

Reasonable Non-generative Super-Resolution and Self-Supervision

Jérémy Anger^{1,2}, Ngoc Long Nguyen², Lara Raad²,
Gabriele Facciolo², Thibaud Ehret², Axel Davy², Pablo Arias²

¹Kayrros SAS

²Centre Borelli, ENS Paris-Saclay, France

Introduction: image restoration and well-posedness

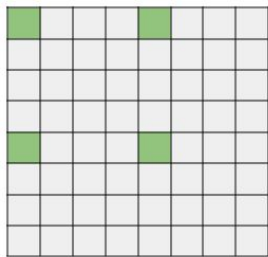
Image restoration problems are often characterized by how ill-posed they are

ill-posed

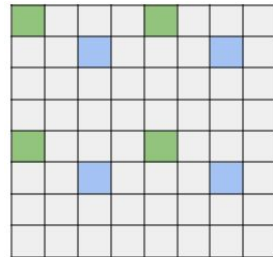
well-posed

→ results driven by regularization / priors

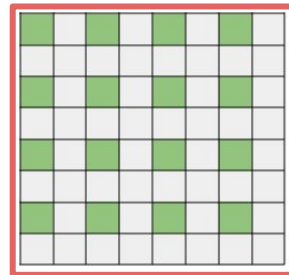
→ results driven by data, accurate



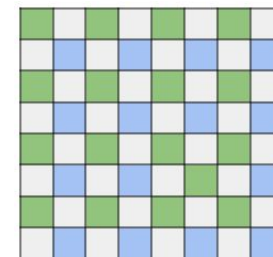
x4
single-frame



x4
multi-frame



x2
single-frame



x2
multi-frame

Introduction: reasonable super-resolution

For single-image super-resolution of 10m Sentinel-2, what can we do?

- If the user wants to recover real details: $\leq x2$ SR
- If the user wants to generate a plausible image: $> x2$ SR

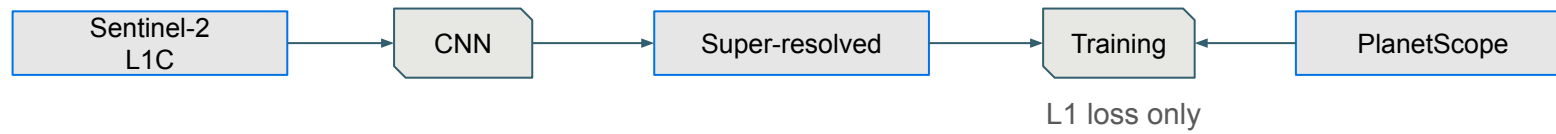
The choice depend on the end-goal.

We study accurate super-resolution:

