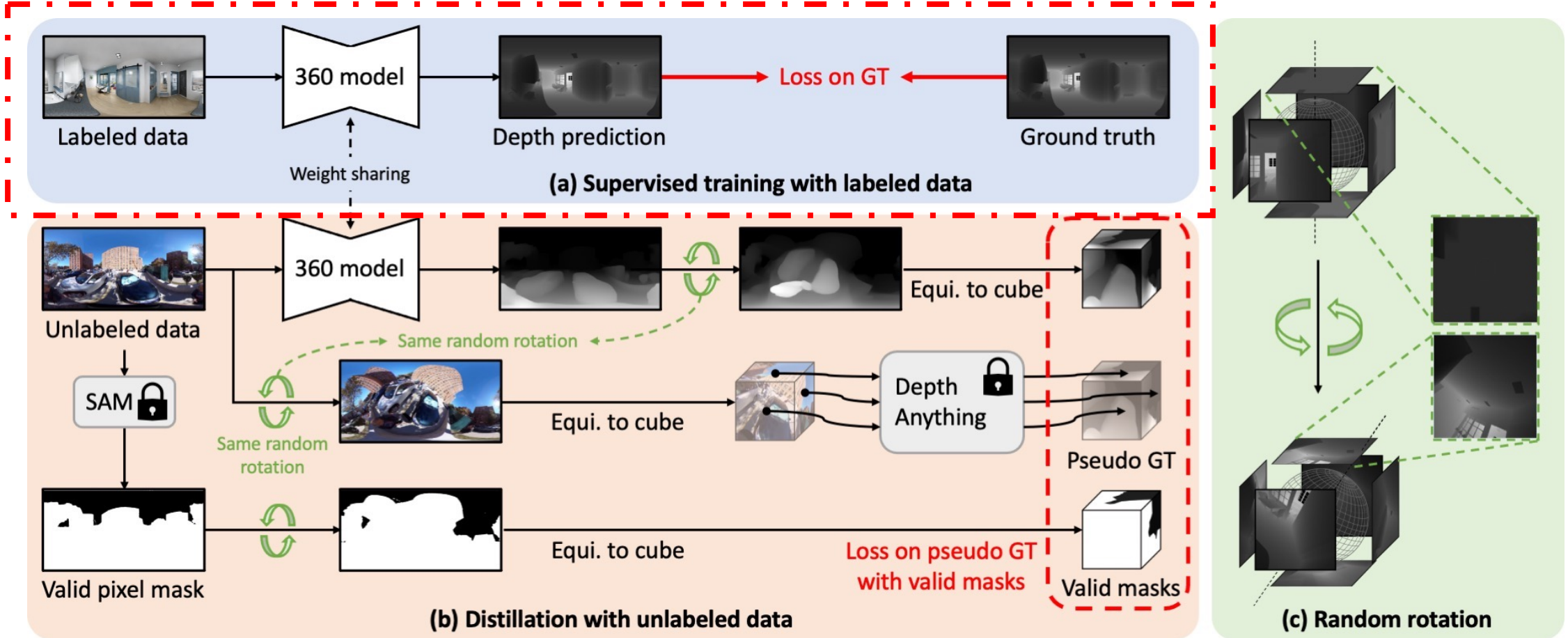


6 labeled datasets (1.5M images)
8 unlabeled datasets (62M+ images).

