KP LABS
Mission Complete

Real-World Sentinel-2 Super-resolution Relying on Task-Driven Training

Maciej Ziaja[1,2], Paweł Kowaleczko[1,3], Michał Kawulok[1,2]

29.05.24

1. KP Labs
2. Silesian University of Technology
3. Warsaw University of Technology

info@kplabs.pl
www.kplabs.pl

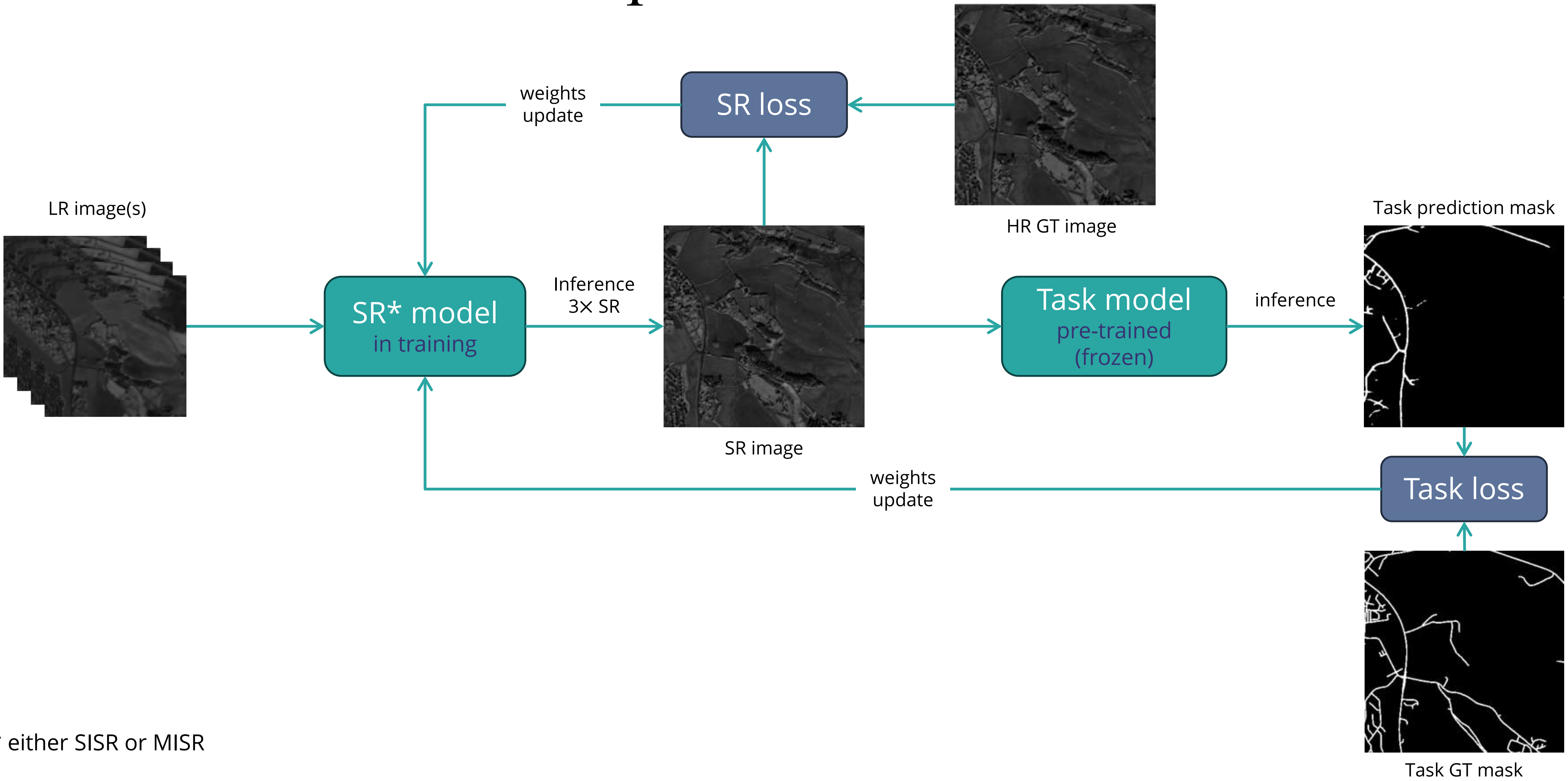
About us

KP Labs is a new space company in Gliwice, Poland

We create space grade hardware, software, and ML models for Earth observation and satellite telemetry.