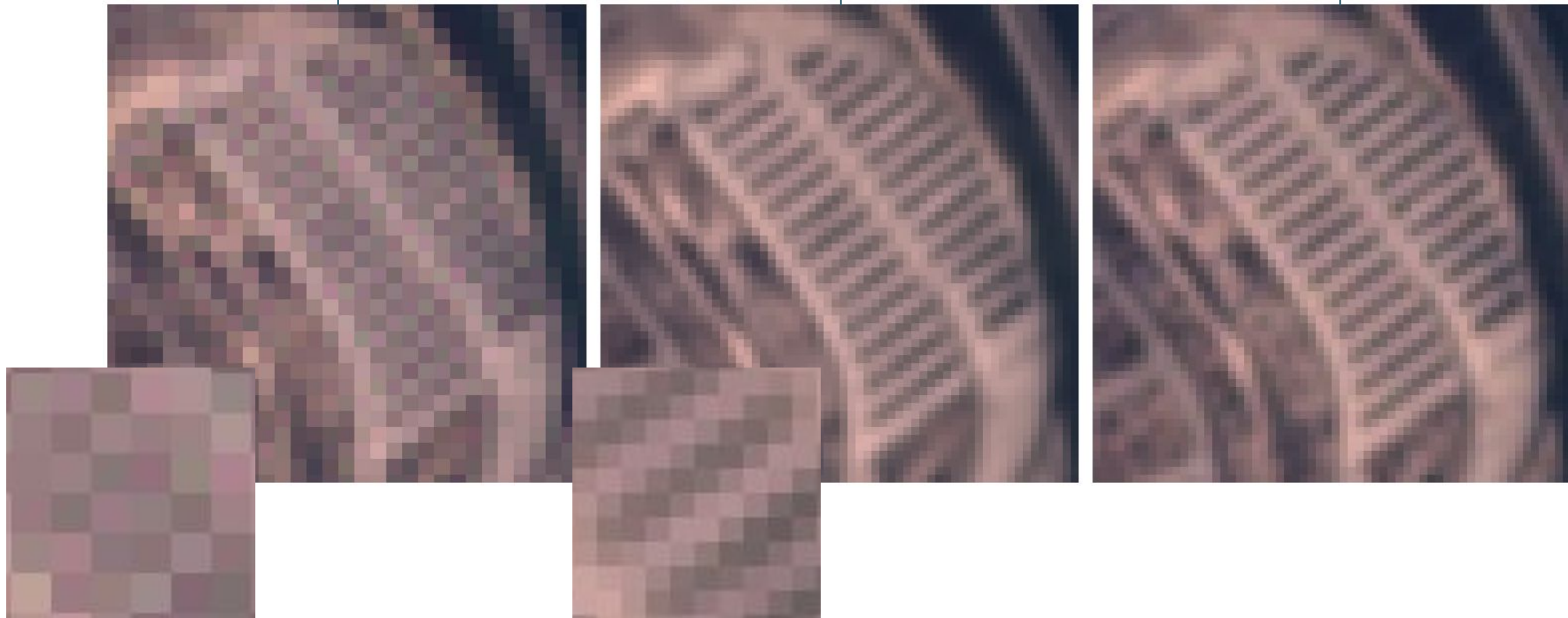
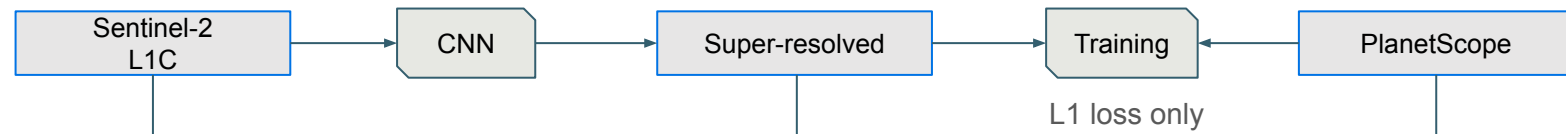
Why is Sentinel-2 adapted for non-generative SISR?

How to train without ground-truth?

