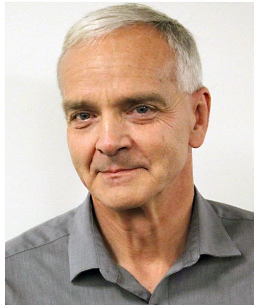
Generative Modeling for Multi-task Visual Learning


Zhipeng Bao


Martial Hebert


Yu-Xiong Wang

THE ROBOTICS INSTITUTE
Carnegie Mellon University

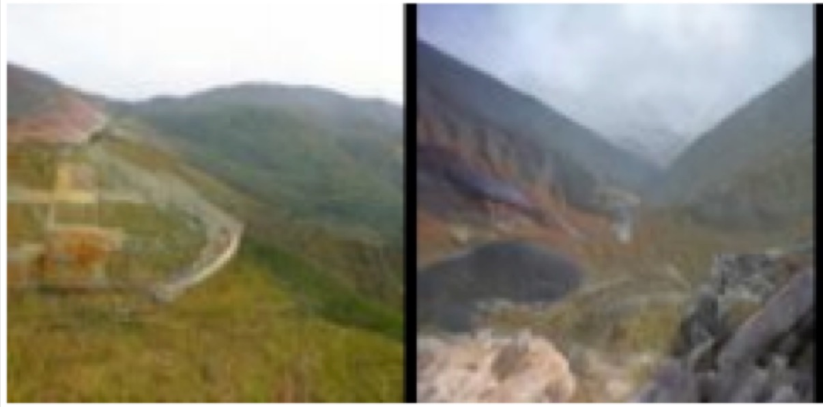
UNIVERSITY OF ILLINOIS URBANA-CHAMPAIGN

ICML
International Conference On Machine Learning

What can generative models do?

StyleGAN v2, CVPR 20

BigGAN, ICLR 19



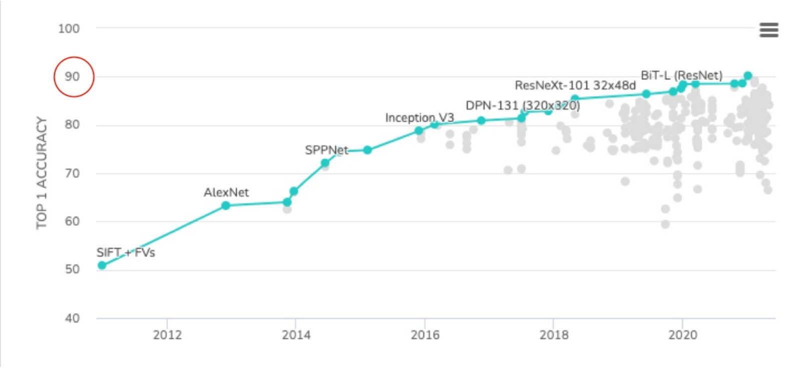
SAGAN, ICML 19

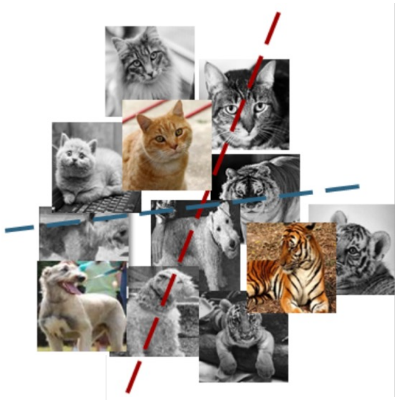
Realistic?

Useful!

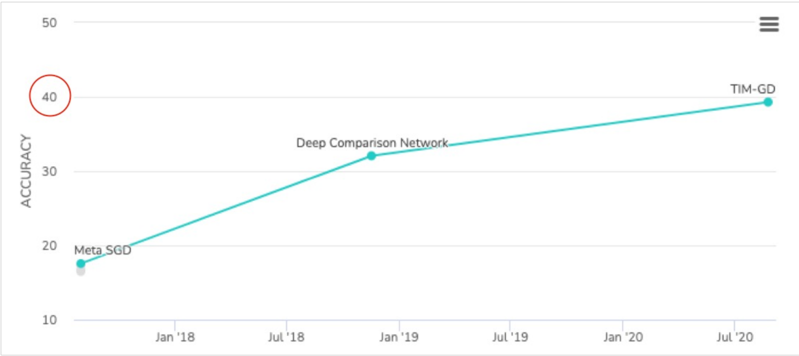
What can generative models do?