www.kplabs.space

Task-driven super-resolution



* either SISR or MISR

Datasets

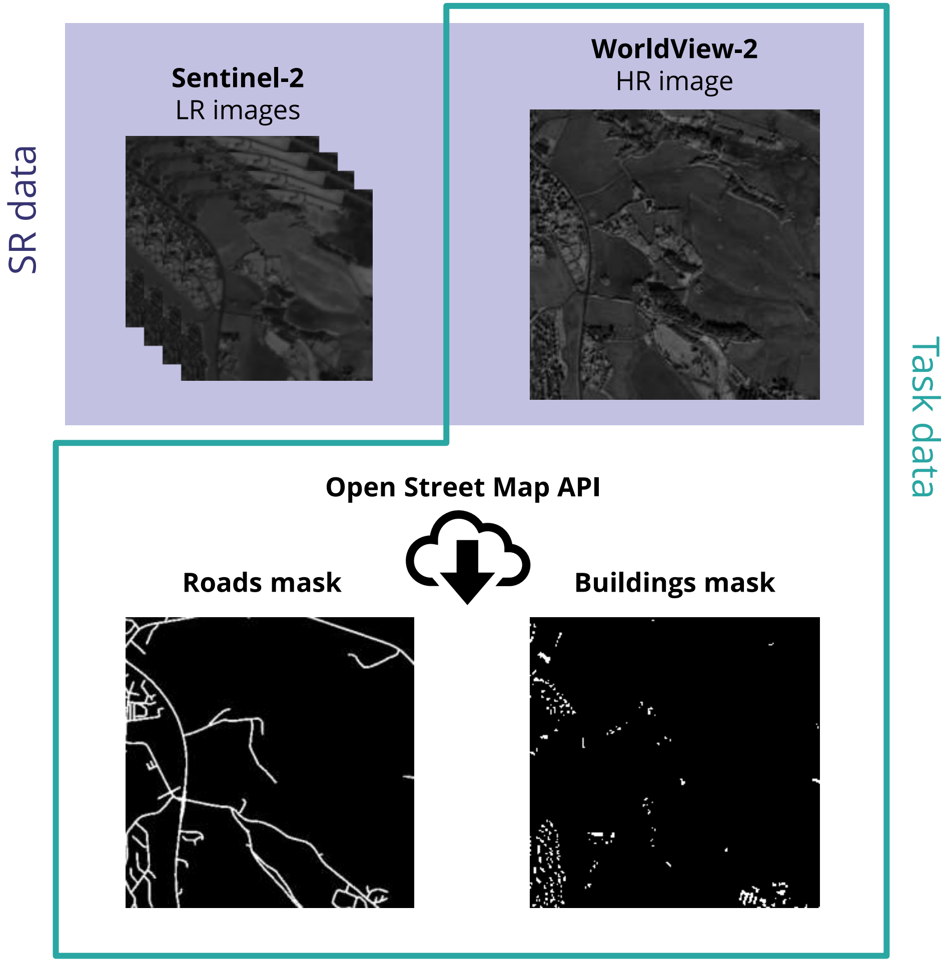
Compiling a dataset for task-driven SR training

