IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Convolutional Neural Network Modelling for Land Surface Temperature Super Resolution

ML Track – Topics in Machine Learning: ML for Earth Observation

Land Surface Temperature Super-Resolution with a Scale-Invariance-Free Neural Approach: Application to MODIS
Ait-Bachir, Romuald et al. (2025?)

Romain LETELLIER – Massina AMMAD

May 2025

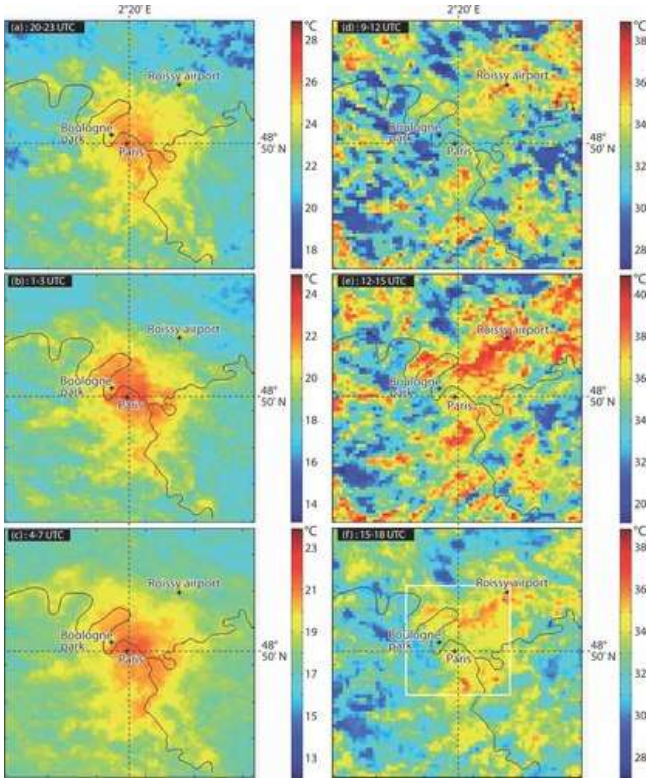
Outline

Context (1/3)

Satellite Monitoring Benefits

Satellite data help:

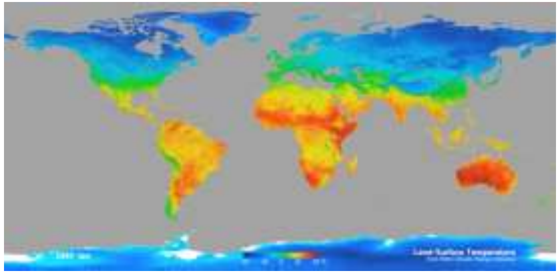
- identify high-risk urban areas,
- informing heat-health warning systems,
- guiding urban planning…

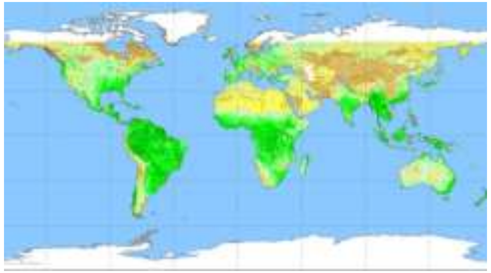


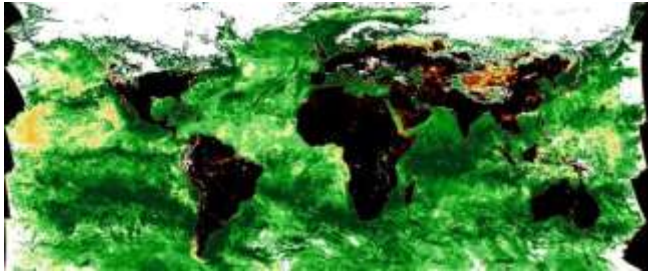
Average LST infrared images (see text) from 4 to 13 August 2003, for each of the diurnal time intervals shown in Figure 2. The colour scale (in degrees Celsius) is optimally enhanced separately for each image. The white square represents the enlarged area of Figure 4(a)

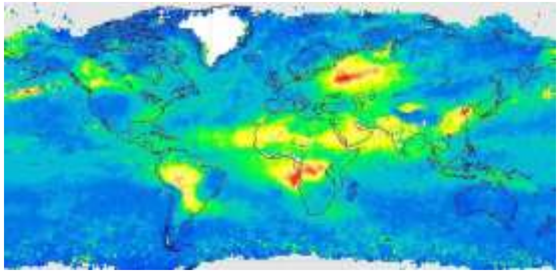
Source: B. Dousset, F. Gourmelon, K. Laaidi, A. Zeghnoun, E. Giraudet, P. Bretin, E. Mauri, and S. Vandentorren, "Satellite monitoring of summer heat waves in the Paris metropolitan area," *International Journal of Climatology*, vol. 31, no. 2, pp. 313–323, Feb. 2011.