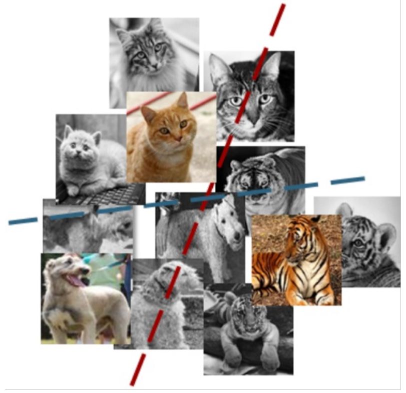
(Papers with codes, leaderboard)





(Papers with codes, leaderboard)

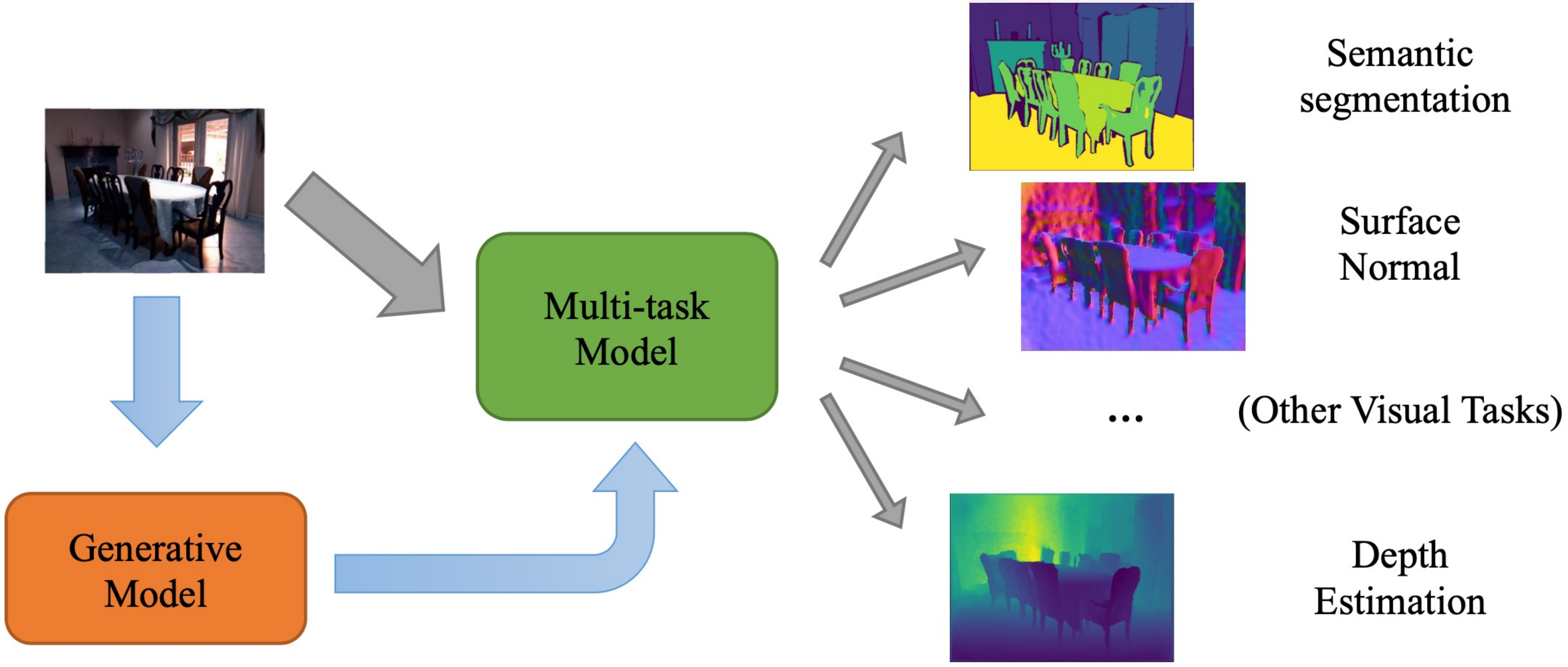




StyleGAN, CVPR 19

[Dent et al., CVPR 09, Vinyals et al., NeurIPS 16]

