SLAM in the Field: An Evaluation of Monocular Mapping and Localization on Challenging Dynamic Agricultural Environment
(Supplementary Material)

Fangwen Shu   Paul Lesur   Yaxu Xie   Alain Pagani   Didier Stricker

Overview of Supplementary Material:

- In Section 1: we provide more detailed experiments results for the **reproducibility** in terms of Stereo, Mono, simulated RGB-D OpenVSLAM on dataset Rosario [4].
- In Section 2: we illustrate 3D trajectories estimated from visual SLAM on dataset Rosario with support of **xyz_view**.
- After Section 2: we present more results regarding to **depth estimation** using Monodepth2 [2] on dataset Rosario.

1. Estimated ATEs on Rosario

Here, we present each single value calculated from 5 runs experiments on each data sequence. Please see Table 1, 2, 3, 4, 5 and 6.

OpenVSLAM	Sequence					
Stereo	01	02	03	04	05	06
test1	1.345510	1.891184	1.382651	1.457684	1.634542	3.261564
test2	1.284283	2.133062	1.690512	1.453359	1.626288	3.442307
test3	1.476829	1.812399	2.026922	1.54984	1.479779	3.171042
test4	1.150358	1.810362	1.841467	1.529698	1.583595	3.166508
test5	1.484909	2.082111	1.804072	1.426992	1.924564	3.999046
Average	1.348378	1.945824	1.749125	1.483515	1.649755	3.408093

Table 1. Stereo OpenVSLAM, **without** scale correction.

OpenVSLAM	Sequence					
Mono	01	02	03	04	05	06
test1	13.395258	30.593833	4.254778	5.460637	22.027573	86.336890
test2	7.963333	26.226151	4.300106	6.045350	20.641006	94.142169
test3	6.219473	29.838208	4.250377	6.137460	27.741376	94.926586
test4	9.95986	28.716891	4.325487	6.205326	26.71388	94.002996
test5	13.395258	25.488622	4.332335	6.874662	21.171525	86.24307
Average	10.186636	28.172741	4.292617	6.144687	23.659072	91.130342

Table 2. Monocular OpenVSLAM, **with** scale correction.

2. Estimated 3D Trajectory on Rosario

The illustrated 3D trajectory with xyz_view please see Figure 1, 2, 3 and 4.

OpenVSLAM	Sequence					
Mono+D$_{GT}$	01	02	03	04	05	06
test1	8.278413	5.084219	5.823108	4.371406	5.634116	13.279659
test2	8.355741	4.891129	5.569139	4.203990	6.152829	14.148369
test3	7.791463	5.049453	6.331629	4.417157	5.948545	13.807191
test4	8.33554	4.856672	5.629635	4.037858	5.795523	13.33655
test5	8.857831	4.840701	5.125538	4.497018	6.049574	13.410093
Average	8.323798	4.944435	5.695810	4.305486	5.916117	13.596372

Table 3. Mono+D$_{GT}$ OpenVSLAM, **without** scale correction.

OpenVSLAM	Sequence					
Mono+D$_{GT}^{scale}$	01	02	03	04	05	06
test1	7.358559	2.025614	0.651759	0.233053	2.516591	5.764439
test2	7.452016	1.945064	0.731986	0.220936	2.494906	6.008664
test3	6.970739	2.054693	0.700865	0.303697	2.458734	5.811211
test4	7.458441	2.16323	0.66136	0.247094	1.940523	5.653927
test5	7.976594	1.966845	0.64334	0.241459	2.525089	5.650581
Average	7.443270	2.031090	0.677862	0.249248	2.387169	5.777764

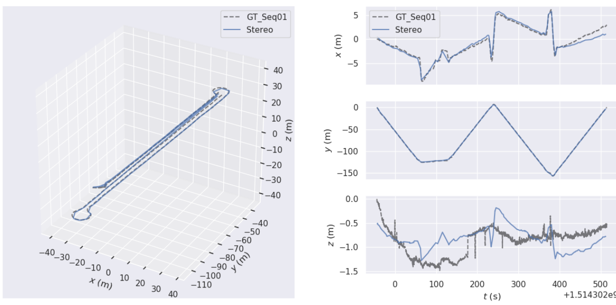
Table 4. Mono+D$_{GT}^{scale}$ OpenVSLAM, **with** scale correction.