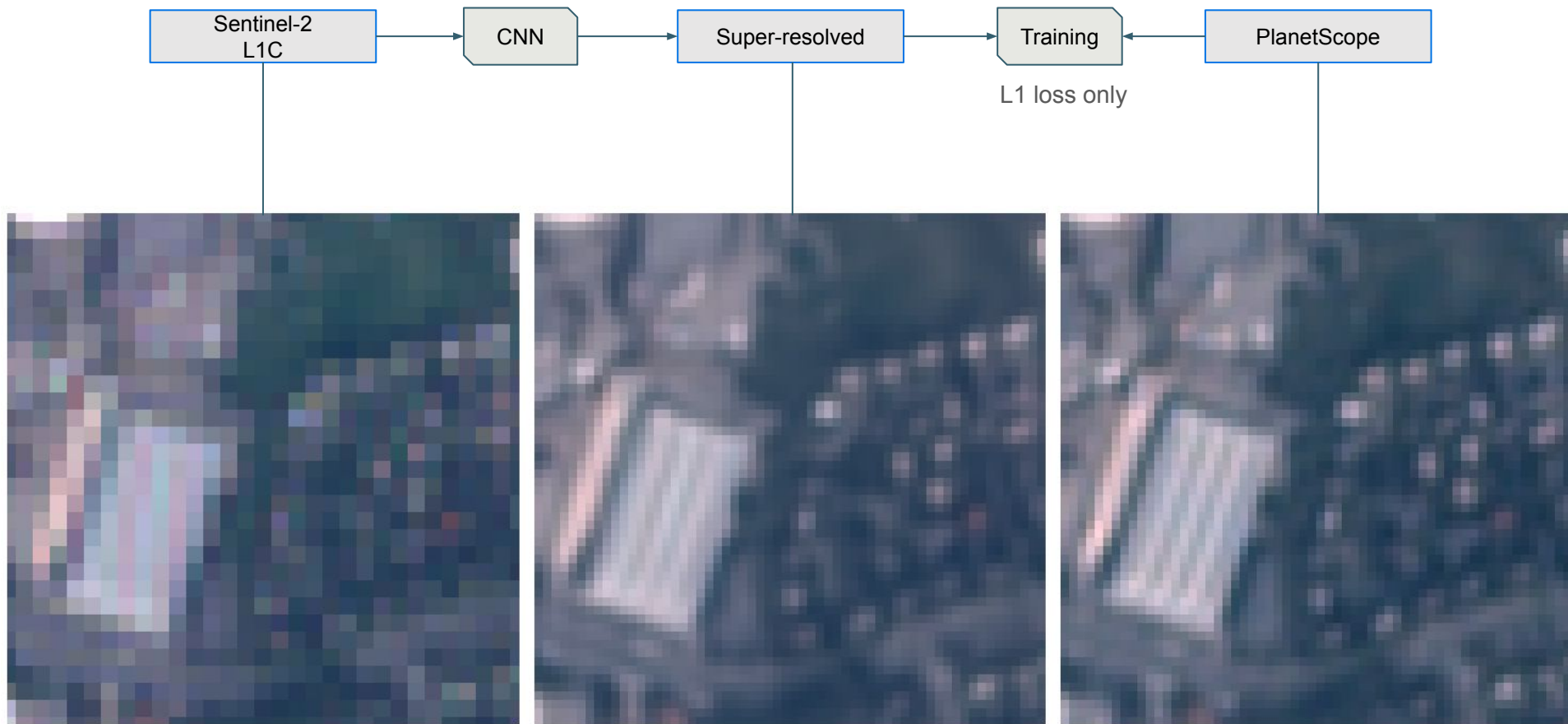
Sentinel-2 super-resolution



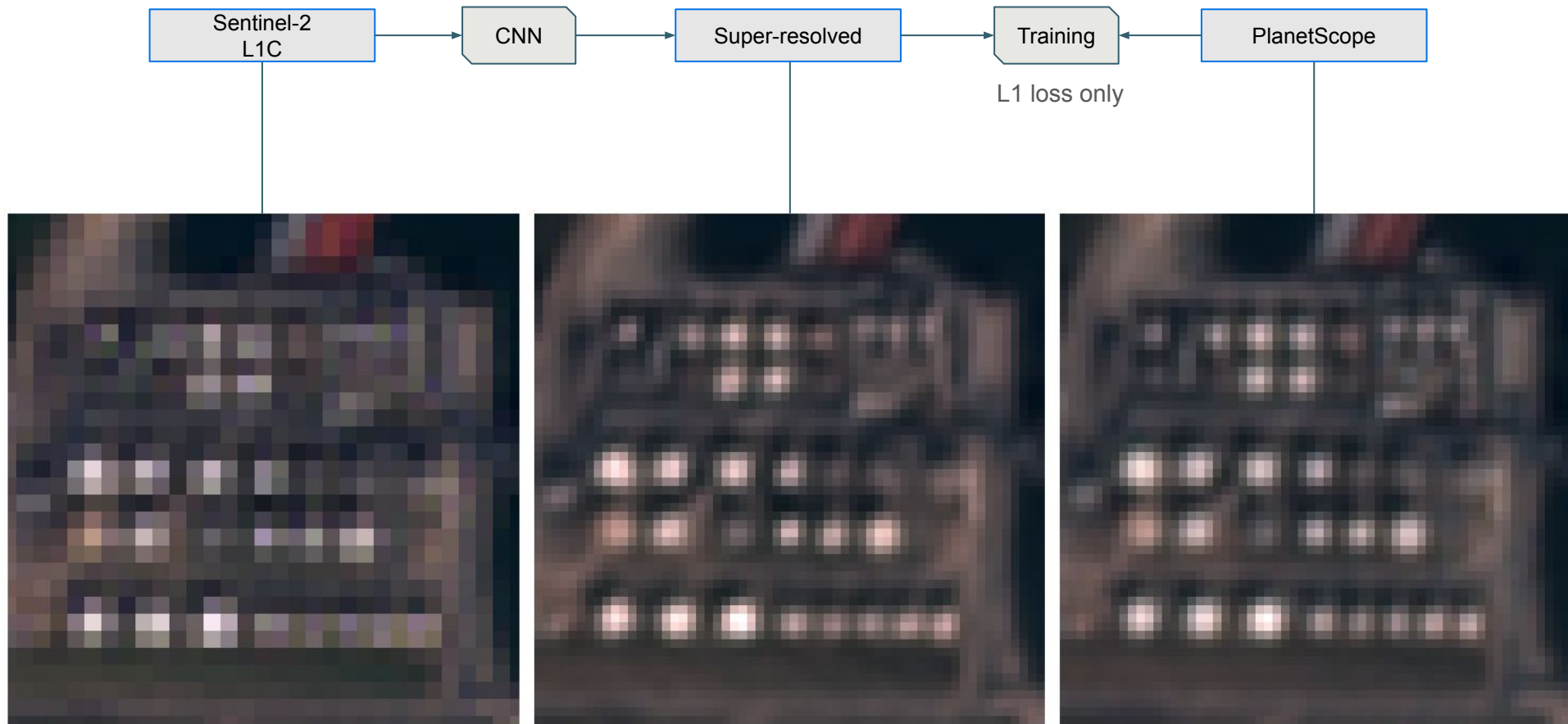
Sentinel-2 super-resolution: