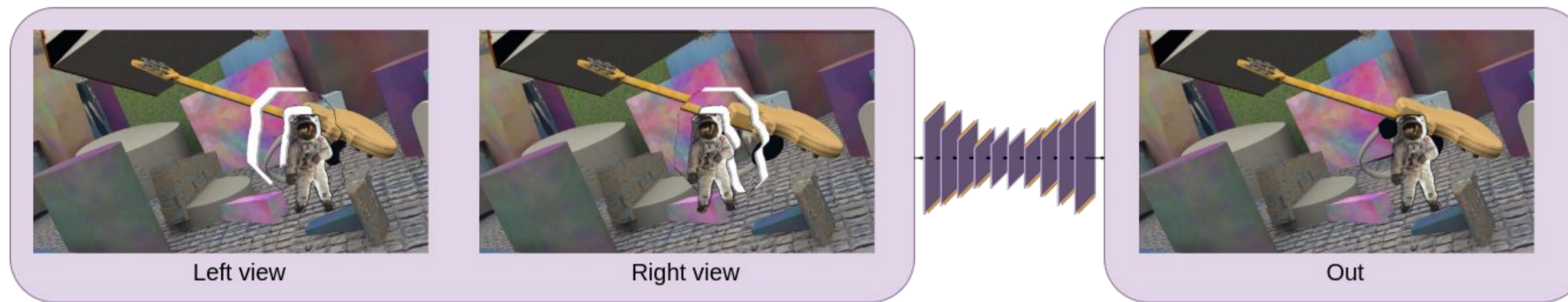


What do we do?

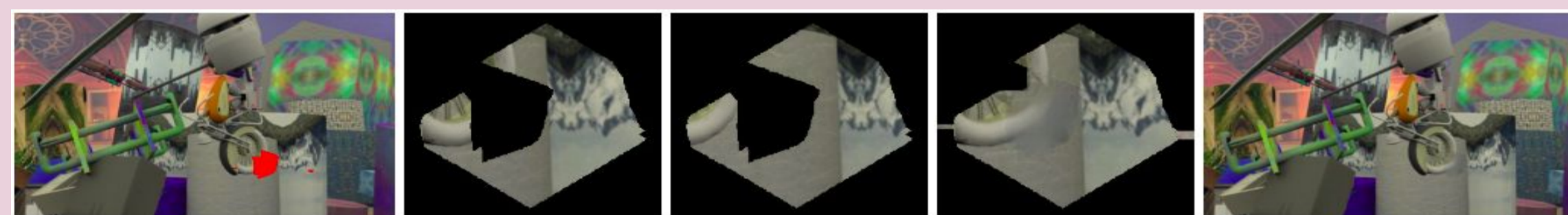


In multi-view settings we want to recover the missing information due to object disocclusions in a reasonable and cross-camera consistent way. Our approach uses object background information and stereo consistency to inpaint behind objects. In addition, our model converts a difficult unsupervised problem into an easier supervised one. Making it possible to train on larger stereo datasets.

Datasets

- Good quality, natural stereo datasets are very hard to come by.
- Random sampling of context-synthesis areas performs **data augmentation**.

FlyingThings3D[4] containing a variety of objects flying around in a randomised way.



Driving[4] naturalistic-looking dynamic street scene resembling the *KITTI* dataset.



Middlebury[5] 33 natural scenes.



Quantitative results

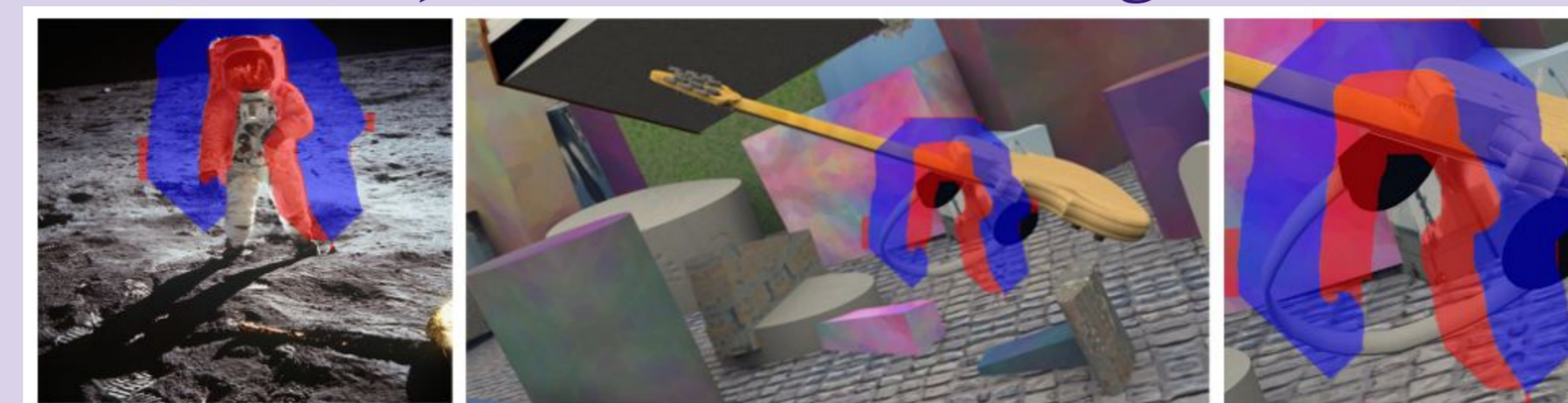
- Better across all metrics.
- Competitive consistency in more challenging problem.