OpenVSLAM	Sequence					
Mono+D$_{CNN}$	01	02	03	04	05	06
test1	5.031927	3.230896	2.574448	2.634001	2.706000	7.527925
test2	5.567432	3.444037	2.607057	3.088945	2.487649	7.643749
test3	5.501320	3.612542	2.574382	3.265632	2.927987	8.220594
test4	5.582937	3.121221	2.792248	2.991206	2.509777	7.642227
test5	5.172661	3.644995	2.567709	2.733721	2.523757	7.893021
Average	5.371255	3.410738	2.623169	2.942701	2.631034	7.785503

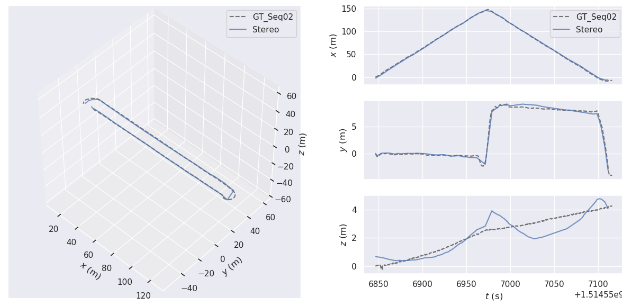
Table 5. Mono+D$_{CNN}$ (trained model MS*) OpenVSLAM, **without** scale correction.

OpenVSLAM	Sequence					
Mono+D$_{CNN}^{scale}$	01	02	03	04	05	06
test1	2.736342	1.302296	0.527764	0.227795	1.944690	5.725328
test2	3.843360	1.157363	0.486486	0.263382	1.388780	6.001030
test3	3.288103	1.344954	0.556806	0.231819	2.315588	6.766430
test4	3.592999	1.138049	0.615431	0.335533	1.601014	6.121325
test5	2.978706	1.363066	0.656692	0.227342	1.498768	6.263446
Average	3.287902	1.261146	0.568636	0.257174	1.749768	6.175512

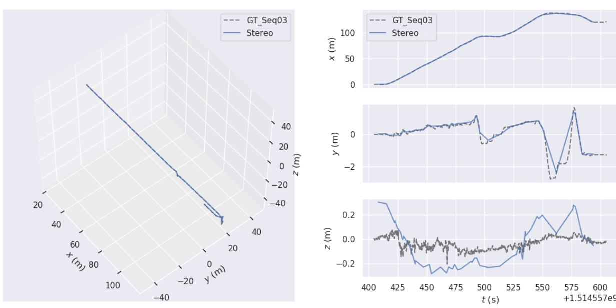
Table 6. Mono+D$_{CNN}^{scale}$ (trained model MS*) OpenVSLAM, **with** scale correction.