Dataset features

- 90 scenes in total (72 training, 9 validation, 9 tests)
- Scenes with sizes over 1000 × 1000 pixels
- SR Data per scene (from MuS2 dataset):
  - Multiple LR Sentinel-2 images
  - HR WorldView-2 image (downsized)
- Multiple overlapping S-2 and WV-2 modalities but we mainly work with NIR
- Task-oriented data per scene projected onto WV-2 (from Open Street Map):
  - Buildings segmentation masks
  - Roads segmentation masks

Dataset versions

- Real-world data (WV-2 and S-2 images) → challenging dataset, low temporal consistency
- Simulated data (WV-2 and LR images simulated from WV-2 using S-2 PSF and downsampling) → easier dataset, high temporal consistency



Datasets

Visual preview of the compiled dataset

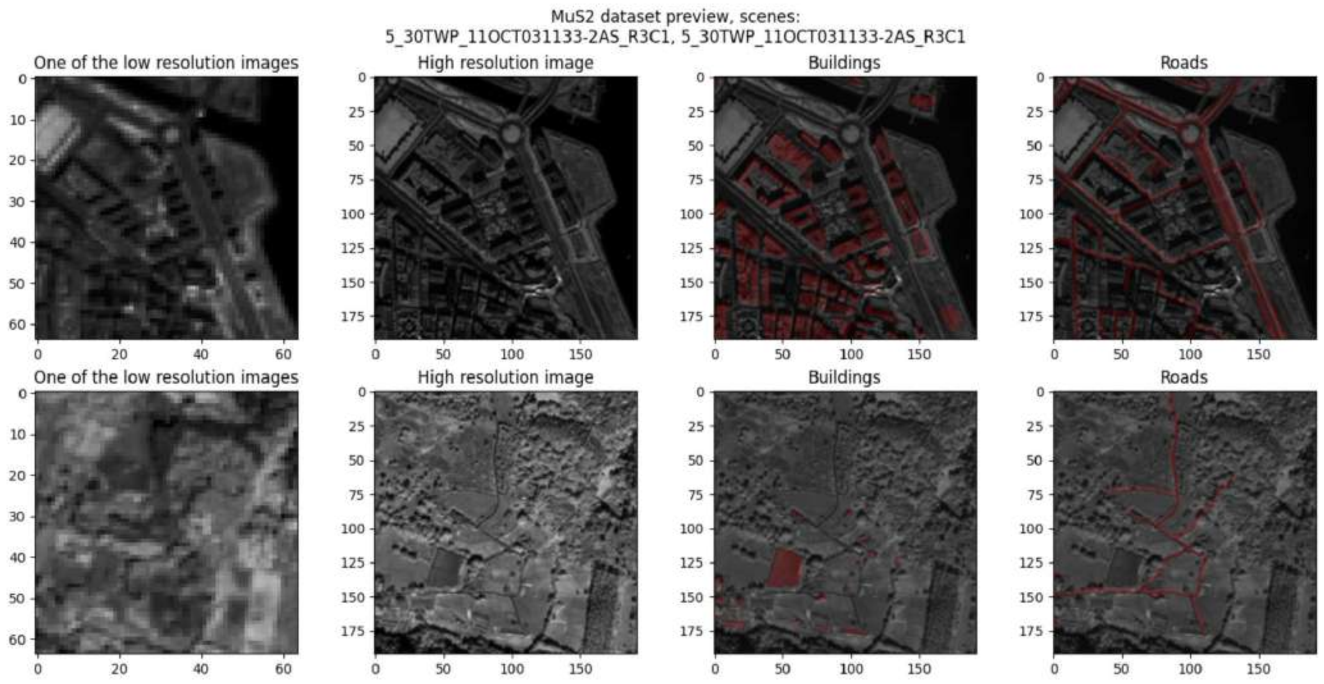


Figure 10: Sample scene fragments from the MuS2 SR dataset adapted to the requirements of task-driven training.