Context (2/3)


LST


NDVI


NDWI


AOD

LAI

Context (3/3)

They can't be produced as often / with a high enough quality due to their sensitivity:

Image Type	Sensitive To
NDVI	None
NDWI	Clouds, Atmospheric Humidity, Dry Soil
LST	Clouds, Water Vapor
AOD	Clouds, Ground Reflectance
LAI	Clouds, Aerosol

Purpose of the project (1/4)

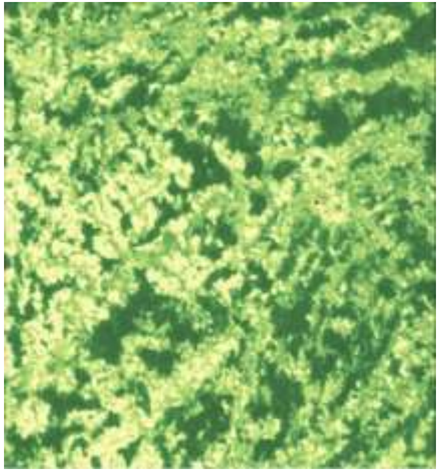
MODIS satellite provides us with images at a **good frequency** and **bad resolution** (LST → 1km * 1km)

ASTER satellite provides us with good resolution but very few images (1 per 16 days)

https://fr.wikipedia.org/wiki/Terra_%28satellite%29

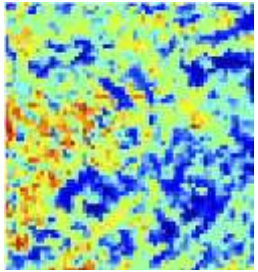
Purpose of the project (2/4)

⇒ Let's use DL to enhance the quality of these images!



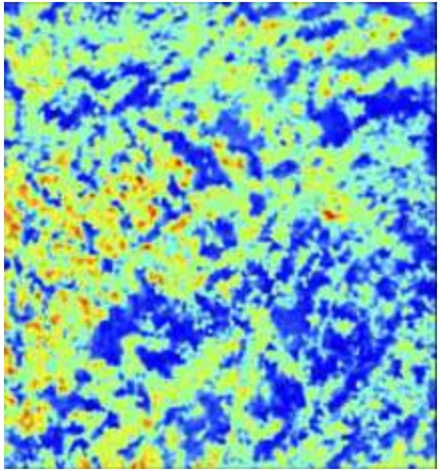
*NDVI 250m * 250m*

+



*LST 1km * 1km*

⇒

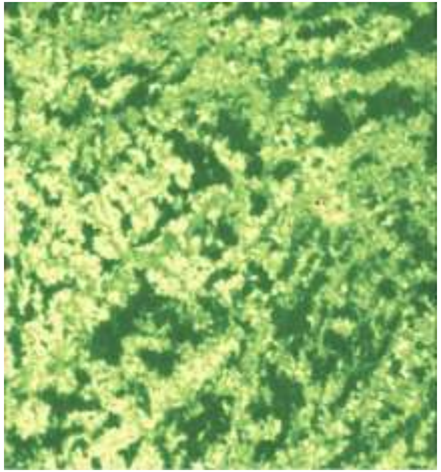


*LST 250m * 250m*

Source: *Land Surface Temperature Super-Resolution with a Scale-Invariance-Free Neural Approach: Application to MODIS*. Ait-Bachir, Romuald et al.

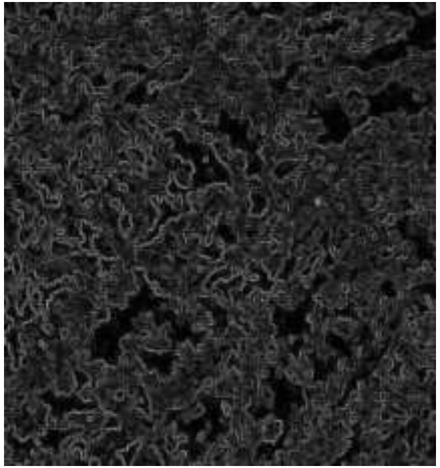
Purpose of the project (3/4)

We consider a correlation between the gradient of the NDVI and the gradient of the LST ⇒ the textures are equivalent.