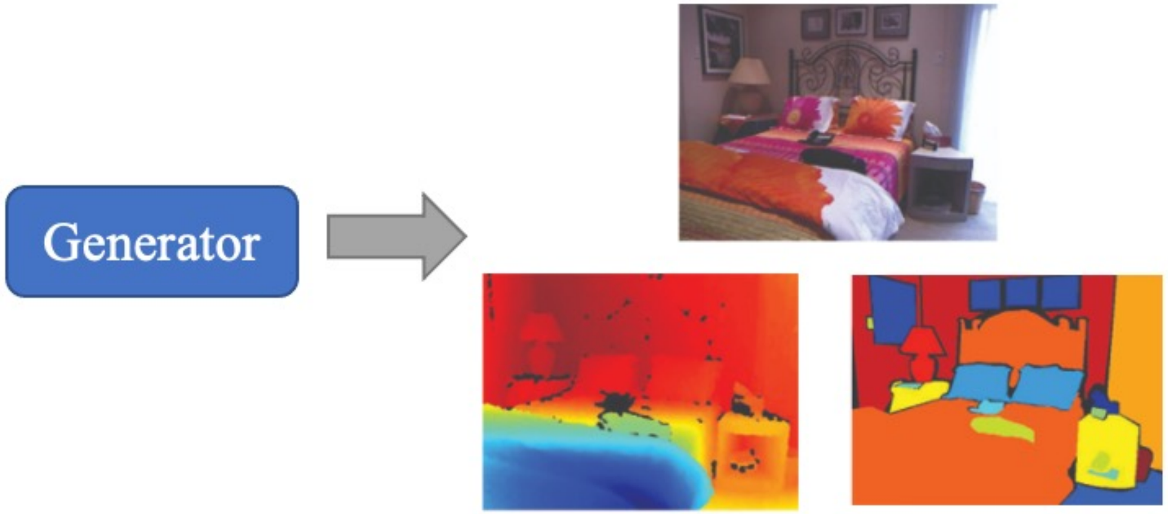
Leverage knowledge across multiple tasks: beyond a shared encoder



Multi-task learning with generative modeling:

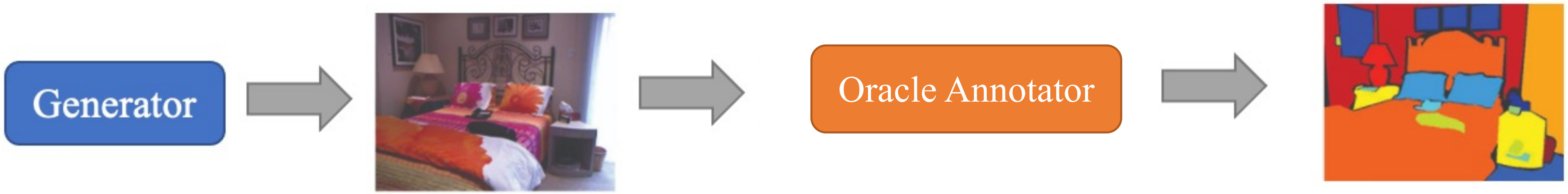
- Facilitate the flow of knowledge across tasks
- Synthesize data as augmentation to benefit multiple tasks

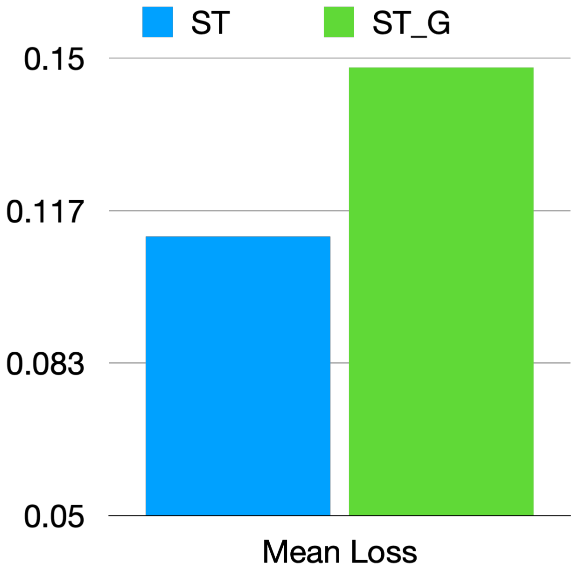
Naïve solution: synthesize paired image and pixel-wise annotations



Key challenge: difficult to synthesize images with pixel-wise annotations

Pilot study: can an oracle annotator help?