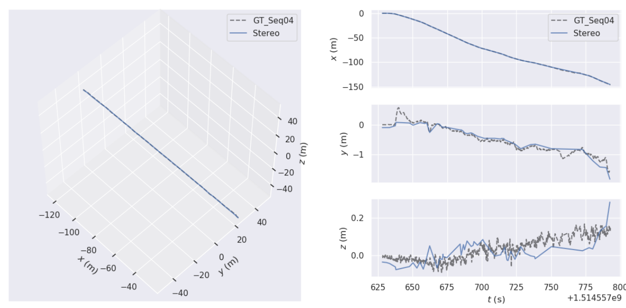
(a) Seq. 01



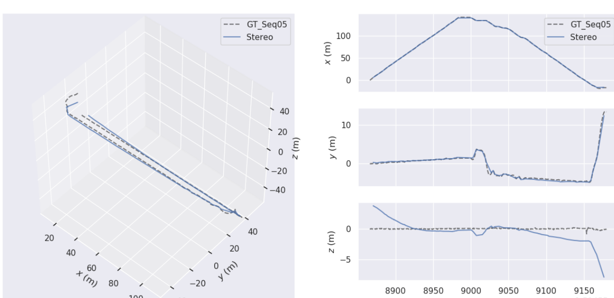
(b) Seq. 02



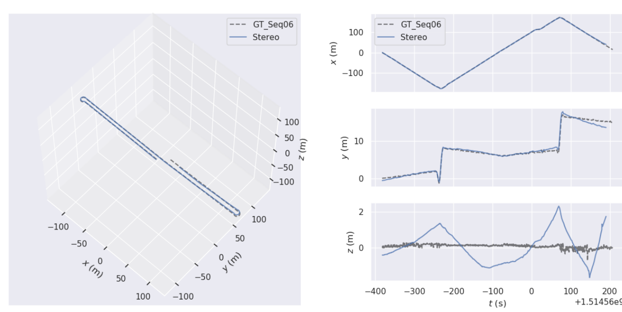
(c) Seq. 03



(d) Seq. 04

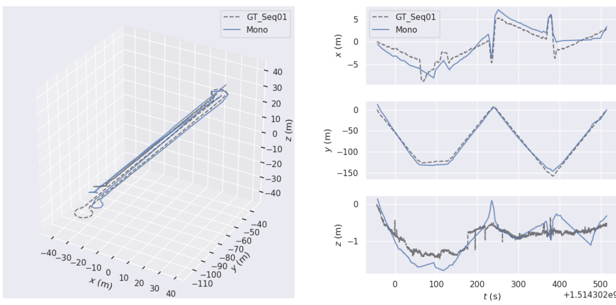


(e) Seq. 05

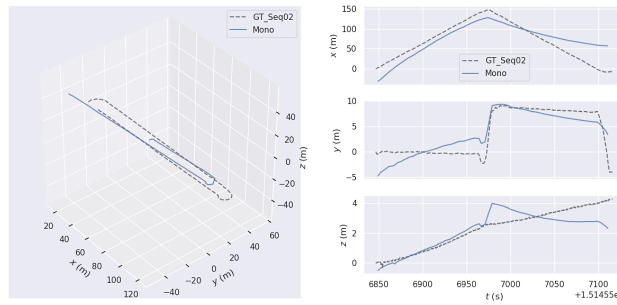


(f) Seq. 06

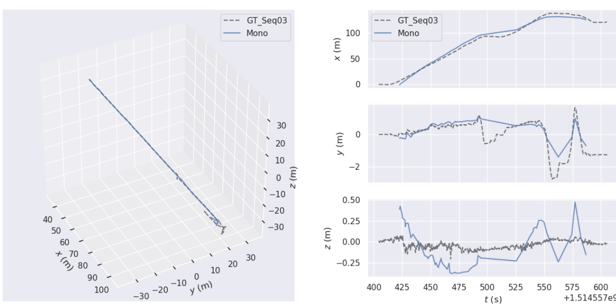
Figure 1. 3D trajectory estimated from **Stereo** OpenVSLAM.