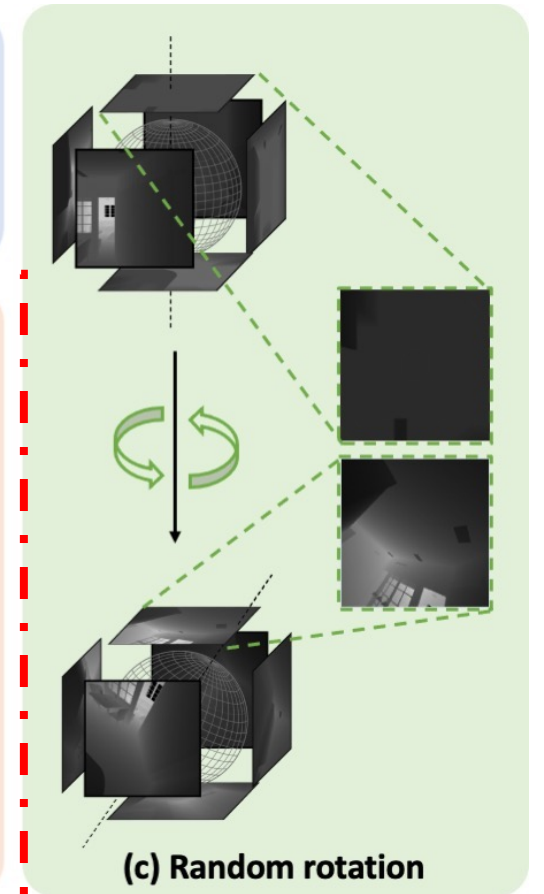
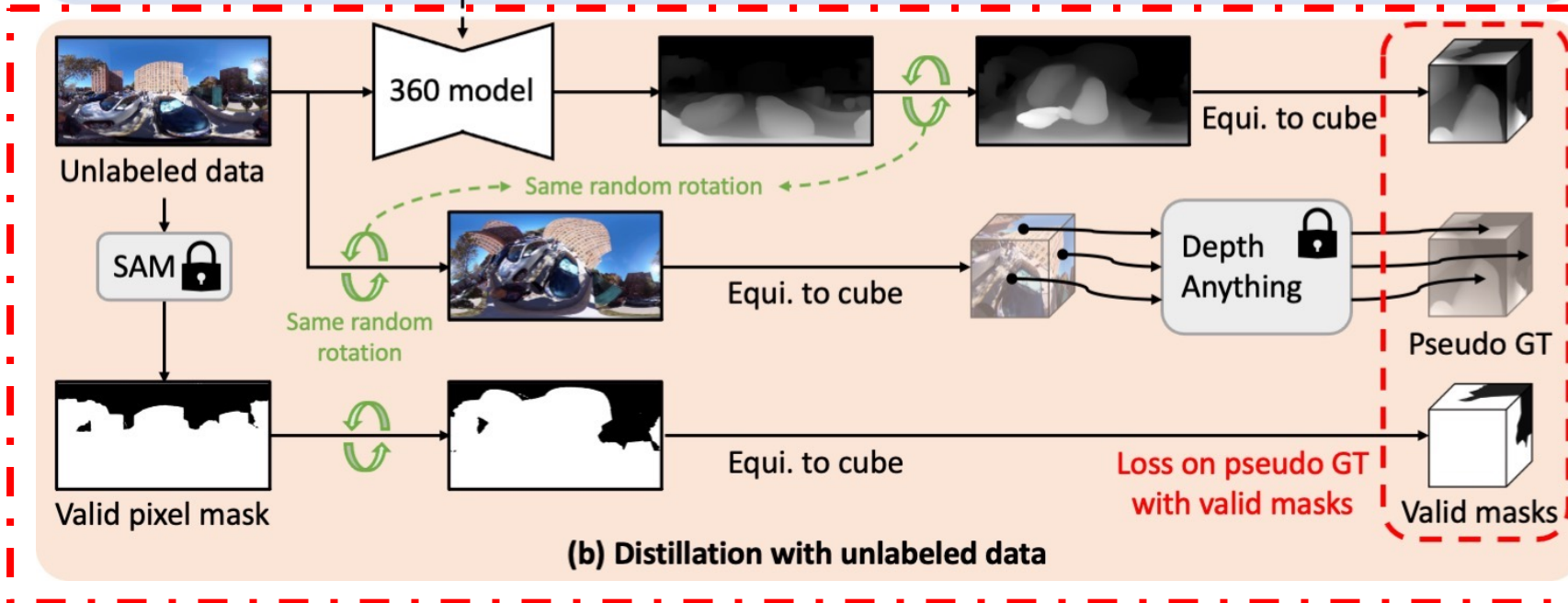
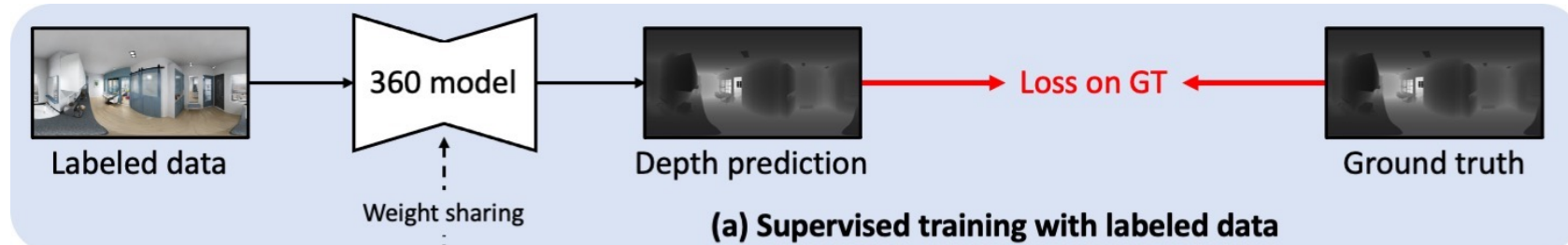
Cross Camera Model Knowledge Distillation