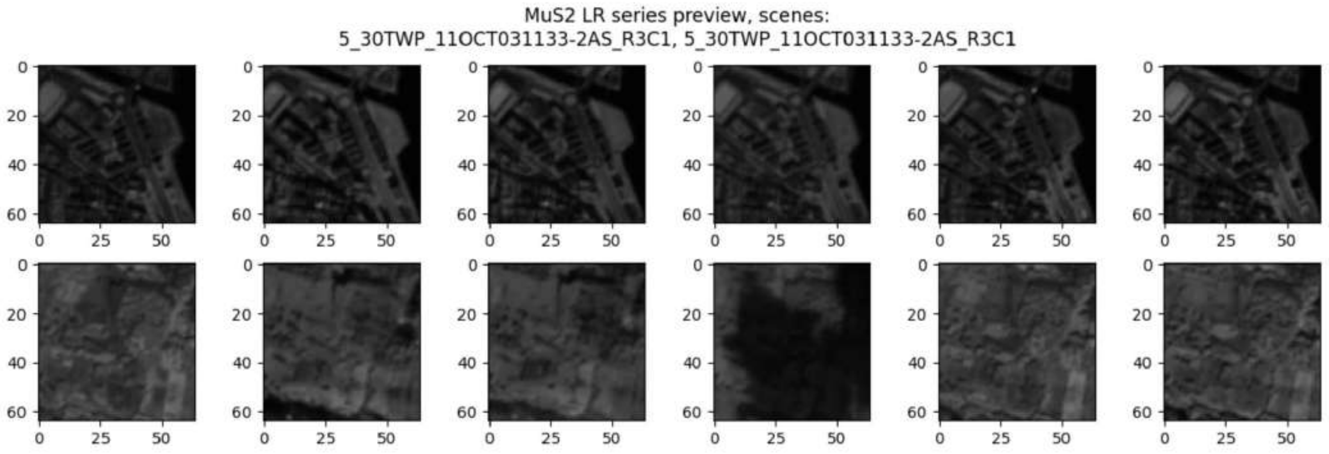


Figure 11: Sample series of multiple LR imagery fragments from the MuS2 SR dataset adapted to the requirements of task-driven training.

Training task models

For the future task-driven SR trainings

- Task scenarios: roads and buildings segmentation
- We train on the demonstrated dataset with consistent train/val/test split in all (trainings tasks and SR)
- We use Unet++ architecture
- Training in patches, evaluation on complete scenes
- Dice loss (1 – dice score)

Test metrics

Model	Dice coefficient	Accuracy	Precision	Recall
Buildings	0.491	0.958	0.499	0.507
Roads	0.465	0.947	0.429	0.511