results with PlanetScope



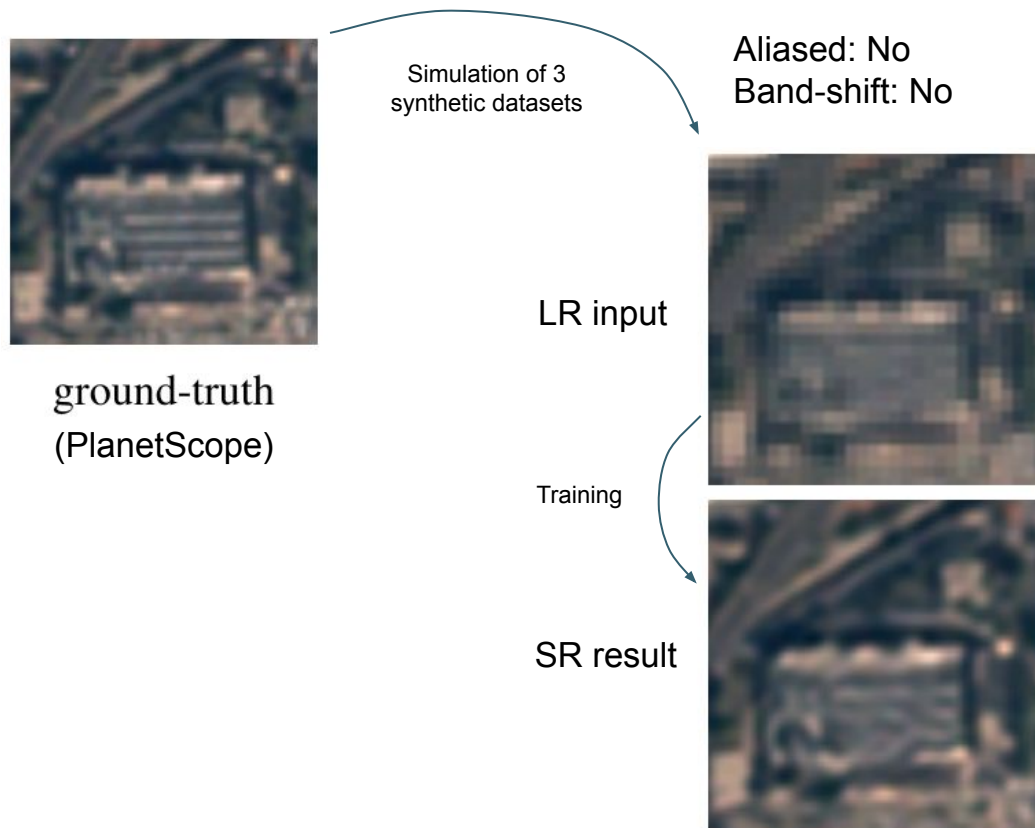
Sentinel-2 super-resolution: results with PlanetScope