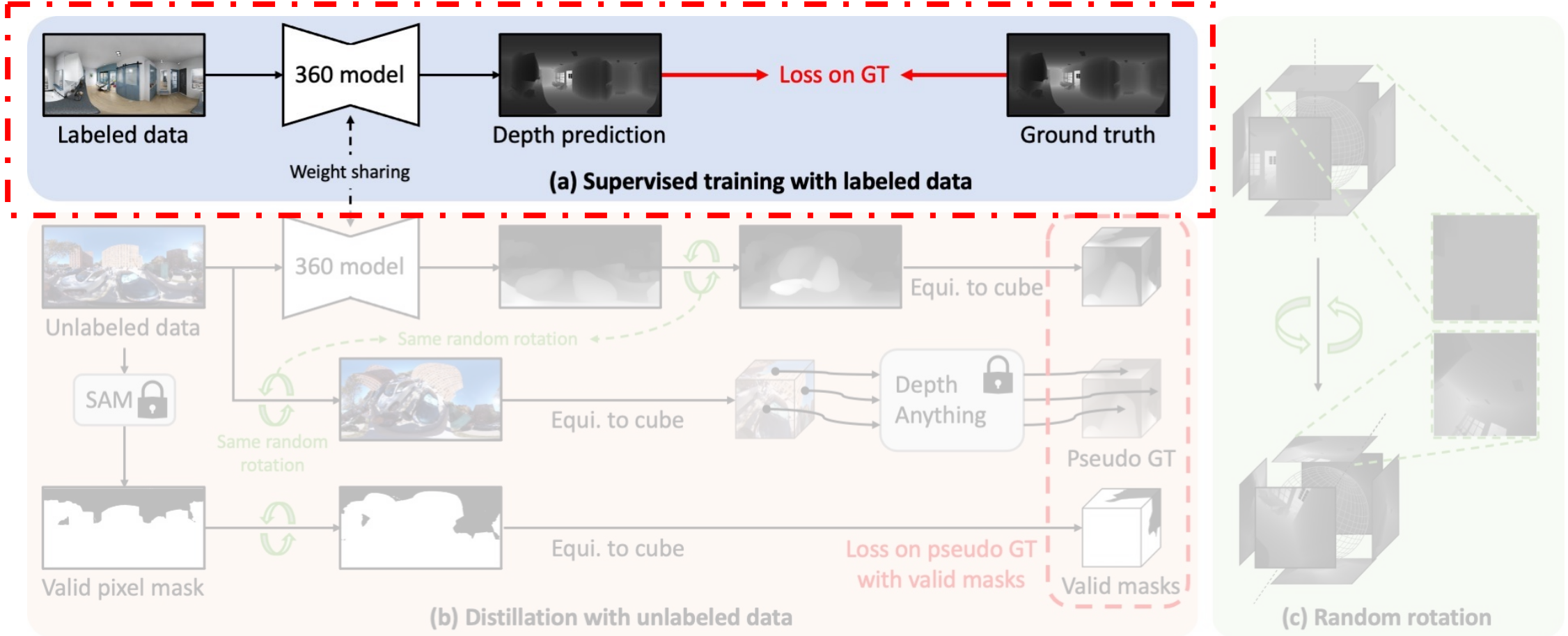
Cross Camera Model Knowledge Distillation