* Tiny-Taskonomy dataset: Mean Loss (↓)

- ST: Single Task model & real data
- ST_G: Single Task model & real + synthesized data

Target Task:

Semantic Segment

Failure Reasons:

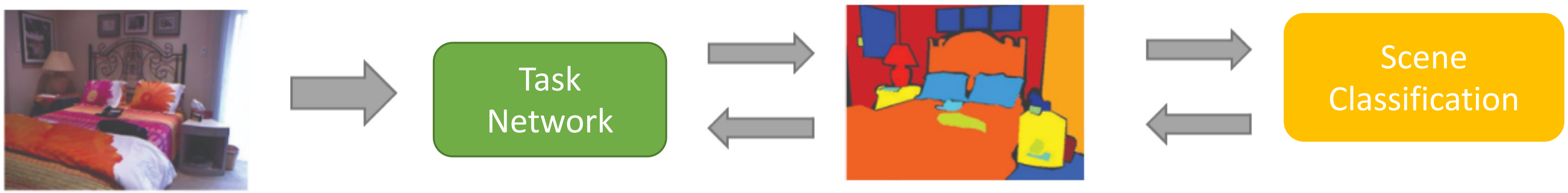
- No downstream signal

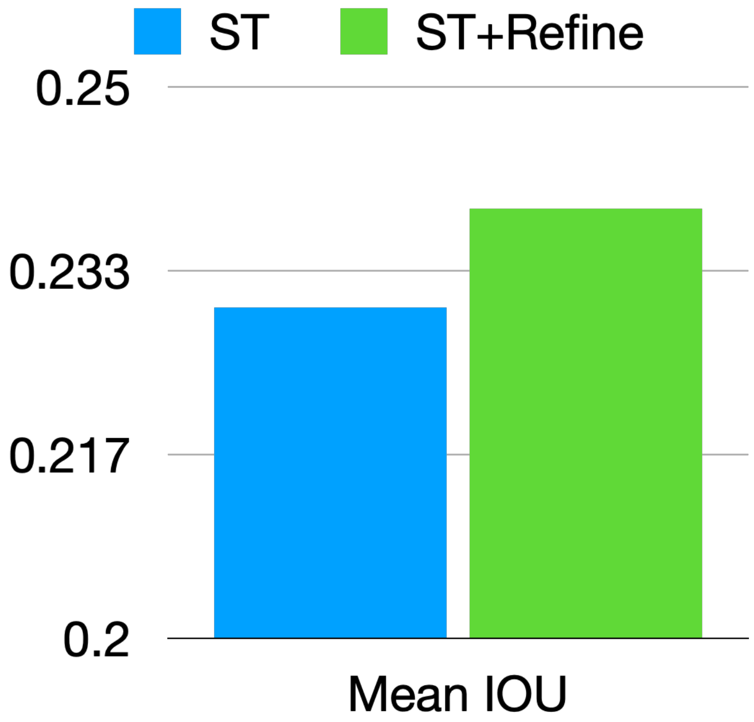
Other Concerns:

- Oracle annotators

How to utilize the synthesized examples?

(1) Refinement





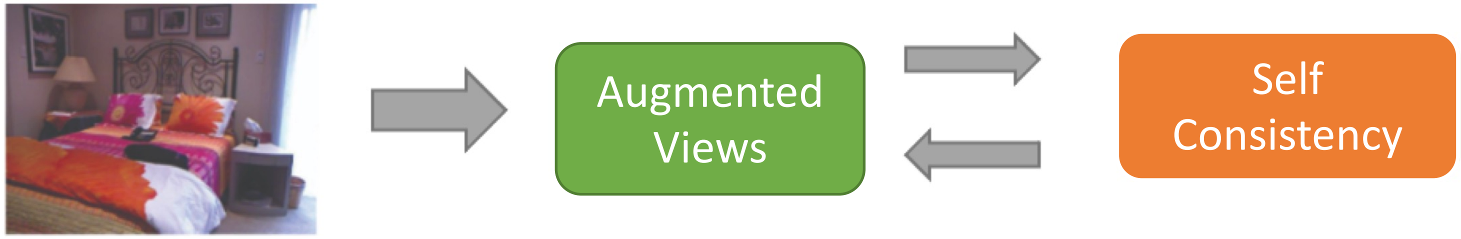
* Mean IOU (↑)