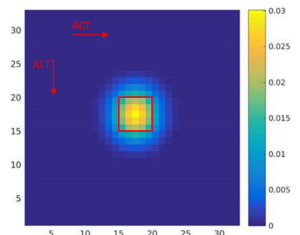
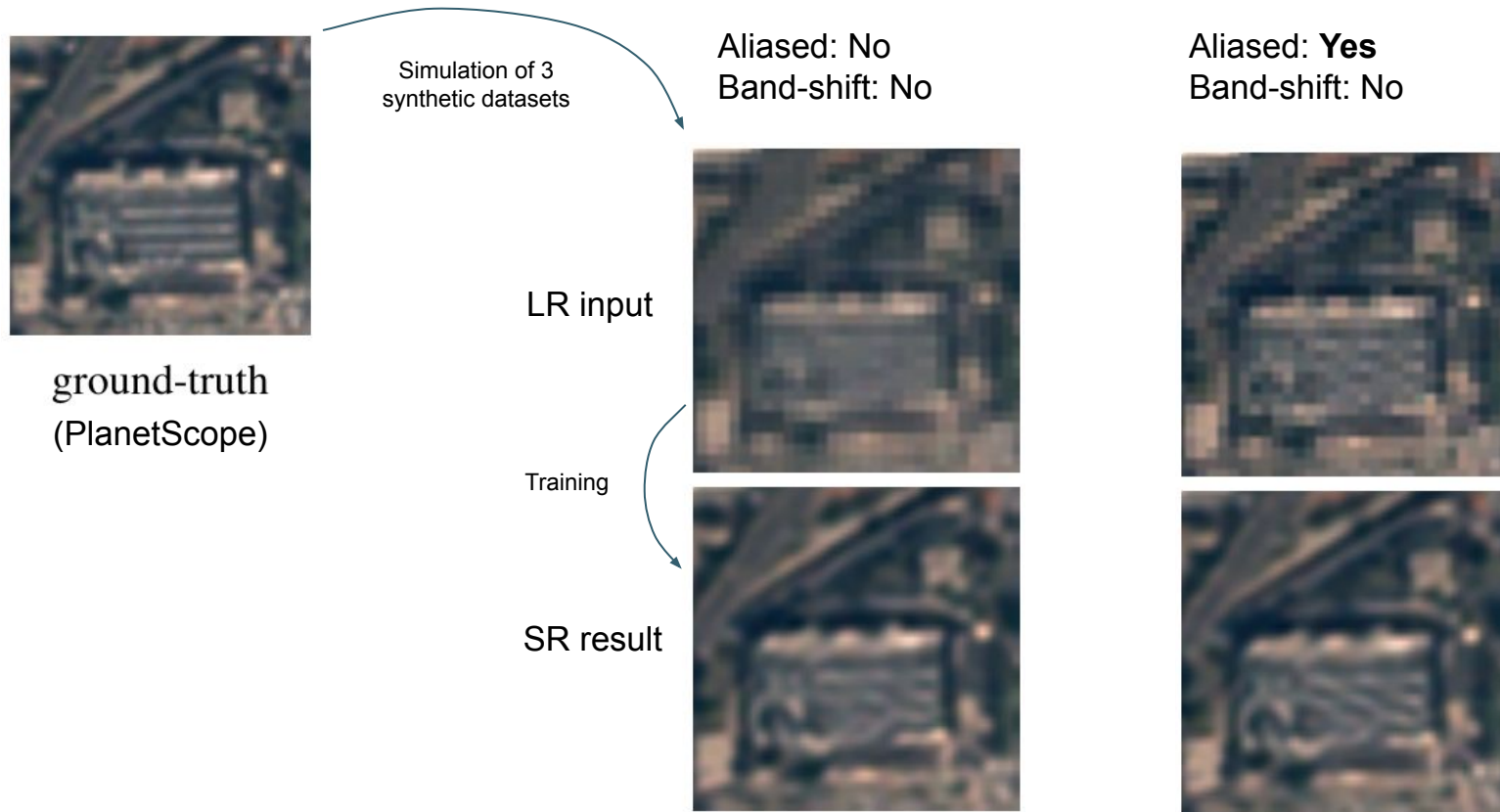
Sentinel-2 super-resolution: results with PlanetScope