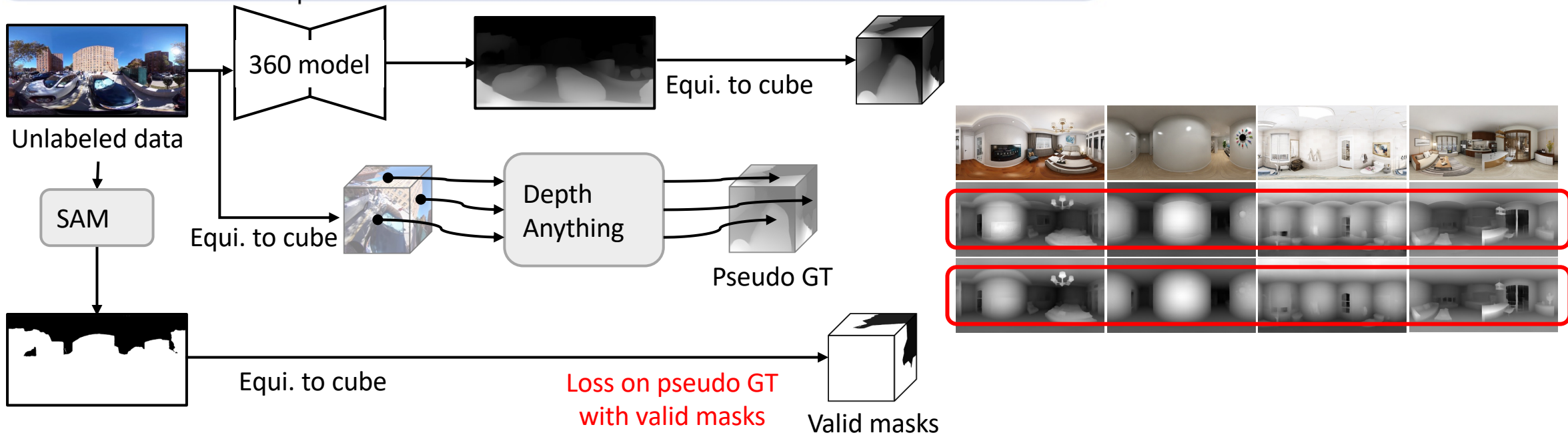
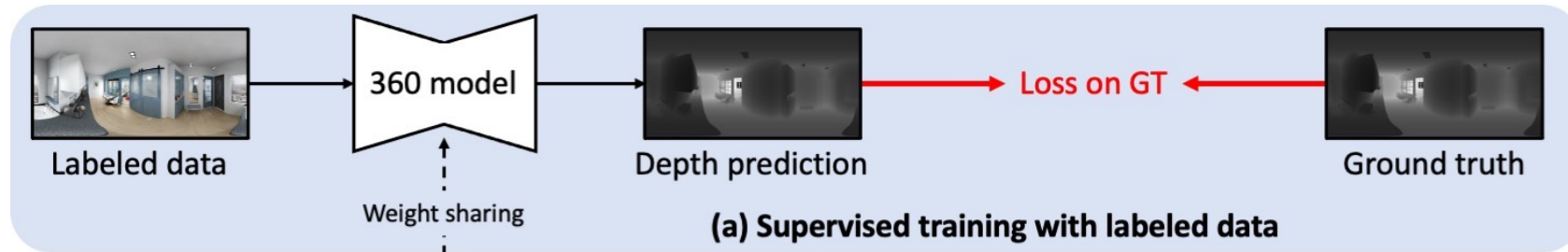
Cross Camera Model Knowledge Distillation



