

EdgeSplats: Robust 3D Edge Reconstruction for In-The-Wild Data

Supplementary Material

A. Implementation Details

Our Gaussian training follows the 3DGS-MCMC protocol [12], with scale and opacity regularization coefficients set to 0.01 and positional noise scale 5×10^5 . For all scenes we cap the number of Gaussians at 5×10^5 and train for 30,000 iterations. We supervise edge-aligned Gaussians with PiDiNet [23] edge maps. Other detectors such as DexiNed [21] are also compatible; in qualitative tests, we observed similar reconstructed structure, with DexiNed often producing denser but less noise-robust edge maps. We use PiDiNet consistently for both supervision and reprojection-based evaluation to avoid introducing bias from mismatched edge priors and to ensure fair comparison across methods. All baselines were evaluated against the same PiDiNet-derived targets under the same protocol.

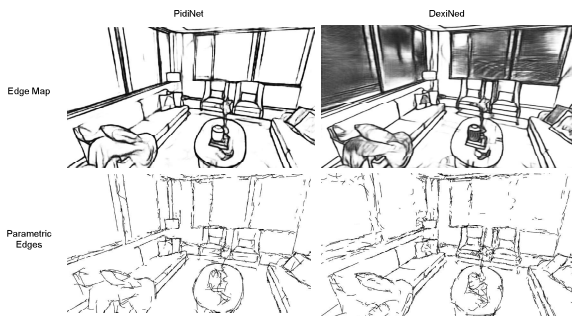


Figure 11. Qualitative comparison of supervisory edge detectors on Replica Room0. DexiNed produces denser edge responses and recovers more fine detail, but also amplifies high-frequency texture, increasing noise. PiDiNet yields cleaner and more stable reconstructions by prioritising dominant structural contours, at the expense of some fine-scale completeness.

Our GAT encoder uses three hidden layers of width 128 and an output dimension of 64. On Replica [22], DBSCAN uses $\varepsilon = 0.09$ for the first two passes (constituent Gaussians) and $\varepsilon = 0.13$ for the third pass (mega-Gaussian clustering). On ETH3D [20] and Living Lab [17], we use $\varepsilon = 0.065$ and $\varepsilon = 0.1$ for these stages, respectively. In practice, ε is selected empirically using a small validation subset to balance over- and under-segmentation in the learned embedding space. If clusters are overly fragmented or sparse, ε is increased to encourage connectivity; if distinct structures are incorrectly merged, ε is reduced. We found the reported values to transfer consistently within each dataset without scene-specific retuning.¹

¹Reported ε values correspond to passes 1–2, pass 3, in order.

For Replica and Living Lab we initialize poses and points with COLMAP estimates to mimic in-the-wild capture; for ETH3D we use the poses and points provided by the dataset authors.

B. Additional Results

To provide full comparisons with EMAP [13] and EdgeGaussians [4], we include additional results on *room1* and *room2* of Replica (Figs. 12, 13). Our method achieves comparable completeness with noticeably lower noise, while running substantially faster. Under the described setup, EdgeGaussians often closely follows the supervised PiDiNet maps, yielding multiple fragmented segments rather than longer semantic structures, whereas EMAP shows incomplete regions and the longest training times.

We also present qualitative results on *terrace* and *door* from ETH3D [20]. In Fig. 14, EdgeGaussians reconstructs the general layout but exhibits spurious floating artifacts; Fig. 15 shows a failure case where EdgeGaussians collapses to noise. In contrast, EdgeSplats remains robust across both scenes without method-specific tuning.