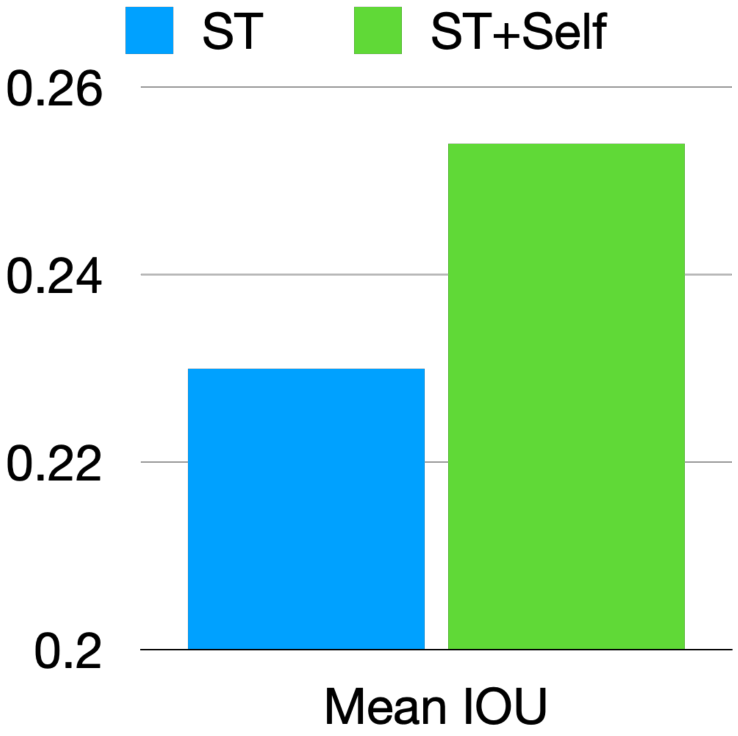
- ST: Single Task model & real data
- ST+Refine: Downstream classification task with synthesized images

Downstream signal gives guidance for synthesizing data!

How to utilize the synthesized examples?

(2) Self-Supervision





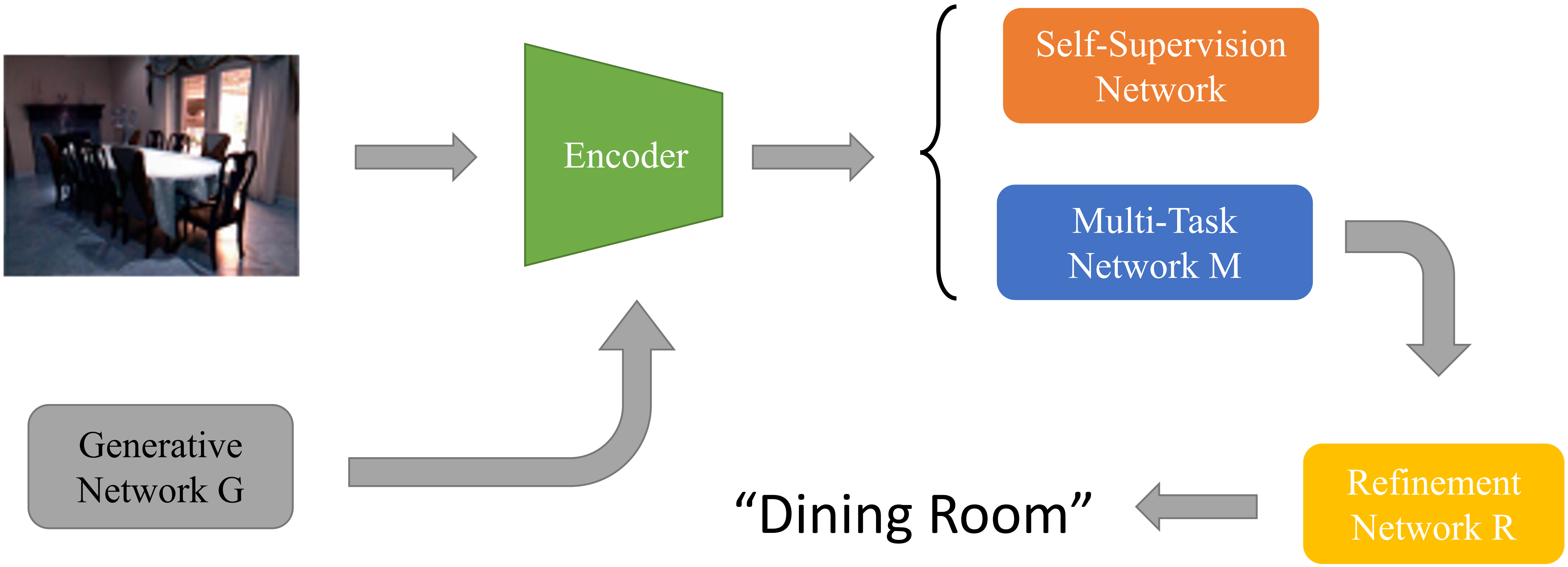
* Mean IOU (↑)