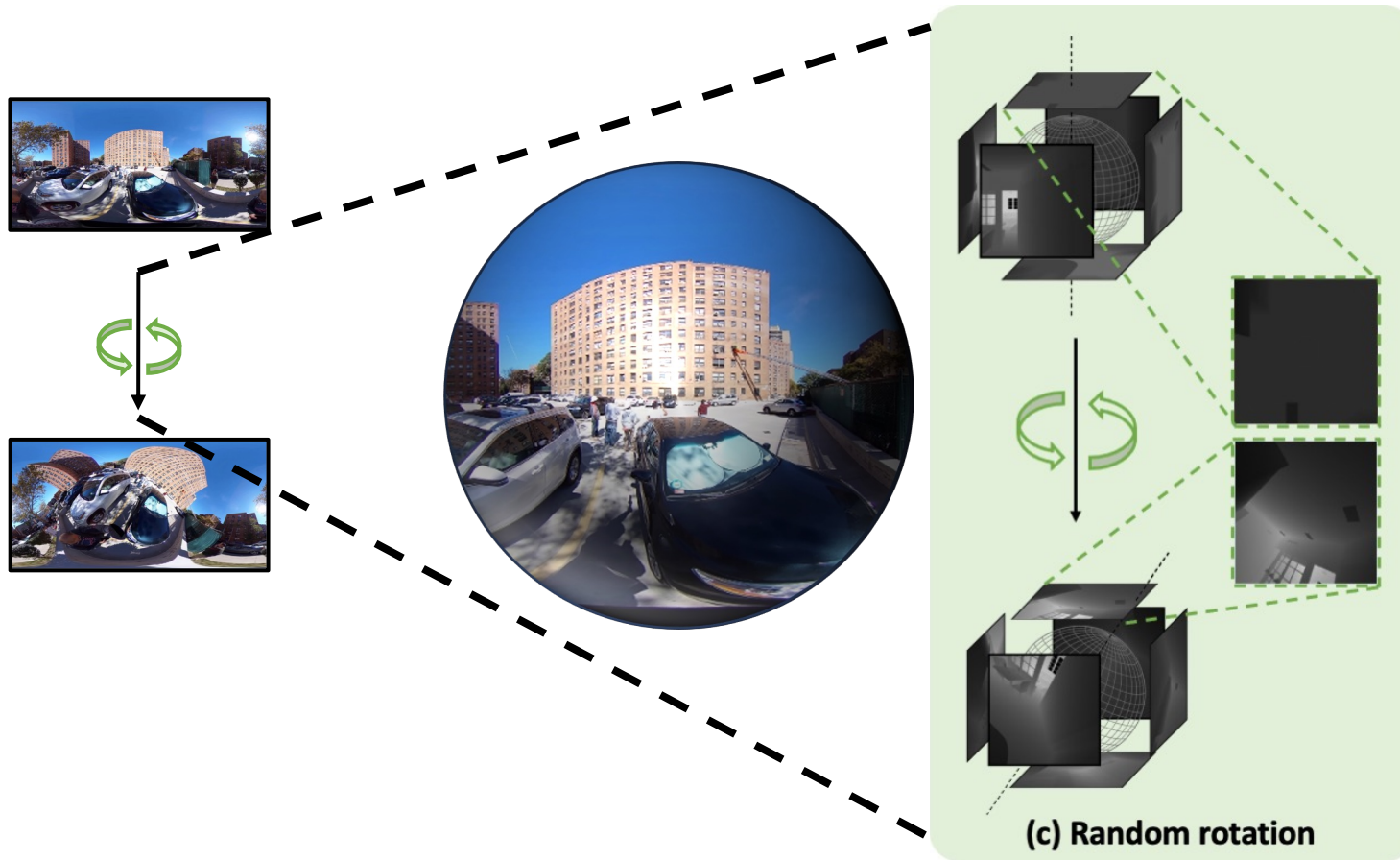
Cross Camera Model Knowledge Distillation