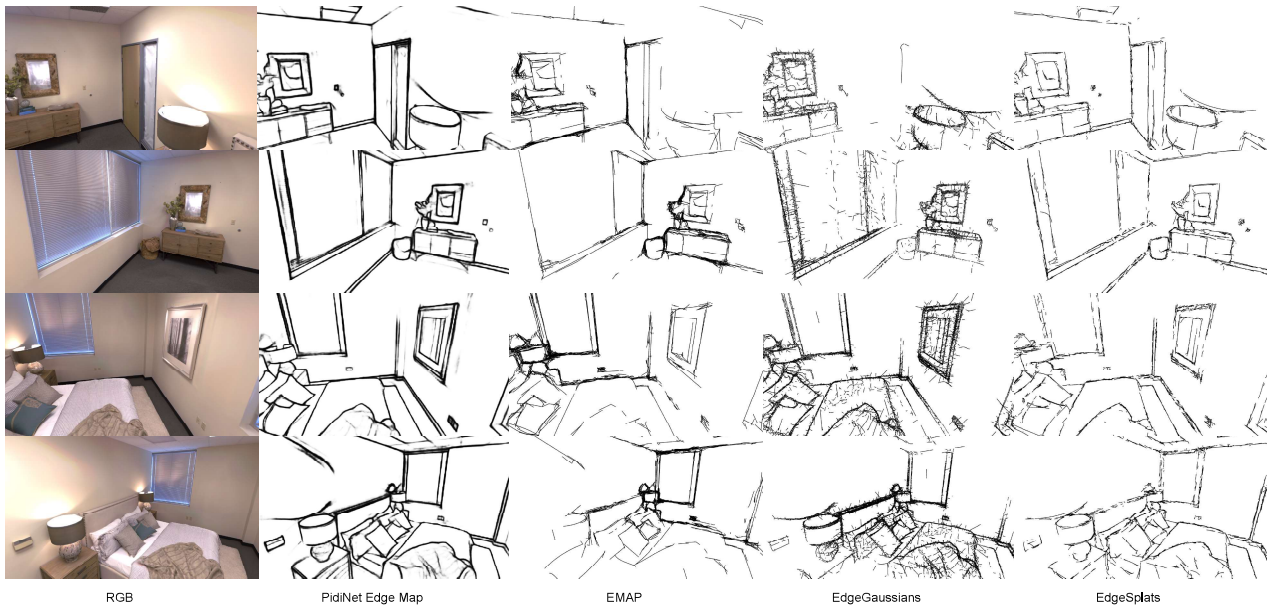


Figure 12. Qualitative results on *room 1* from the Replica dataset. EdgeSplats produces more complete renders than EMAP, whilst being less noisy than EdgeGaussians.