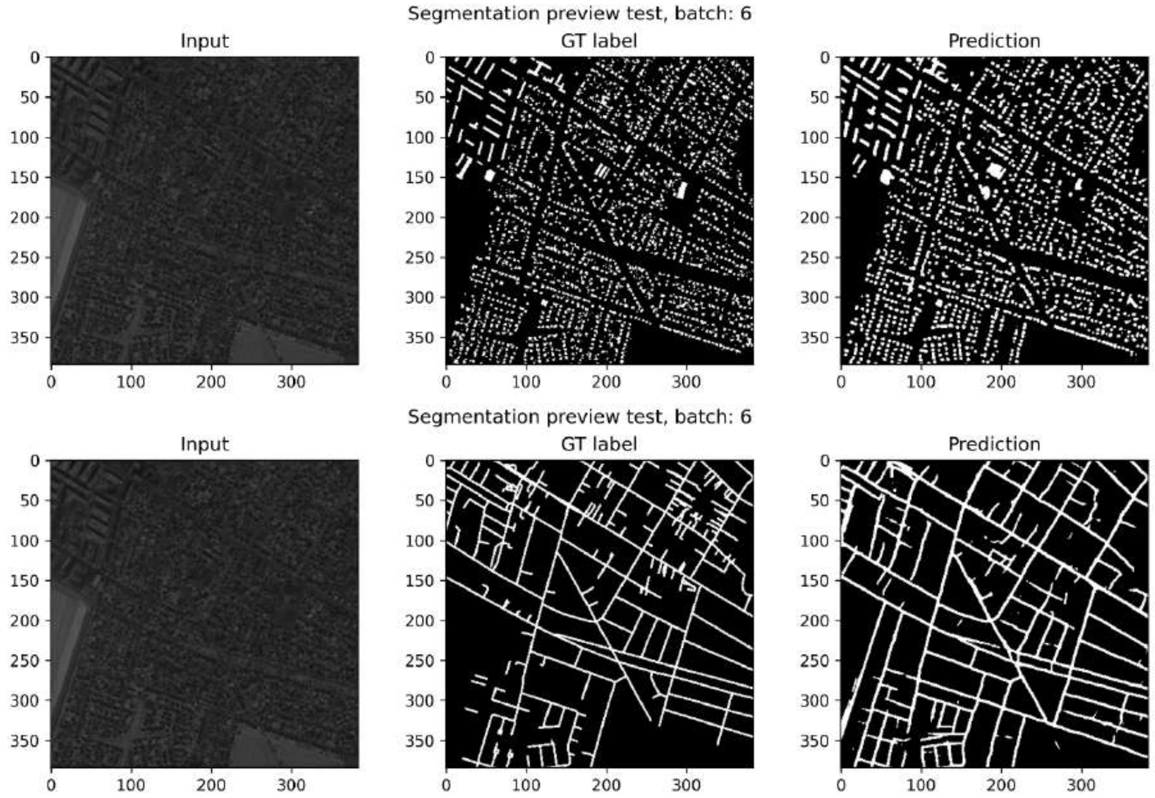


Figure 1: Exemplary segmentation networks results for buildings (top row) and roads (second row) obtained from the HR image from MuS2 dataset. The first column contains input images, the second the ground-truth (GT) segmentation masks (obtained via OSM) and the last one the predicted segmentation masks.

Task-driven SR trainigs

Considerations and goals for the experiments

Aspects to consider in the trainings:

- Establish SR baseline (conventional training with cL1 loss)
- Try to train SR with task-driven loss only (unlikely to succeed)
- Introduce training with cL1 and a single task-driven loss weighted:
  - Investigate how a single task-driven loss training (e.g. buildings segmentation) impacts segmentation results for a different task
  - How to weight multiple losses (e.g., cL1 and segmentation dice?)
    - Static weighting (weighted sum with fixed weights)
    - Dynamic weighting (fixed proportion between losses, weights updated on each epoch end to keep the given proportion in regard to a reference loss)
- Train with conventional SR loss and multiple task-driven losses
- Utilize segmentation information for patches selection (prioritize training patches with roads & buldings presence)
- Try fine-tuning options (e.g., train with traditional loss, fine-tune with task-driven)
- Investigate training on simulated (temporally consistent) vs real-world data (less temporally consistent)
- Compare MISR (RAMS) vs SISR (HAT) networks with task-driven scenarios

Experiments and results for MISR

Table 3: Results of task-driven multi-image super-resolution trainings. Metrics in columns are color coded with gradient from worst highlighted as red, to best indicated with green; color-coding for simulated and real-world data is done separately. B delineates building segmentation oriented losses, R indicates the roads segmentation ones. Notation $W(l)$ designates weight of given loss l per experiment. Experiment #2 fails to learn and minimize the loss, we don't report the numeric values for this training.