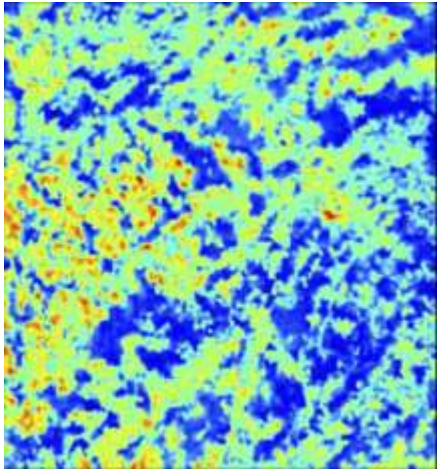
*NDVI 250m * 250m*

⇔



*NDVI gradient 250m * 250m*

⇔



*LST 250m * 250m*

Source: *Land Surface Temperature Super-Resolution with a Scale-Invariance-Free Neural Approach: Application to MODIS*. Ait-Bachir, Romuald et al.

Purpose of the project (4/4)

State of art

TsHARP improves spatial resolution in remote sensing data.

Approach

Combining multi-index data and gradient correlation.

Objective

Boost accuracy and applicability in Earth observation

Source: *Land Surface Temperature Super-Resolution with a Scale-Invariance-Free Neural Approach: Application to MODIS*. Ait-Bachir, Romuald et al – Farshid Sattari, Mazlan Hashim, Mehdi Sookhak, Saeed Banihashemi, Amin Beiranvand Pour, “Assessment of the TsHARP method for spatial downscaling of land surface temperature over urban regions,” *Urban Climate*, Volume 45, 2022, 101265, ISSN 2212-0955, https://doi.org/10.1016/j.uclim.2022.101265.

Our plan (1/1)

DATA

Gather the data: Collect MODIS and ASTER datasets focusing on LST and NDVI.

NEURAL NETWORK

Create the neural network: design architecture exploiting cross-index correlations.

LOSS

Compute the loss: use gradient-based metrics for accurate texture recovery.

COMPARISON

Compare results: Benchmark enhanced images against existing methods.

Source: *Land Surface Temperature Super-Resolution with a Scale-Invariance-Free Neural Approach: Application to MODIS*. Ait-Bachir, Romuald et al

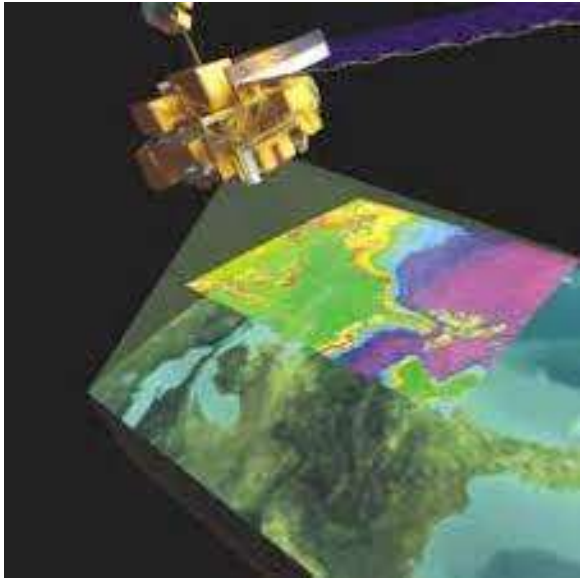
About the data (1/1)

⇒ MODIS Satellite (1-2 images per day)

- MOD21A1: LST | 1 km res
- MOD09GQ: RED & NIR (NDVI) | 250 m res

⇒ ASTER Satellite (1 images per 16 days)

- ASTER AST_08: LST | 90m



https://modis.gsfc.nasa.gov/about/media/modis_brochure.pdf

Pre-processing (1/1)

LST file – hdf format

Quality Control	Temperature Values	UTC
Cloud presence flags indicating image degradation.	Temperature Values	Timestamps for synchronization

Architecture of the SIF-CNN-SR1 (1/3)

	Synthesis	Consistency	Explanation
Classical statistical and ML models	✗	✓	• Depends on the heterogeneity of the studied surface • → Scale-Invariance hypothesis • The high resolution generated image doesn't always present the wished small scale textures from the high resolution concatenated image
SIF-CNN, SC-UNet	✓	✓	• Complexity relatively low • Commonly used for super-resolution applications • Respects the properties of consistency and synthesis (Scale-Invariance-Free)

Source: *Land Surface Temperature Super-Resolution with a Scale-Invariance-Free Neural Approach: Application to MODIS*. Ait-Bachir, Romuald et al Y. Chen, R. Xia, K. Yang, and K. Zou, "MICU: Image super-resolution via multi-level information compensation and U-net," Expert Systems with Applications, vol. 245, p. 123111, 2024. – A. Kalluvila, "Super-resolution of brain MRI via U-net architecture," International Journal of Advanced Computer Science and Applications, vol. 14, no. 5, pp. 26–31, 2023. – X. Hu, M. A. Naiel, A. Wong, M. Lamm, and P. Fieguth, "RUNet: a robust U-net architecture for image super-resolution," in IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Long Beach, CA, USA, 2019, pp.