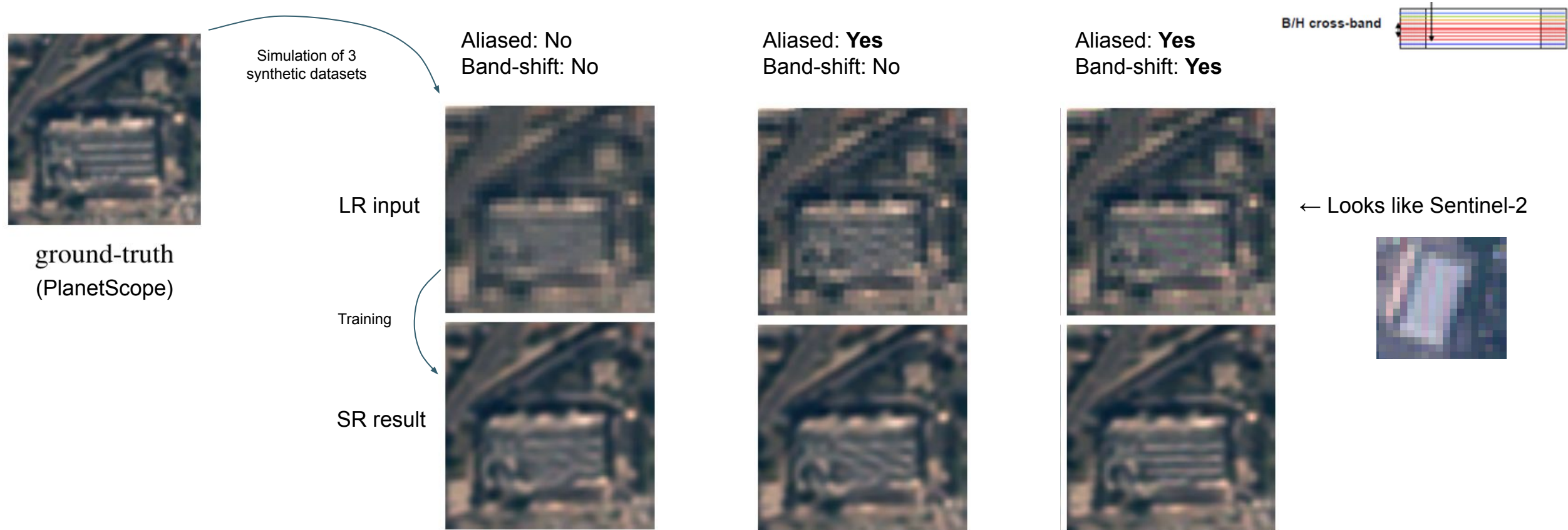
Sentinel-2 super-resolution: impact alias and band shift



