

GRAMO: Geometric Resampling Augmentation for Monocular 3D Object Detection

He GUAN, Chunfeng SONG, Zhaoxiang ZHANG

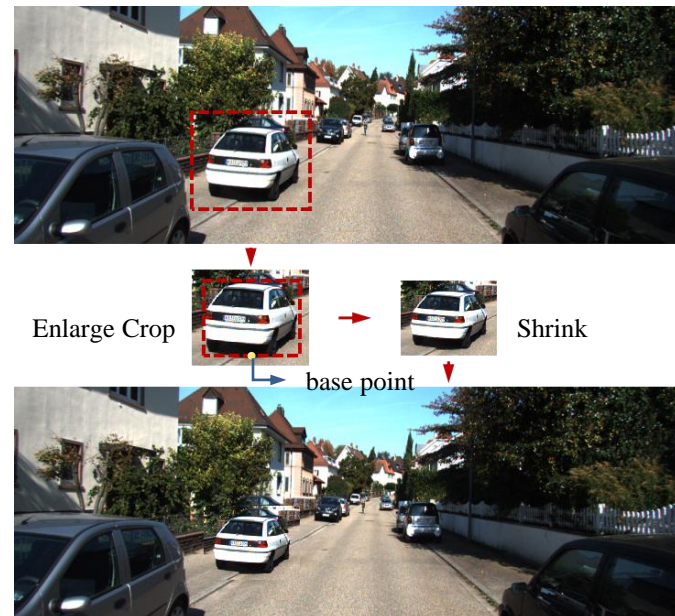
Frontiers of Computer Science, DOI: [10.1007/s11704-023-3242-2](https://doi.org/10.1007/s11704-023-3242-2)

Problems & Ideas

- Problems of conventional stereo matching approaches:
 - Neither single image nor point cloud data augmentation is perfectly adapted to monocular 3D inspection tasks.
 - Geometric consistency in perspective space affects both visual appearance and physical localization.
- Ideas: Geometric consistency are built directly into the proposed augmentation operations to improve sample richness that balances visual and physical plausibility.



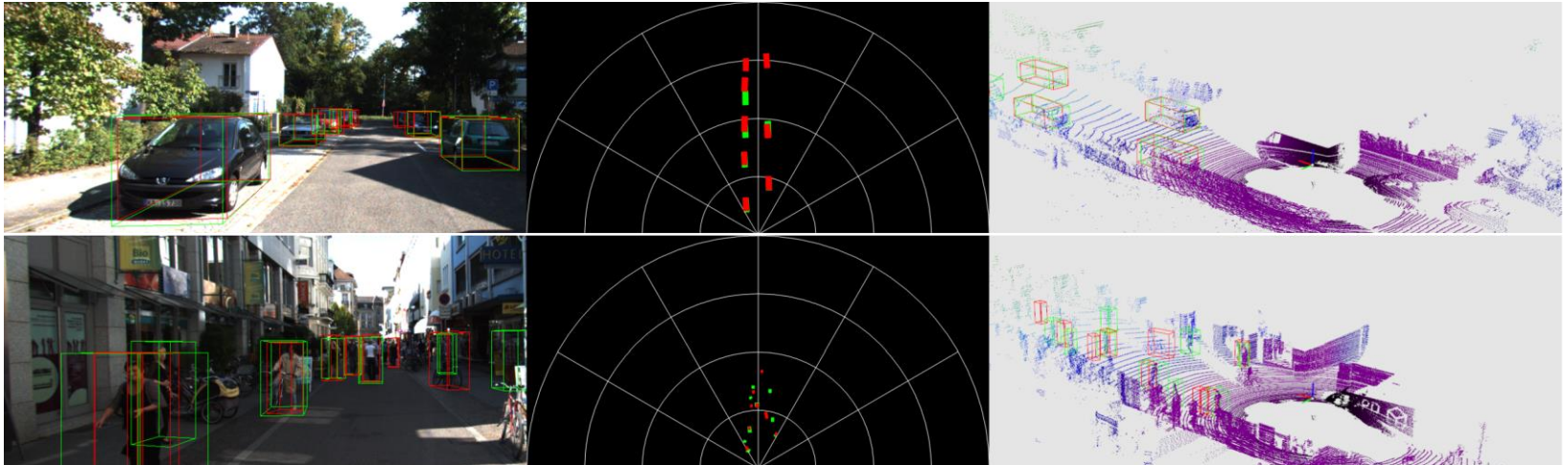
Geometric Copy-Paste



Geometric Crop-Shrink

Main Contributions

- Contributions:
 - A novel geometry-level augmentation operator for monocular 3D that utilizes depth re-ordering and bird's-eye de-overlap to insert new instances into the scene and maintain 2D-3D geometric consistency.
 - A novel geometry-level augmentation operator for monocular 3D that utilizes perspective to simultaneously modify instance appearance scale and physical distance and maintains 2D-3D geometric consistency.



Qualitative results on the KITTI dataset. The prediction results are shown from left to right in the following order: image, BEV and Lidar. We color the predictions red and the ground truths green. LidAR signals are shown for visualization only.