- ST: Single Task model & real data
- ST+Self: ST with self-supervised learning (SimCLR)

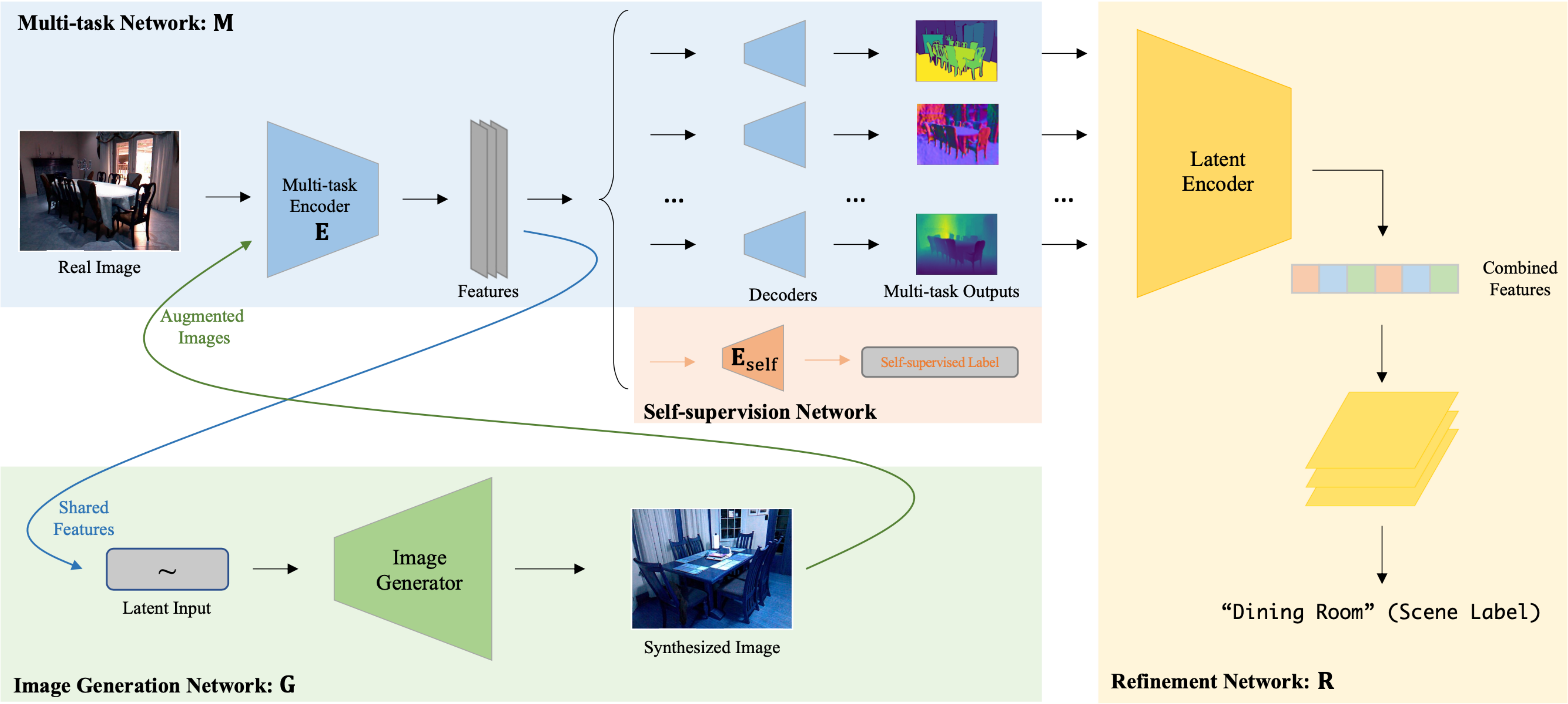
Self-supervision helps build effective features!