Id	Loss	Patches selection	Fine-tune from	Epochs	Test results: cPSNR↑	cSSIM↑	LPIPS↓	Bldg. Dice↑	Rd. Dice↑
Simulated data									
1	SR^{cL1}	N/A		500	32.58	0.856	0.286	0.157	0.278
2	B^{Dice}	No		500					
3	Static $W(SR^{cL1})=1, W(B^{Dice})=1$	No		500	30.93	0.777	0.259	0.377	0.276
4	Dynamic $W(SR^{cL1})=1, W(B^{Dice})=1$	No		500	31.37	0.797	0.269	0.382	0.293
5	Dynamic $W(SR^{cL1})=1, W(B^{Dice})=0.50, W(R^{Dice})=0.50$	No		500	31.69	0.807	0.244	0.391	0.385
6	Dynamic $W(SR^{cL1})=1, W(B^{Dice})=0.25, W(R^{Dice})=0.25$	No		500	32.09	0.833	0.245	0.391	0.390
7	Dynamic $W(SR^{cL1})=1, W(B^{Dice})=0.25, W(R^{Dice})=0.25$	Yes		500	32.09	0.840	0.247	0.392	0.380
8	SR^{cL1}	N/A	1	125	32.47	0.857	0.278	0.177	0.284
9	Dynamic $W(SR^{cL1})=1, W(B^{Dice})=0.25, W(R^{Dice})=0.25$	Yes	1	125	32.39	0.850	0.233	0.374	0.370
10	Dynamic $W(Consistency^{cL1})=1, W(B^{Dice})=0.25, W(R^{Dice})=0.25$	Yes	1	125	29.79	0.711	0.403	0.285	0.265
Real-world data									
1*	Test only from 1	N/A	N/A	N/A	24.44	0.636	0.423	0.021	0.139
7*	Test only from 7	N/A	N/A	N/A	24.33	0.624	0.369	0.362	0.339
11	SR^{cL1}	N/A		125	24.35	0.636	0.491	0.001	0.065
12	Dynamic $W(SR^{cL1})=1, W(B^{Dice})=1$	No		125	24.55	0.622	0.398	0.450	0.278
13	Dynamic $W(SR^{cL1})=1, W(B^{Dice})=0.50, W(R^{Dice})=0.50$	No		125	24.55	0.620	0.365	0.430	0.427
14	Dynamic $W(SR^{cL1})=1, W(B^{Dice})=0.25, W(R^{Dice})=0.25$	No		125	24.69	0.630	0.380	0.411	0.431
15	Dynamic $W(SR^{cL1})=1, W(B^{Dice})=0.25, W(R^{Dice})=0.25$	Yes		125	24.44	0.631	0.369	0.450	0.450
16	SR^{cL1}	N/A	11	125	24.44	0.636	0.484	0.002	0.066
17	$B^{Dice}+R^{Dice}$	Yes	11	125	22.85	0.552	0.370	0.453	0.437
18	Dynamic $W(SR^{cL1})=1, W(B^{Dice})=0.25, W(R^{Dice})=0.25$	Yes	11	125	24.25	0.622	0.384	0.415	0.401
19	Dynamic $W(Consistency^{cL1})=1, W(B^{Dice})=0.25, W(R^{Dice})=0.25$	Yes	11	125	23.42	0.519	0.584	0.296	0.248
20	SR^{cL1}	N/A	1	125	24.84	0.646	0.474	0.006	0.071
21	$B^{Dice}+R^{Dice}$	Yes	1	125	23.35	0.564	0.378	0.482	0.467
22	Dynamic $W(SR^{cL1})=1, W(B^{Dice})=0.25, W(R^{Dice})=0.25$	Yes	1	125	24.74	0.637	0.363	0.468	0.450
23	Dynamic $W(Consistency^{cL1})=1, W(B^{Dice})=0.25, W(R^{Dice})=0.25$	Yes	1	125	23.94	0.558	0.528	0.296	0.273

KP LABS
Mission Complete

Experiments and results for SISR

Table 4: Results of task-driven single-image super-resolution trainings. Color coding is done for each metrics separately for all experiments.