Table 1. Quantitative results. Image quality & stereo consistency of different models. Bold is best. * values are from their paper.

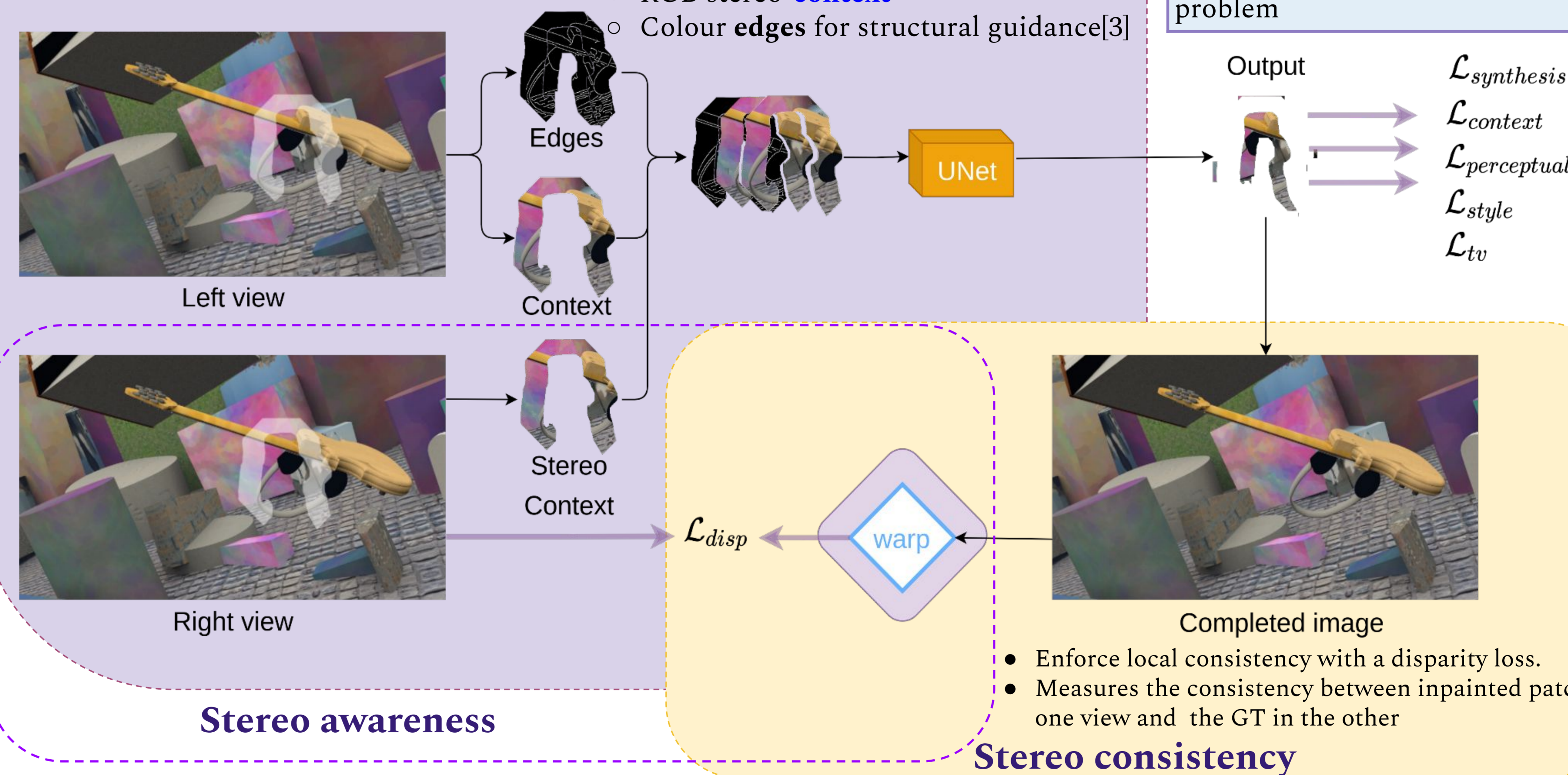
Dataset	Model	PSNR \uparrow	SSIM \uparrow	LPIPS \downarrow	DispE (%) \downarrow
FlyingThings3D	Shih <i>et al.</i> [11]	28.32	0.8589	0.0707	7.96
	Ours	30.50	0.8643	0.0556	7.67
Driving	Shih <i>et al.</i> [11]	30.46	0.969	0.1141	9.94
	Chen <i>et al.</i> [6]*	22.38	0.959	-	7.79
	Ma <i>et al.</i> [7]*	23.20	0.964	-	4.72
	Ours	34.94	0.977	0.0628	8.01