[Chen et al. ICML 20]

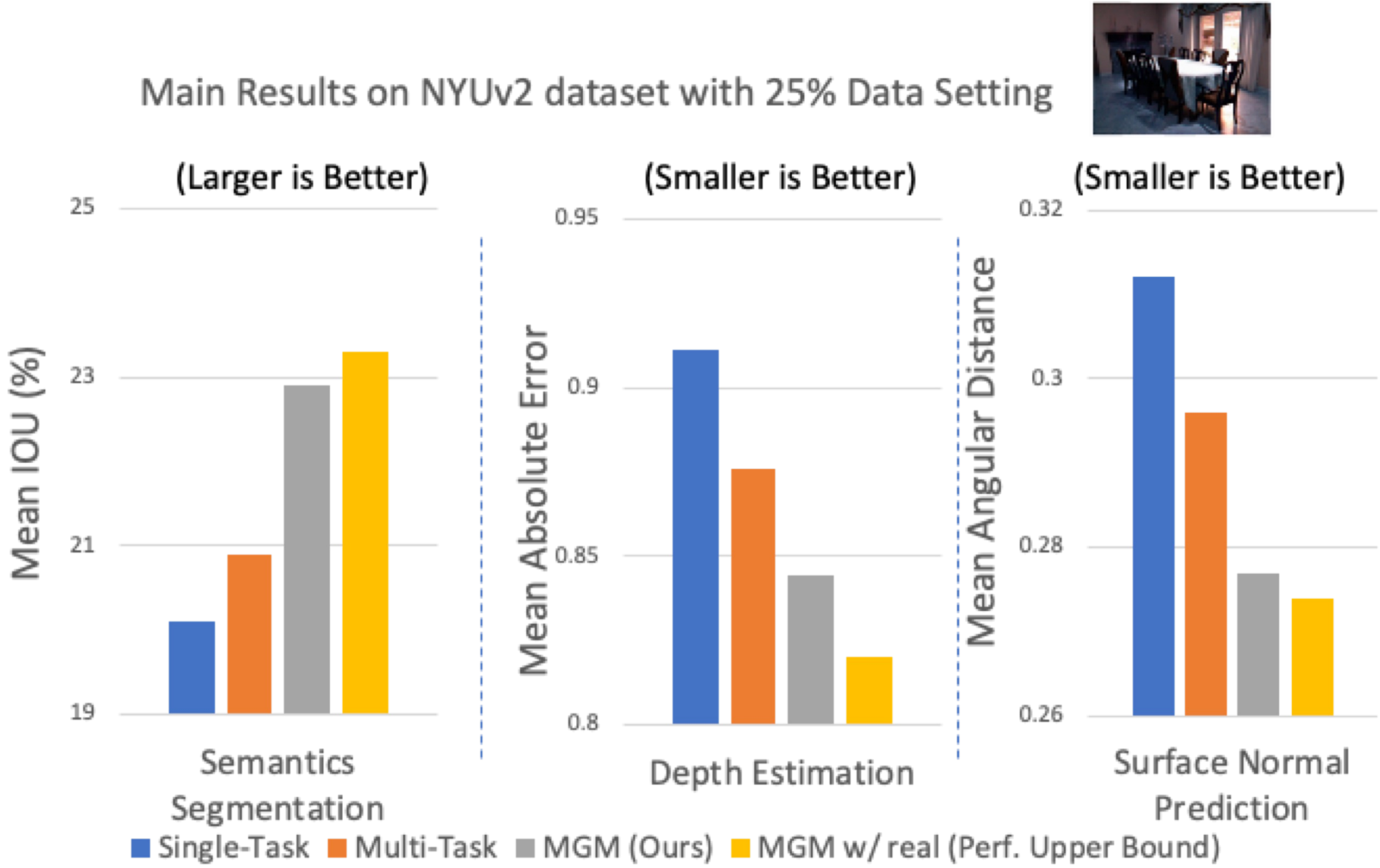
How to utilize the synthesized examples?