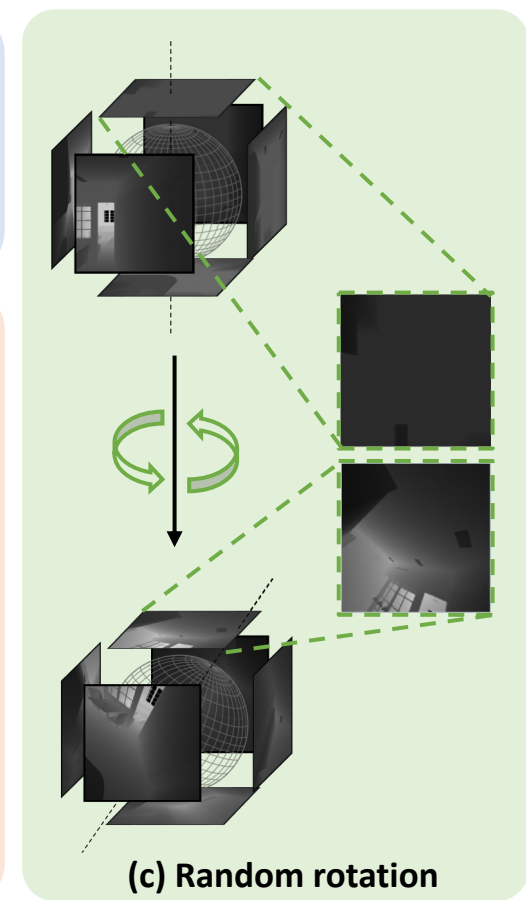
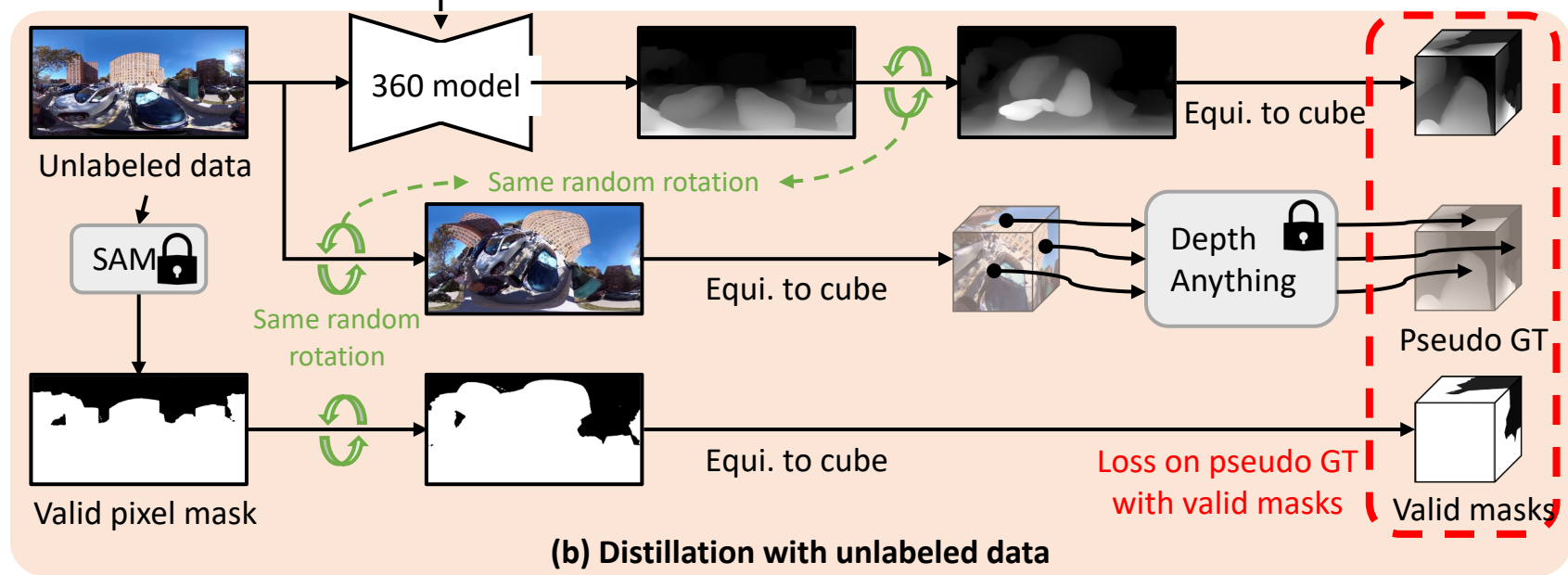
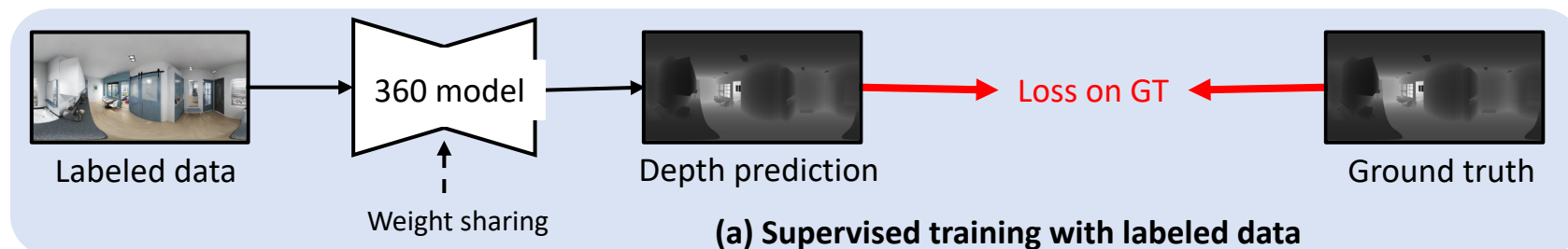