Sentinel-2 super-resolution: impact alias and band shift



Sentinel-2 super-resolution: impact alias and band shift



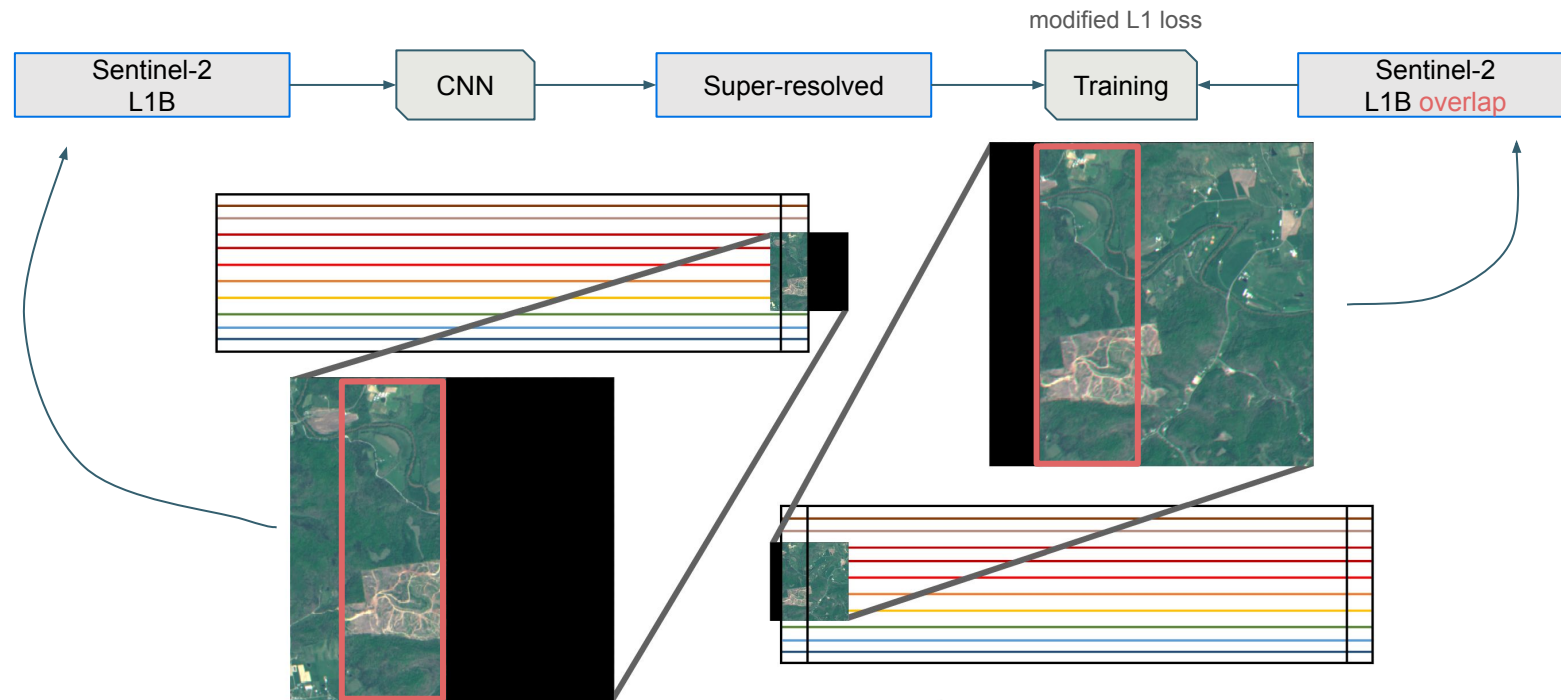
This synthetic experiment shows that alias and band-shift are important for SISR

→ Sentinel-2 MSI is well adapted for modest but accurate SISR!