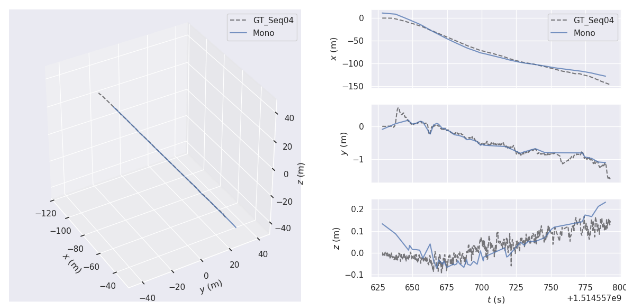
(a) Seq. 01



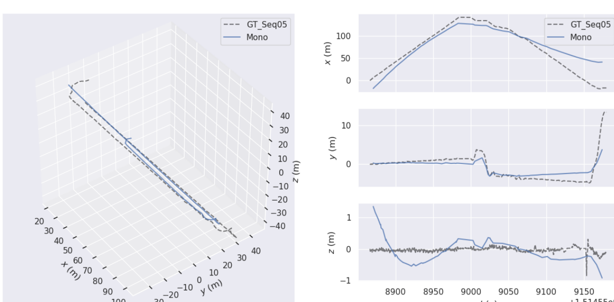
(b) Seq. 02



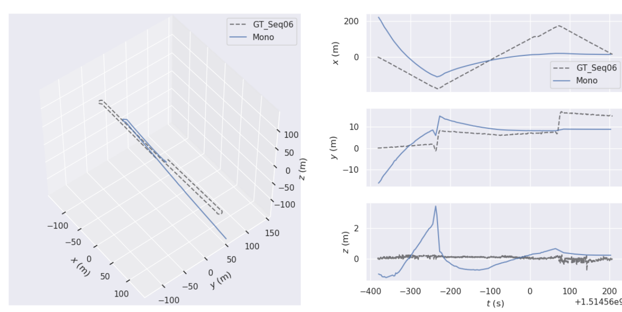
(c) Seq. 03



(d) Seq. 04

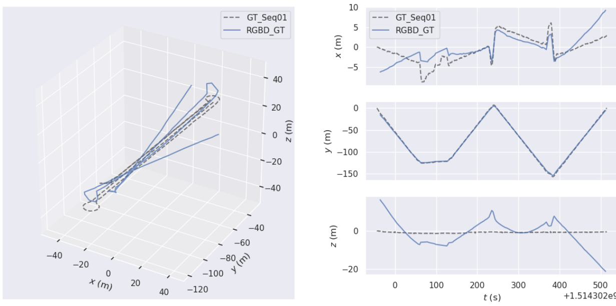


(e) Seq. 05

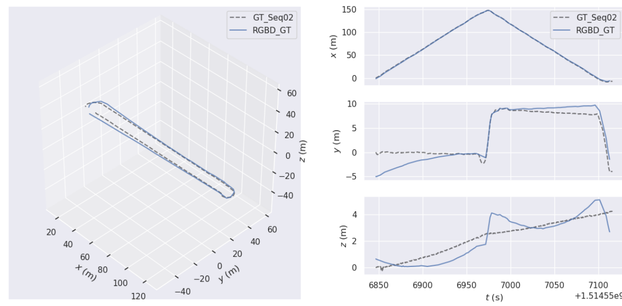


(f) Seq. 06

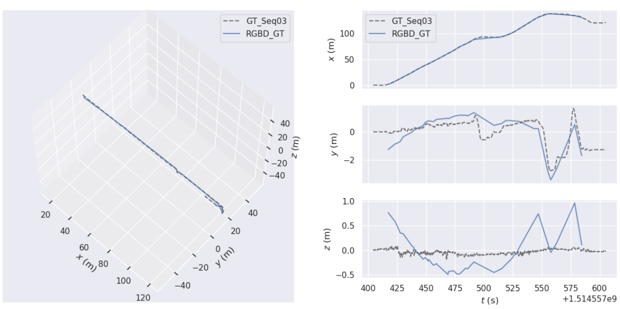
Figure 2. 3D trajectory estimated from **Monocular** OpenVSLAM.