Cross Camera Model Knowledge Distillation



Cross Camera Model Knowledge Distillation





In-Domain Depth Estimation (Matterport3D)

Method	Loss	Train	Test	Abs Rel ↓	$\delta_1 \uparrow$	$\delta_2 \uparrow$	$\delta_3 \uparrow$
BiFuse [47]	BerHu	M	M	-	0.845	0.932	0.963
UniFuse [16]	BerHu	M	M	0.106	0.890	0.962	0.983
SliceNet [31]	BerHu	M	M	-	0.872	0.948	0.972
BiFuse++ [48]	BerHu	M	M	-	0.879	0.952	0.977
HRDFuse [1]	BerHu	M	M	0.097	0.916	0.967	0.984
UniFuse [16]	Affine-Inv	M	M	0.102	0.893	0.970	0.989
UniFuse [16]	Affine-Inv	M, ST-all (p)	M	<u>0.089</u>	0.911	<u>0.975</u>	<u>0.991</u>
BiFuse++ [48]	Affine-Inv	M	M	0.094	<u>0.914</u>	0.974	0.989
BiFuse++ [48]	Affine-Inv	M, ST-all (p)	M	0.085	0.917	0.976	0.991

In-Domain Video Demo



Zeroshot Depth Estimation (Stanford2D3D)

Method	Loss	train	test	Abs Rel ↓	$\delta_1 \uparrow$	$\delta_2 \uparrow$	$\delta_3 \uparrow$
BiFuse [47]	BerHu	M-all	SF	0.120	0.862	-	-
UniFuse [16]	BerHu	M-all	SF	0.094	0.913	-	-
BiFuse++ [48]	BerHu	M-all	SF	0.107	0.914	0.975	0.989
Depth Anything [59]	Affine-Inv	Pers.	SF	0.248	0.635	0.899	0.97
Marigold [19]	Affine-Inv	Pers.	SF	0.195	0.692	0.942	0.982
UniFuse [16]	Affine-Inv	M-all	SF	0.090	0.914	0.976	0.990
UniFuse [16]	Affine-Inv	M-all, ST-all (p)	SF	<u>0.086</u>	0.924	0.977	0.990
UniFuse [16]	Affine-Inv	M-all, SP-all (p)	SF	0.090	0.920	0.978	0.990
BiFuse++ [48]	Affine-Inv	M-all	SF	0.090	0.921	0.976	0.990
BiFuse++ [48]	Affine-Inv	M-all, ST-all (p)	SF	0.082	0.931	0.979	<u>0.991</u>
BiFuse++ [48]	Affine-Inv	M-all, SP-all (p)	SF	<u>0.086</u>	<u>0.926</u>	0.979	<u>0.991</u>
HoHoNet [45]	Affine-Inv	M-all	SF	0.095	0.906	0.975	0.991
HoHoNet [45]	Affine-Inv	M-all, ST-all (p)	SF	0.088	0.920	0.979	0.992
EGFormer [64]	Affine-Inv	M-all	SF	0.098	0.906	0.972	0.989
EGFormer [64]	Affine-Inv	M-all, ST-all (p)	SF	<u>0.086</u>	0.923	0.976	0.990