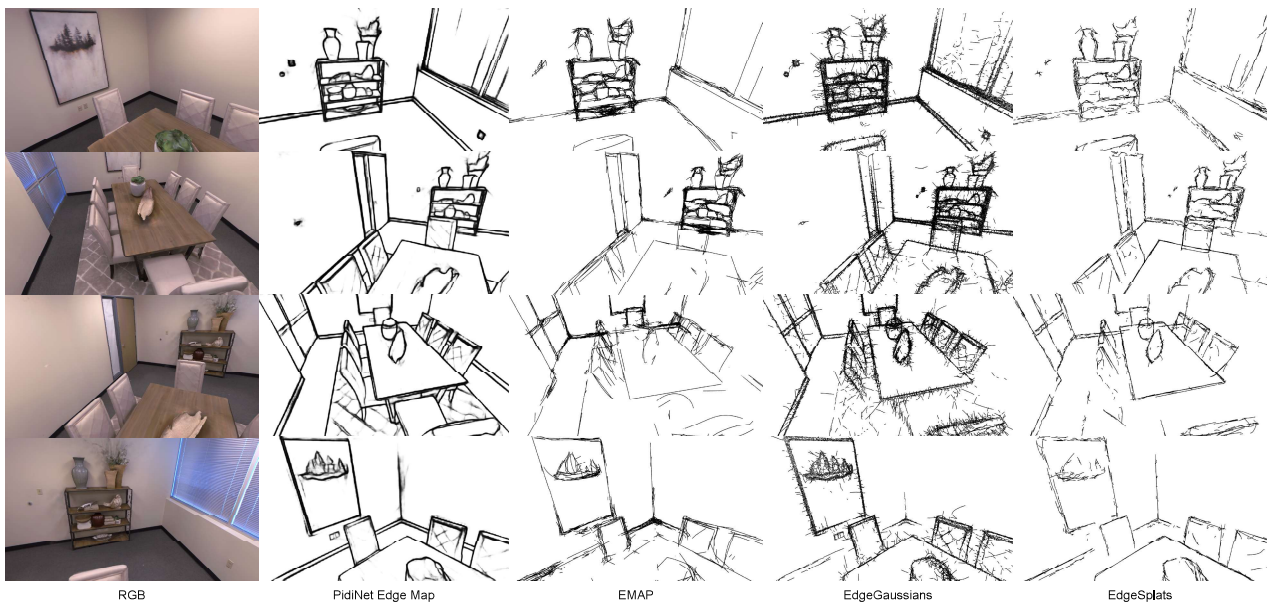


Figure 13. Qualitative results on *room 2* from the Replica dataset. EdgeGaussians may look more 'accurate' to the training images, however closer inspection shows this is largely due to fragmented edge sections.

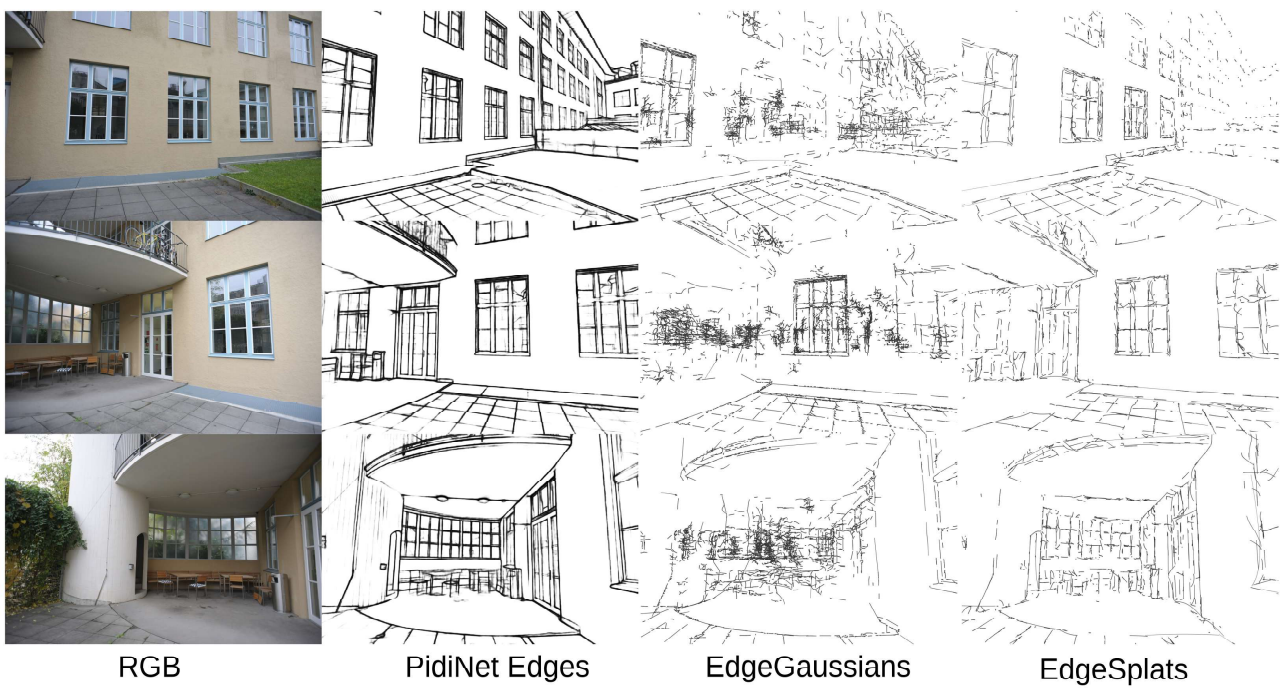


Figure 14. Results on the *terrace* scene from ETH3D, a more challenging scene than those provided in Replica. Both options produce semi-complete reconstructions, however EdgeSplats suffers from less noise.