Architecture of the SIF-CNN-SR1 (2/3)

Early processing

2 Convolutional blocks (Conv2d, BatchNorm2d, ReLU)

Encoder

3 levels
- Convolution 2D
- Batch Norm
- ReLU

Residual connexion for the first 2 blocks

Decoder

3 levels
- Convolution 2D
- Batch Norm.
- ReLU

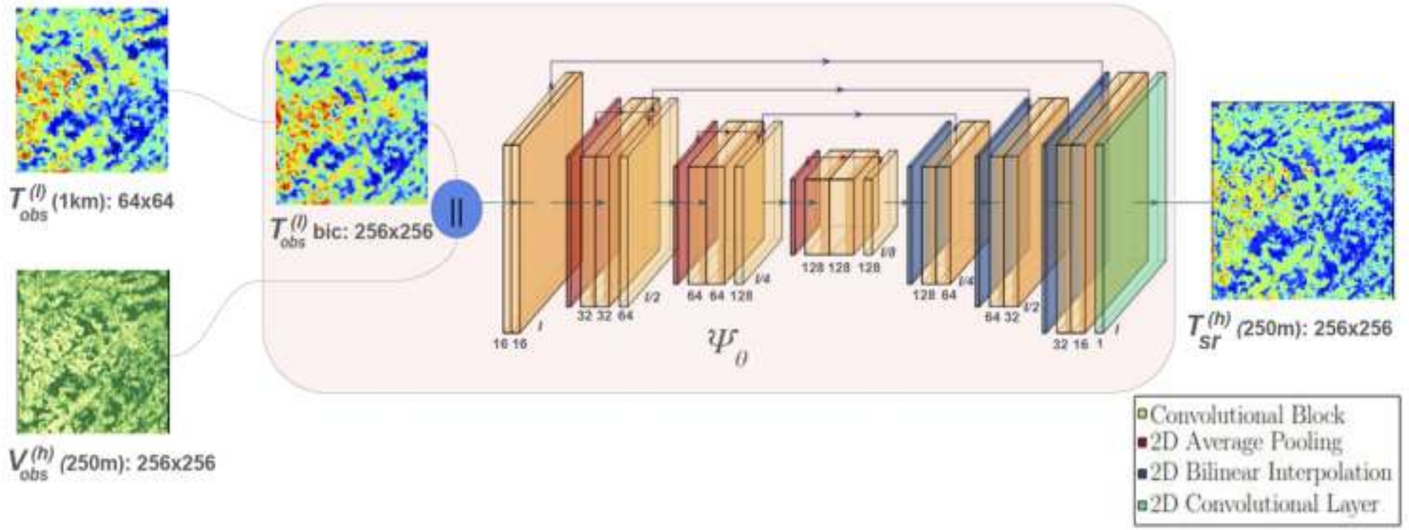
Skip connection

Final processing

One last Convolutional Block to reach the desired size of the image

Architecture of the SIF-CNN-SR1 (3/3)

SIF-CNN-SR1



Parameters:

417K

Loss Function (1/1)