How do we do it?

Object occlusion regions



- Represent real objects
- More challenging as not necessarily bounded
- Generate masks from object boundaries:
 - **context** is the background around the object
 - **synthesis** is the hole behind the object
- Data augmentation:
 - Random sampling of context-synthesis areas
 - Sample from stereo view using disparity
- Input to the network
 - RGB **context**
 - RGB stereo-**context**
 - Colour **edges** for structural guidance[3]



What is new?

Problem	Solution
Novel view synthesis requires peeking behind objects and trying to recover information which is not present in an image	Leverage information from stereo-views
Stereo data is scarce	Perform data augmentation
Most inpainting approaches focus on randomly shaped inpainting masks of limited complexity	Use geometrically-meaningful object masks
Inpainting large and irregular non-bounded holes is a difficult problem	Use structure guided inpainting and extra stereo-information

Conclusions

- **Novel** stereo-aware learned inpainting model.
- Enforces **stereo consistency**
- Trained in a **self-supervision** fashion
- Uses **geometrically meaningful masks** representing object occlusions.
- Improvement over state-of-the-art models by up to **50% PSNR**.
- Good performance over several **diverse datasets**.
- Possible future work: extend model to cope with the challenges of wide-baseline non-parallel cameras.

References

- [1] Meng-Li Shih, Shih-Yang Su, Johannes Kopf, and Jia-Bin Huang. 3D Photography using Context-aware Layered Depth Inpainting. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Apr. 2020.
- [2] Guilin Liu, Fitsum A. Reda, Kevin J. Shih, Ting-Chun Wang, Andrew Tao, and Bryan Catanzaro. Image Inpainting for Irregular Holes Using Partial Convolutions. Proceedings of the European Conference on Computer Vision (ECCV), Apr. 2018.
- [3] Kanyar Nazeri, Eric Ng, Tony Joseph, Faisal Qureshi, and Mehran Ebrahimi. EdgeConnect: Structure Guided Image Inpainting using Edge Prediction. In IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), Oct. 2019.
- [4] Nikolaus Mayer, Eddy Ilg, Philip H. Isinger, Philipp Fischer, Daniel Cremers, Alexey Dosovitskiy, and Thomas Brox. A Large Dataset to Train Convolutional Networks for Disparity, Optical Flow, and Scene Flow Estimation. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016.
- [5] D. Scharstein, H. Hirschmüller, York Kitajima, Greg Krathwohl, Nera Nesić, Xi Wang, and P. Westling. High-Resolution Stereo Datasets with Subpixel-Accurate GroundTruth. In GCPR, 2014.
- [6] Shen Chen, Wei Ma, and Yue Qin. CNN-Based Stereoscopic Image Inpainting. In Int. Conf. on Image and Graphics (ICIG), Nov. 2019.
- [7] Wei Ma, Mana Zhang, Wenguang Ma, Shibao Xu, and Xiaopeng Zhang. Learning across views for stereo image completion. IET Computer Vision, 2020.
- [8] Jia-Ren Chang and Yong-Sheng Chen. Pyramid Stereo Matching Network. In IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 2018.

Qualitative results

- Superior image quality.
- More challenging masks.
- Sharp edges.
- Visually pleasing.
- Struggle with intricate unseen structures.