Id	Loss	Patches selection	Fine-tune from	Epochs	Test results cPSNR↑	cSSIM↑	LPIPS↓	Bldg. Dice↑	Rd. Dice↑
Simulated data									
1-S	SR^{cL1}	N/A		50	33.42	0.839	0.297	0.081	0.140
7-S	Dynamic $W(SR^{cL1}) = 1$, $W(B^{Dice}) = 0.25$, $W(R^{Dice}) = 0.25$	Yes		50	32.66	0.817	0.267	0.263	0.302
Real-world data									
11-S	SR^{cL1}	N/A		50	26.15	0.643	0.471	0.006	0.011
15-S	Dynamic $W(SR^{cL1}) = 1$, $W(B^{Dice}) = 0.25$, $W(R^{Dice}) = 0.25$	Yes		50	25.79	0.630	0.343	0.242	0.259

Visual results

Task-driven trainings

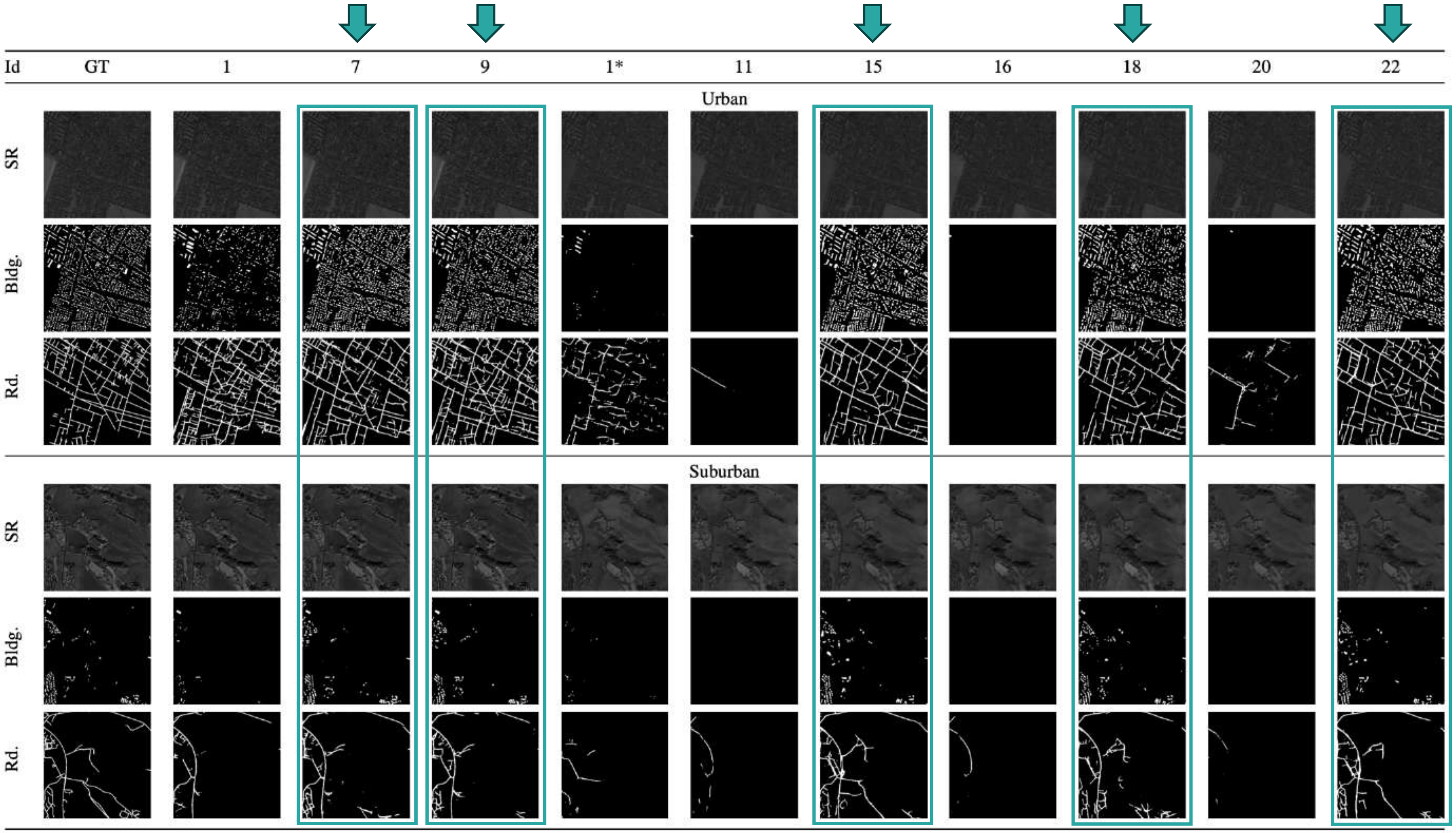
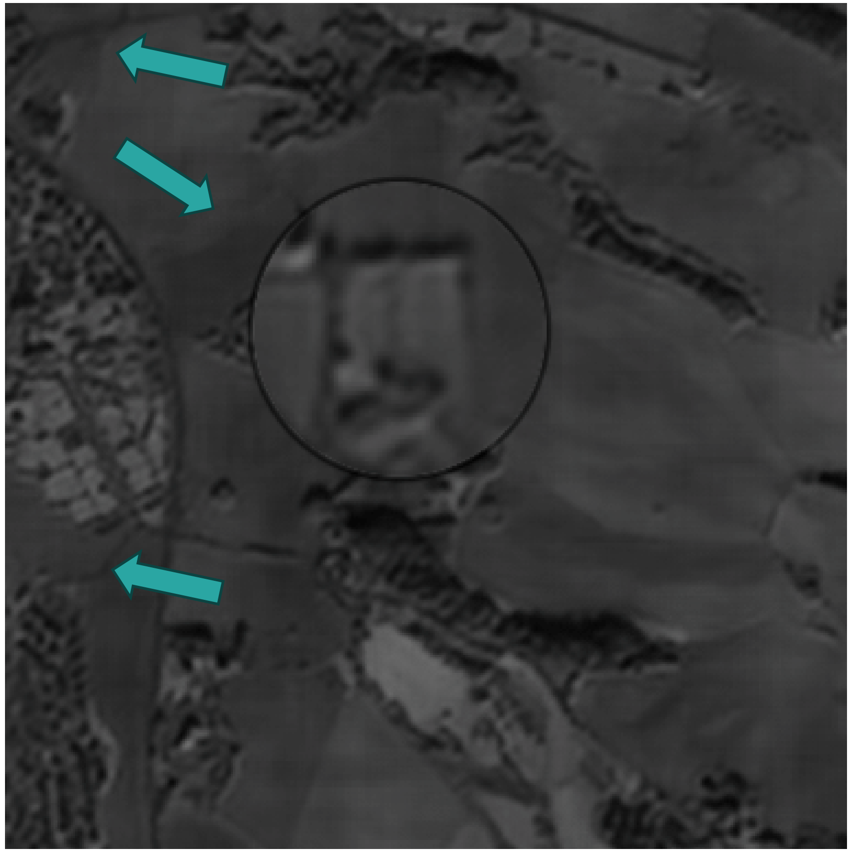


Figure 12: Visual preview of multi-image super-resolution and segmentation results on MuS2. The images are close-ups of test scenes, cropped to 384×384 patches.

Visual results

Close-up, baseline vs task-driven RAMS training on real-world data

11 (baseline)



15 (task-driven)

KP LABS
Mission Complete

Results

Perspectives on task-driven training and evaluation of SR networks

- Conventional SR networks may not produce SR results sufficient for further processing out-of-the box
- Task-driven trainings improve segmentation results on test data a lot
- Various tasks seem to improve congruently with task-driven training
- Methods like dynamic loss weighting and patches selection improve results
- Works for SISR and MISR, MISR seems to benefit more from task-driven trainings
- Task-driven trainings make more significant impact when training on challenging real-world data
- Task-driven trainings lead to more distinct man-made structures in real-life data
- Methods to be expanded and developed further, especially in the context of foundational models (both SR and task ones)

Thank you

www.kplabs.space