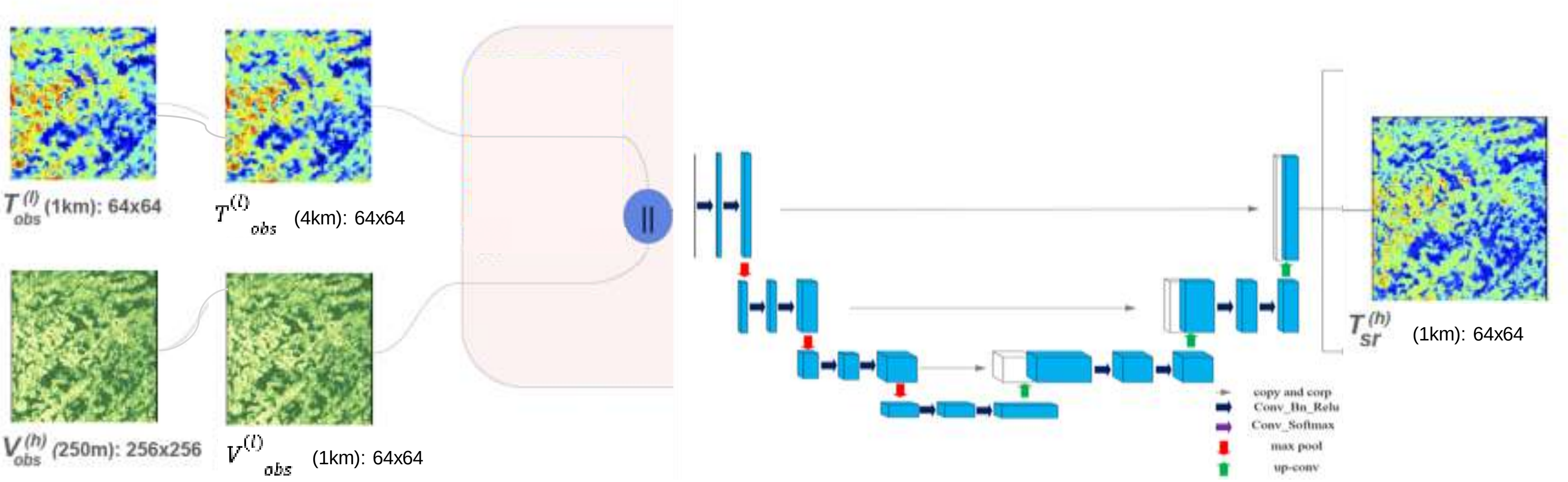
$$Loss = \alpha * \mathbf{MSE}(grad(NDVI_{input}), grad(LST_{output})) \\ + (1-\alpha) * \mathbf{MSE}(LST_{input}, H(LST_{output}))$$

- $grad : Sobel\ Gradient$
  $H : Degradation$

$$\alpha = 0.99$$

Architecture of the SC-UNet (1/1)

SC-Unet



Parameters:

1M

Source: *Land Surface Temperature Super-Resolution with a Scale-Invariance-Free Neural Approach: Application to MODIS*. Ait-Bachir, Romuald et al.
Wang, Chuanxia & He, Yuting & Qi, Xiaoming & Zhao, Ziteng & Yang, Guanyu & Zhu, Xiaomei & Zhang, Shaobo & Dillenseger, Jean-Louis & Coatrieux, Jean-Louis. (2019). BiSC-UNet: A fine segmentation framework for kidney and renal tumor. 10.24926/548719.013.

Loss Function (1/1)

$$Loss = \mathbf{MSE}(LST_{1km*1km}, \Phi(LST_{4km*4km}))$$

⇒ The output (1km * 1km) must match the input (1km * 1km)

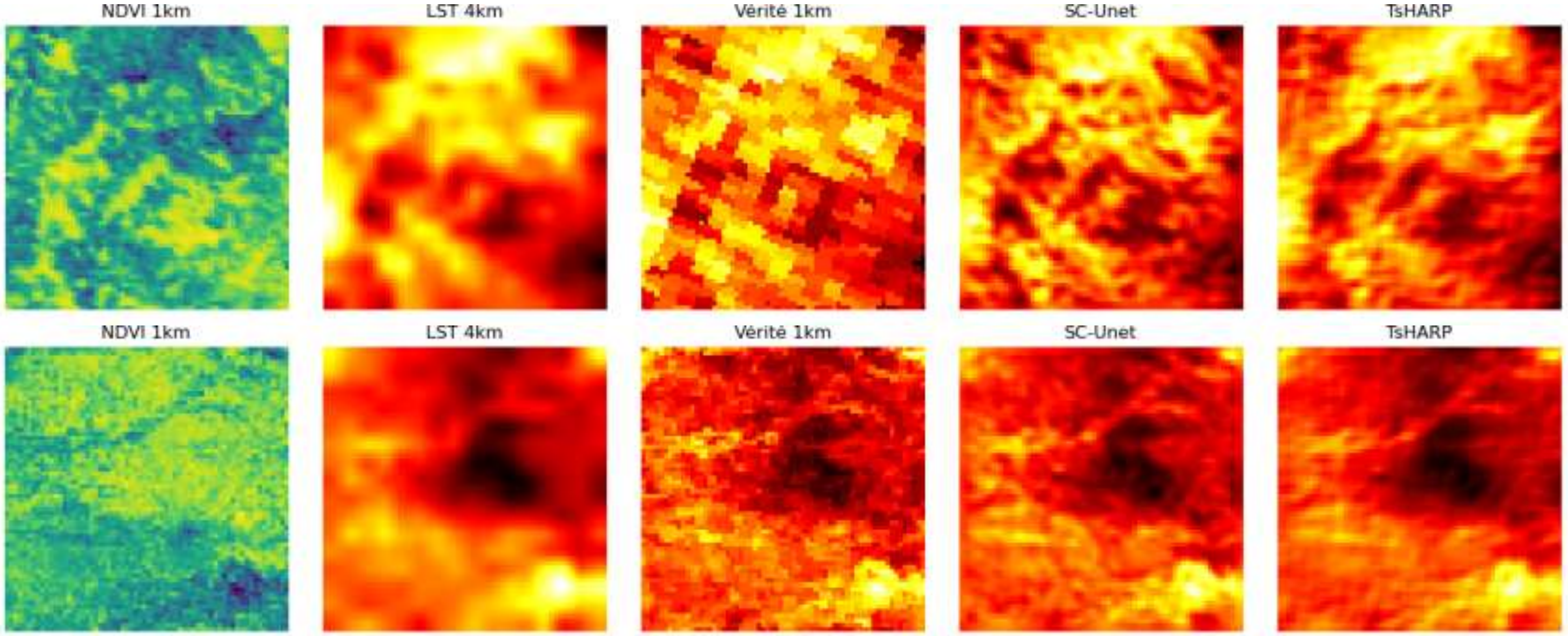
When it's tuned, we give the network the non-degraded NDVI and LST images to raise the resolution to 250m * 250m.