Multi-task oriented generative modeling (MGM)



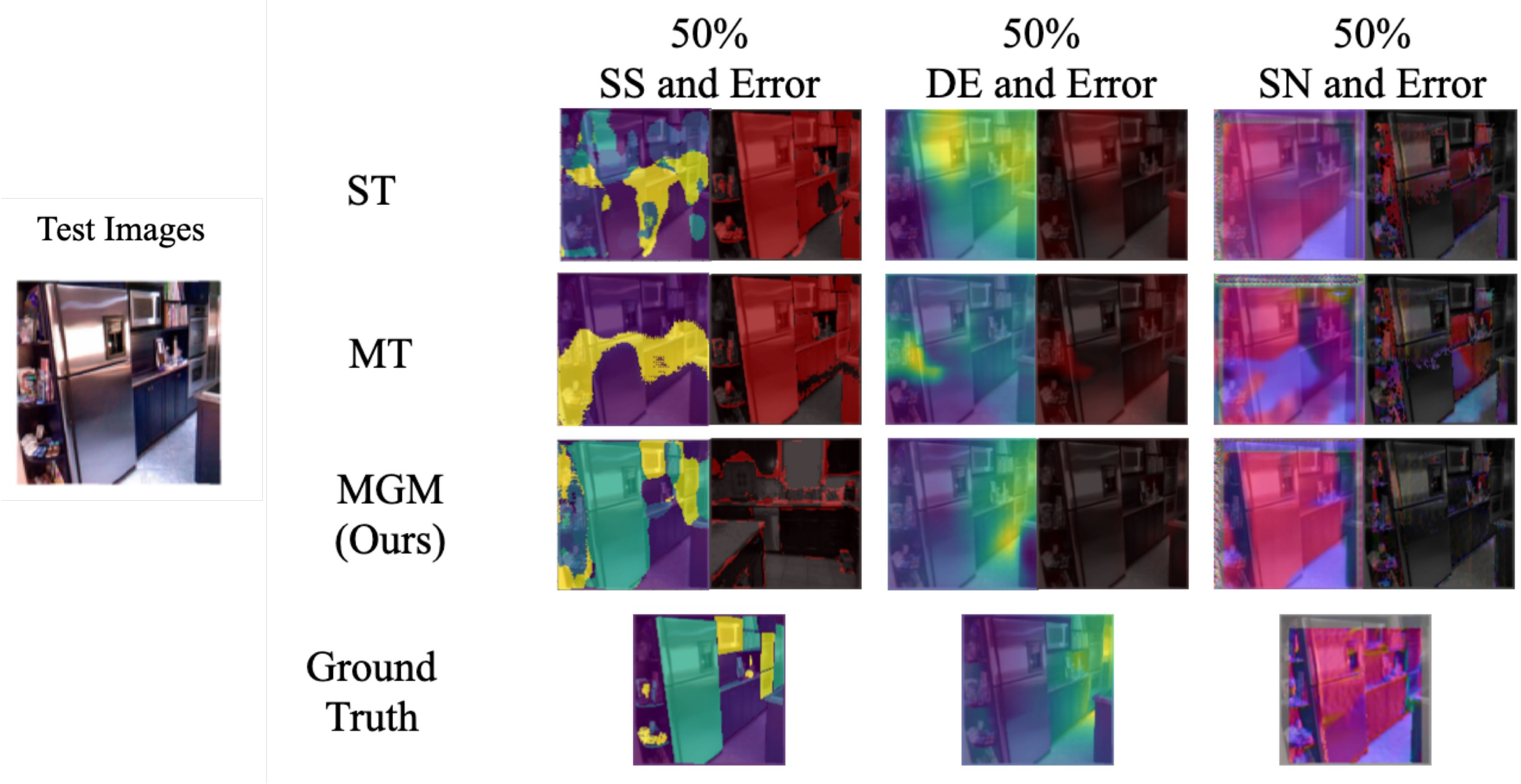
Performance improvements with MGM



- Tasks:
  - Semantic Segmentation (SS)
  - Depth Estimation (DE)
  - Surface Normal Prediction (SN)
- 25% Data Setting:
  - 25% Real images for ST / MT
  - 25% Real + 25% Synthesized images for MGM

[Silberman et al., ECCV 12, Zamir et al., CVPR 18]

Performance improvements with MGM



[Zamir et al., CVPR 18]

Code Available

Thank you!

Welcome to our poster (Session 2 Track 3)!