

336 A Gaussian Splatting Details

337 We initialize the 3DGS start points with points deprojected from DROID-SLAM disparity, de-
338 duplicated. To accomodate small amounts of SLAM drift, we enable camera optimization, similar
339 to (Seiskari et al., 2024; Zhao et al., 2024), which refines poses during 3DGS training. To deal with
340 the large scene scale in our captures, we need to modify a few important hyperparameters inside
341 Splatfacto, which was primarily tuned for use with object-centric captures. We use the AbsGS split-
342 ting heuristic introduced by (Ye et al., 2024) with a gradient threshold of 0.0006 as well as lowered
343 learning rate on the Adam optimizer controlling 3DGS means to $5e - 5$ as recommended by the orig-
344 inal implementation. We also extend the length of the culling/splitting portion of 3DGS training from
345 15000 to 25000 steps, and lower the densify size threshold to 0.1% from 1%, ensuring fine details
346 are reconstructed.

347 B Hardware Specifications

348 The PAWS data collection system consists of three cameras, one on each shoulder, and a third camera
349 pointed at the feet. The shoulder-mounted cameras are stereo ZedX Mini cameras, while the down-
350 ward facing camera is a stereo Zed Mini. For depth estimation we opt for the provided neural depth
351 model, which is slightly less accurate than RAFT-Stereo (Lipson et al., 2021) yet considerably faster.
352 The three cameras are connected to a rigid plate attached to the chest, and their precise transforms
353 are derived from the CAD model. All three cameras are connected to a NVIDIA Jetson AGX Orin
354 Developer Kit which hardware encodes and saves the three camera streams for later processing.

355 C Rendering Performance Analysis

356 3D Gaussian Splatting models support fast batch rendering, capable of rendering thousands of frames
357 per second at certain resolutions.

Resolution	FPS (RGB + Depth)
64×64	1487 ± 12
96×96	1023 ± 9
256×256	756 ± 9
512×512	431 ± 6
1024×1024	185 ± 3

Figure C.1: Rendering frames per second for various resolutions. Reported times are for computing both RGB and Depth values from the 3DGS.