Φ = Neural Network

Metrics for evaluation (1/1)

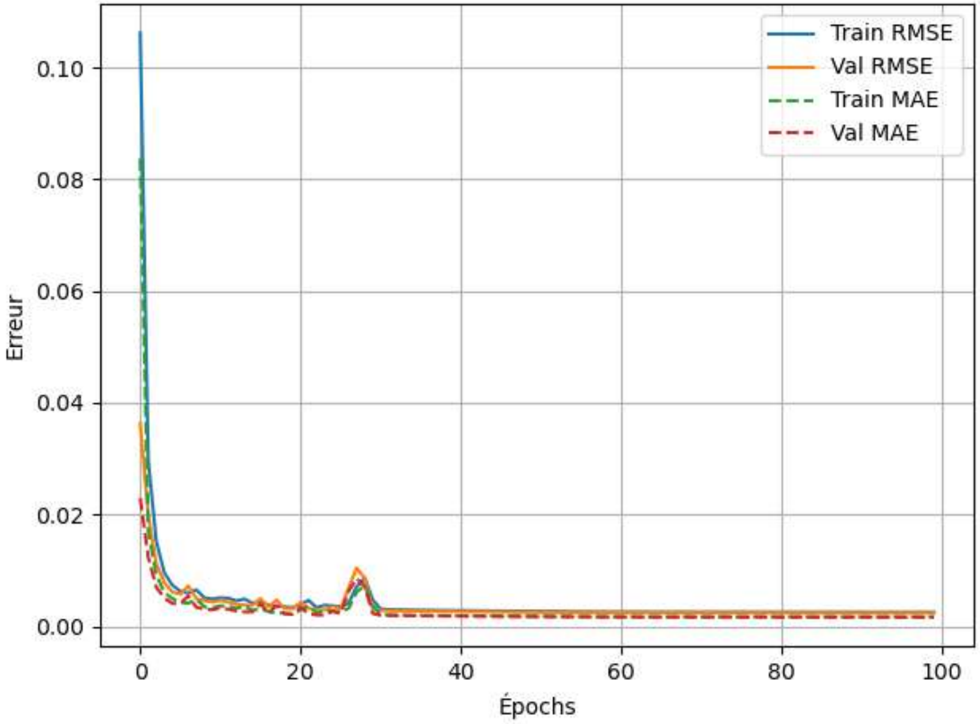
RMSE	• Good for understanding overall prediction accuracy.
MAE	• Gives a clearer idea of typical prediction error.
SSIM	• Correlates well with human perception.

Results : qualitative visualization (1/1)