Self-supervised super-resolution

PlanetScope is not a perfect ground-truth. Sen2Venµs is better.

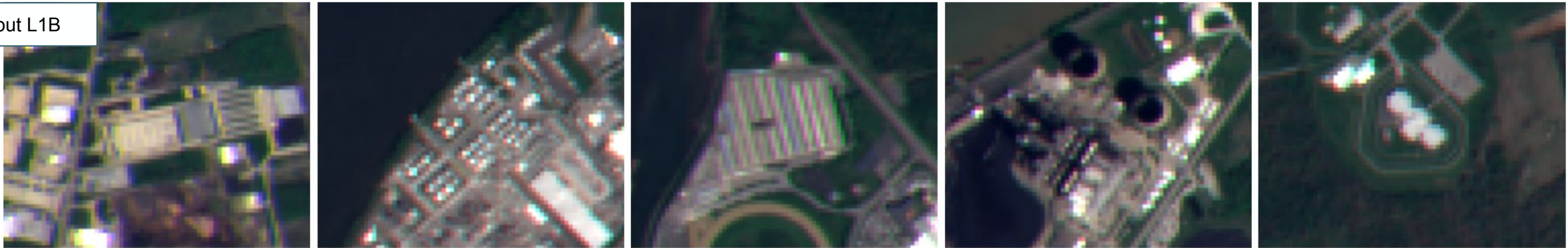
But L1B products offers us new opportunities: Self-supervision on overlaps
→ training x2 SR with only 10m bands as target



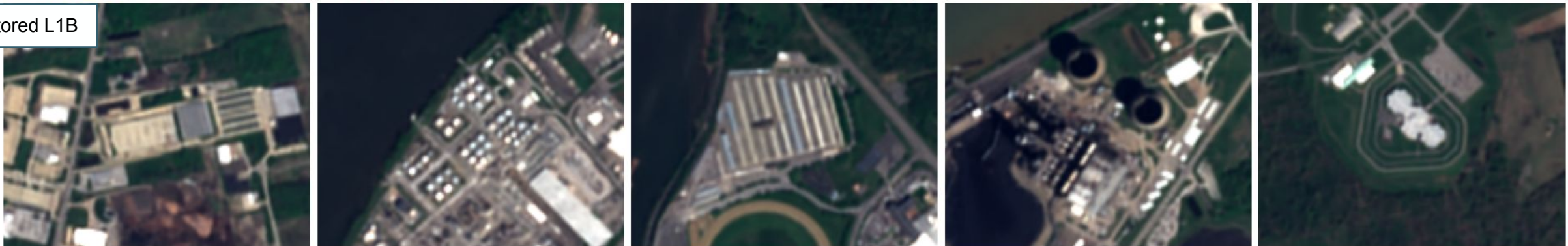
Self-supervised super-resolution

The networks learns to register the bands, super-resolve, denoise, sharpen...

Input L1B



Restored L1B



Conclusion

We have to discriminate SR methods into two categories:

- restore real details in order to extract information
 - help for downstream tasks and human interpretation
 - the output quality is driven by the instrument design
 - is limited by the input characteristics (optical system, SNR, etc)
typically it does not help to go beyond x2
- improve the visual aspect by generating plausible imagery
 - beautification, prone to hallucination
 - the output is driven by strong priors
 - is limited by the network capability (architecture, training procedure, target distribution)
we can do x20 super-resolution

Non-generative super-resolution is achievable for Sentinel-2 image, thanks to unique characteristics of the instrument.

Self-supervision (on L1B) allows to train SR without ground-truth nor simulation.