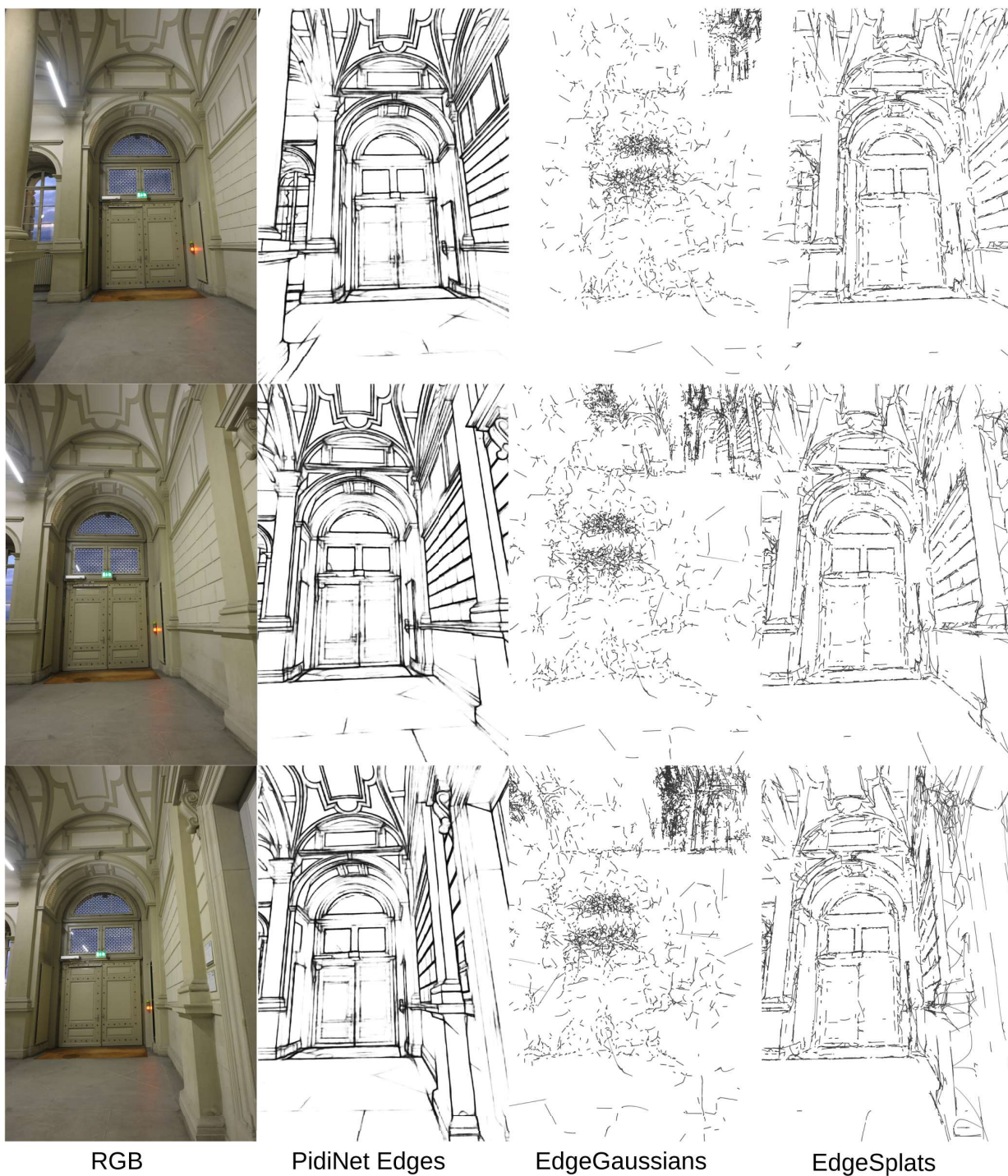


Figure 15. Results on the *door* scene from ETH3D, a scene composed of only 8 training views. EdgeGaussians falls into a degenerate configuration, whilst EdgeSplats remains robust.