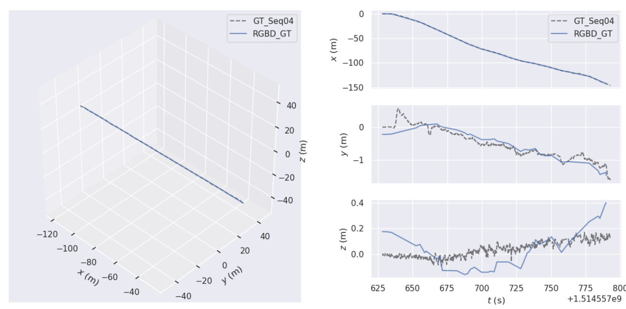
(a) Seq. 01



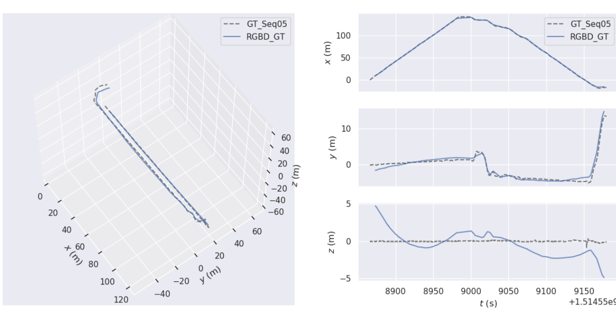
(b) Seq. 02



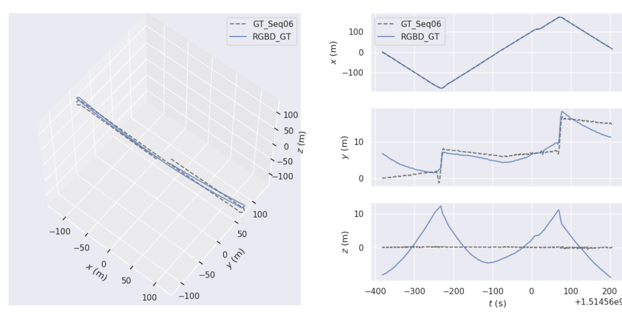
(c) Seq. 03



(d) Seq. 04

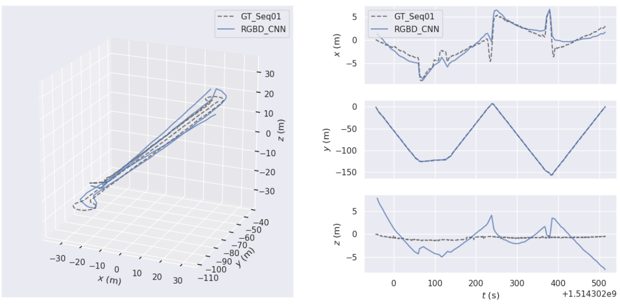


(e) Seq. 05

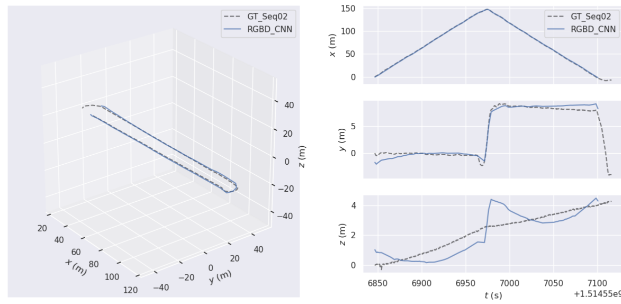


(f) Seq. 06

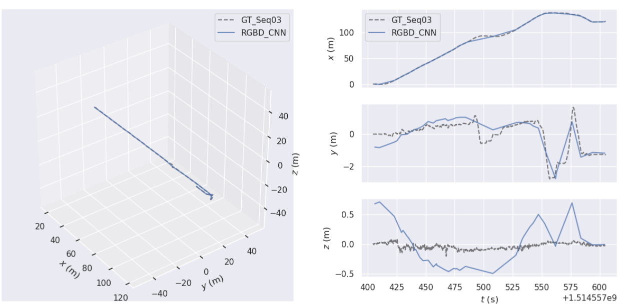
Figure 3. 3D trajectory estimated from **Mono+D**$_{GT}^{scale}$ OpenVSLAM.



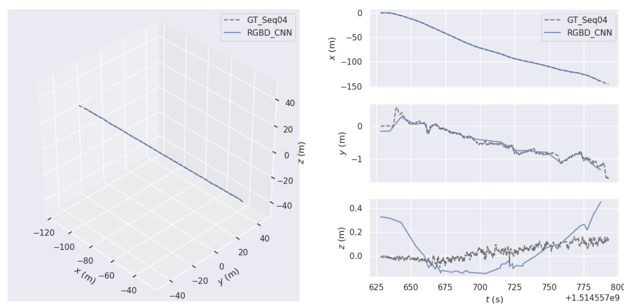
(a) Seq. 01



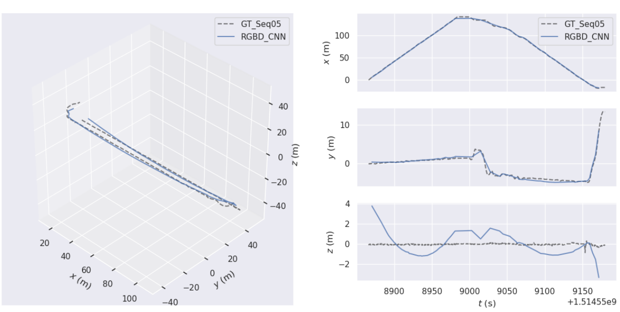
(b) Seq. 02



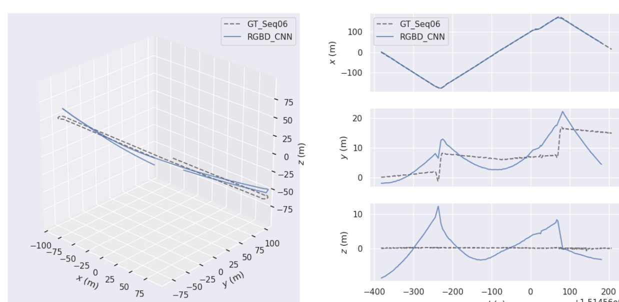
(c) Seq. 03



(d) Seq. 04



(e) Seq. 05



(f) Seq. 06

Figure 4. 3D trajectory estimated from **Mono+D$_{CNN}^{scale}$ (trained model MS*)** OpenVSLAM.