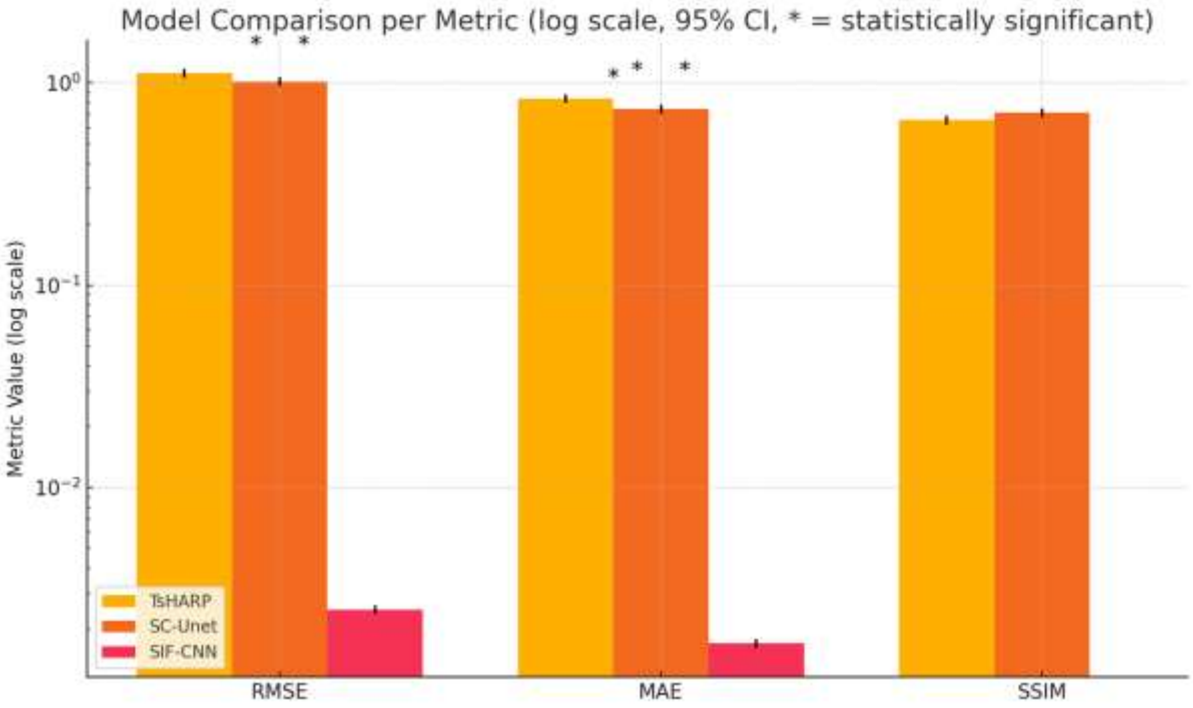


Results: SIF-CNN-SR1 loss evolution (1/4)

=== FINAL METRICS ===
RMSE: 0.0025
MAE: 0.0017
VAL_RMSE: 0.0024
VAL_MAE: 0.0016
TRAIN_LOSS: 0.0000



Results: quantitative comparison (2/4)