Zeroshot Video Demo



In-the-wild

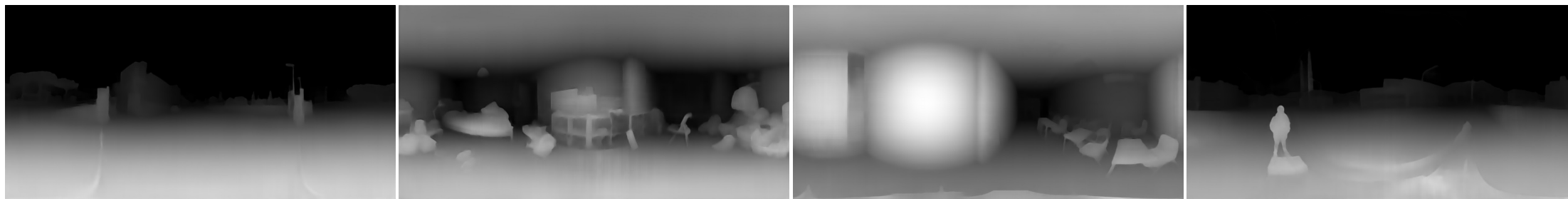
RGB



UniFuse



UniFuse (p)

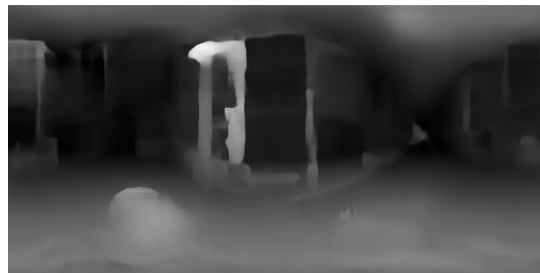
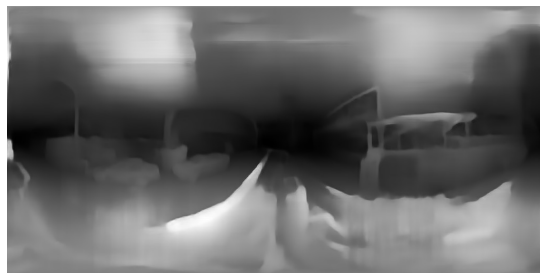
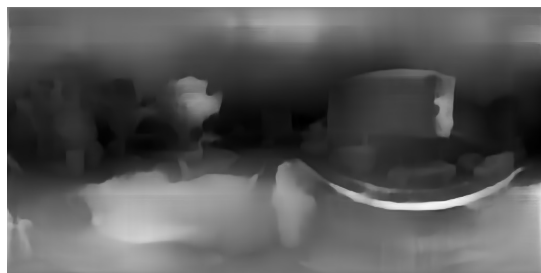


In-the-wild

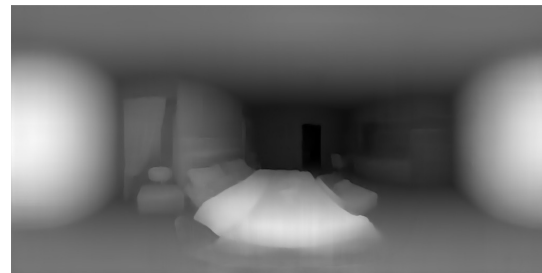
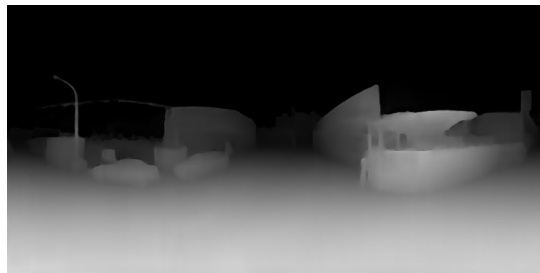
RGB



UniFuse



UniFuse (p)



Video Demo In-the-Wild



Depth Anywhere: Enhancing 360 Monocular Depth Estimation via Perspective Distillation and Unlabeled Data Augmentation

Welcome to visit us at Poster Session 4

Thank you

Project Page [Depth Anywhere](#)



 [GitHub](#)



 [Hugging Face Spaces](#)