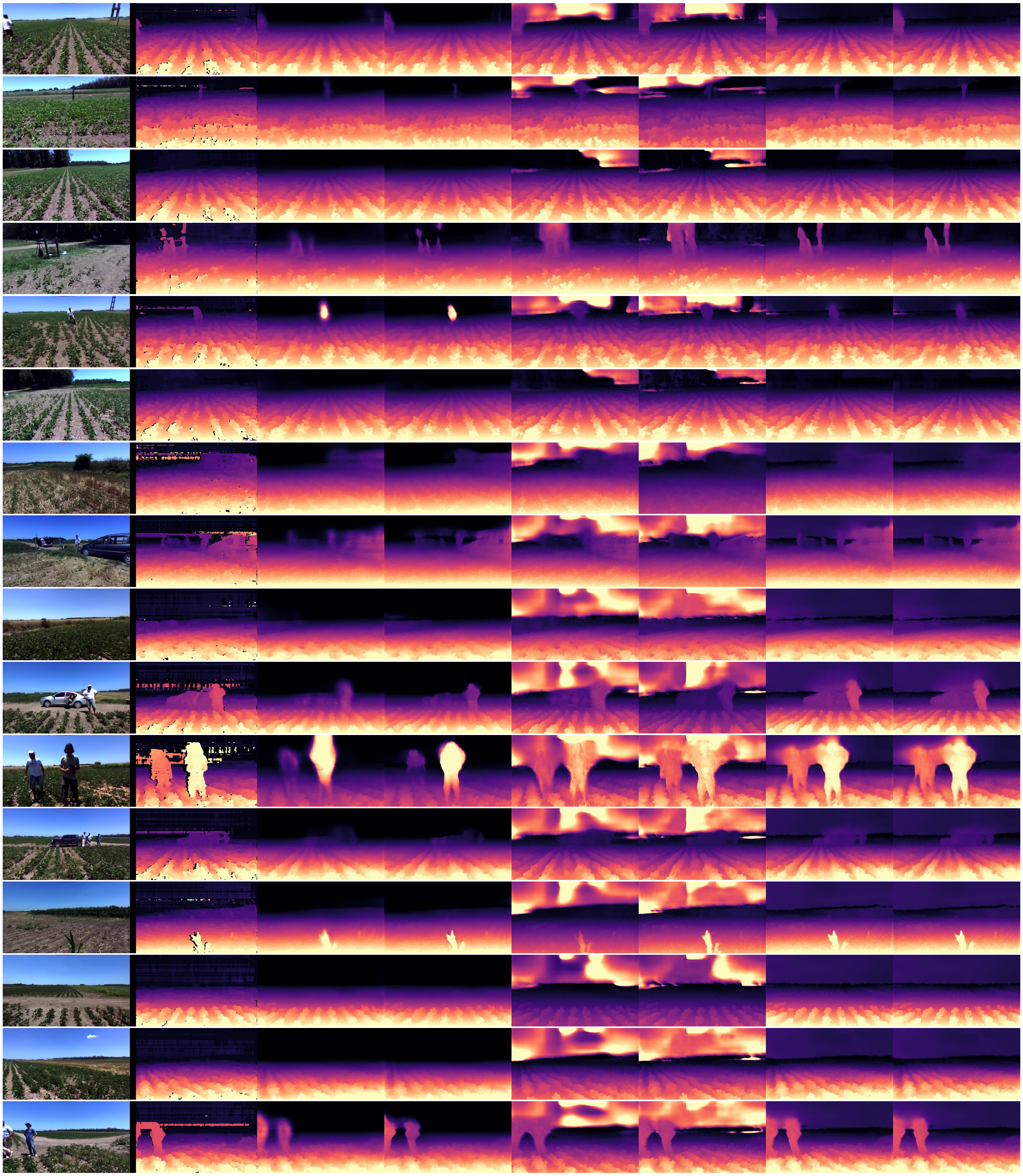


Figure 5. **Qualitative results** of self-supervised monocular depth estimation (Monodepth2) on Rosario, an illustration of estimated depth map using different training strategy. **Legend: S** - Self-supervised stereo supervision; **M** - Self-supervised mono supervision; ***** - start with model pretrained on KITTI [1], **SGBM** (ground truth) - semi-global batch matching [3].

References

[1] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. Vision meets robotics: The kitti dataset. *International Journal of Robotics Research (IJRR)*, 2013.

[2] Clément Godard, Oisin Mac Aodha, Michael Firman, and Gabriel J Brostow. Digging into self-supervised monocular depth estimation. In *Proceedings of the IEEE international conference on computer vision*, pages 3828–3838, 2019.

[3] Heiko Hirschmuller. Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on pattern analysis and machine intelligence*, 30(2):328–341, 2007.

[4] Taihú Pire, Martín Mujica, Javier Civera, and Ernesto Kofman. The rosario dataset: Multisensor data for localization and mapping in agricultural environments. *The International Journal of Robotics Research*, 0(0):1–9, 2019.