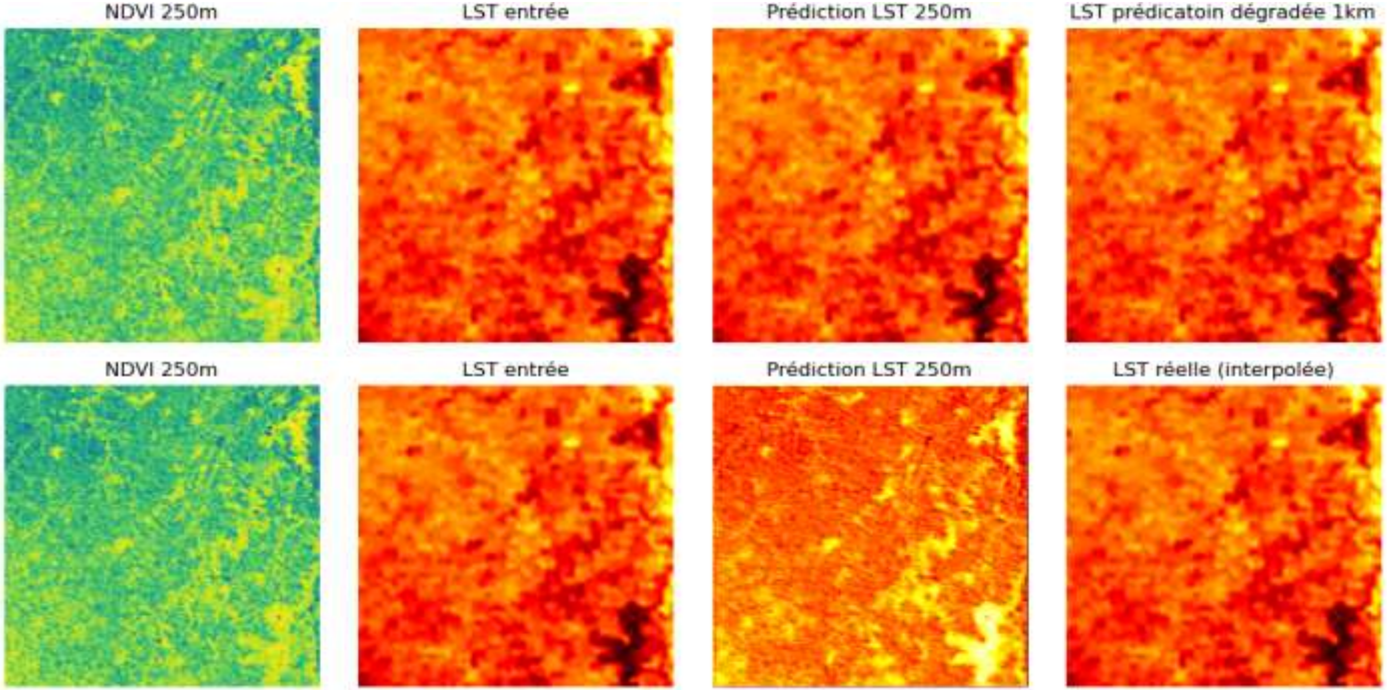


	TsHARP	SC-Unet	SIF-CNN-SR1
RMSE	1.1175	1.0164	0.0025
MAE	0.8370	0.7425	0.0017
SSIM	0.6551	0.7133	

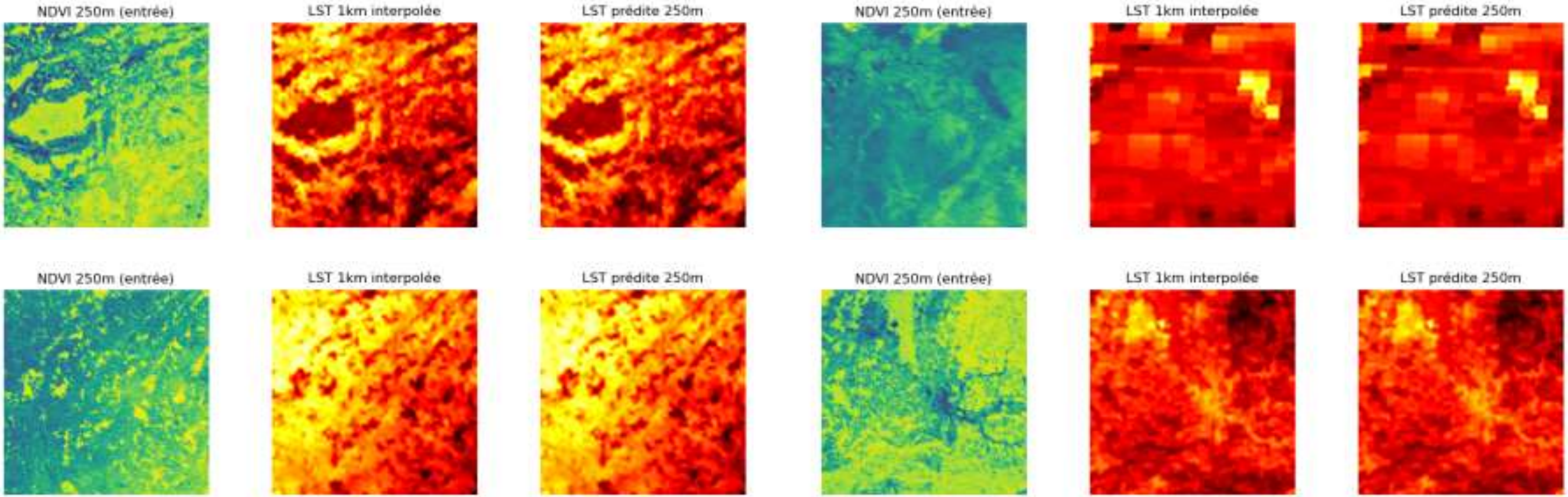
Lower **RMSE** and **MAE** → better prediction accuracy

Higher **SSIM** → better preservation of spatial structures

Results : qualitative visualization (3/4)



Results : qualitative visualization (4/4)



Perspectives (1/1)

Do a hyperparameter optimization

Explore other architectures

Gather ASTER data for validation

$$\hat{\theta} = \arg\min_{\theta} \alpha \left[J\left(\gamma G(V_{obs}^{(h)}), G(\Psi_\theta(V_{obs}^{(h)}, T_{obs}^{(l)})) \right) \right] + \\ + (1-\alpha) \left[J\left(T_{obs}^{(l)}, H(\Psi_\theta(V_{obs}^{(h)}, T_{obs}^{(l)})) \right) \right]$$

Model	α	γ	G
SIF-CNN-SR1	0.99	-0.5	Sobel filter
SIF-CNN-SR2	0.10	-0.25	High pass filter

Source: *Land Surface Temperature Super-Resolution with a Scale-Invariance-Free Neural Approach: Application to MODIS.* Ait-Bachir, Romuald et al.

Conclusion (1/1)

SIF-CNN-SR1 seems to do a predict a good upscaled LST image.

The scale-invariance-free approach offers strong potential for operational Earth observation applications.

Compare to ASTER to validate the prediction and compute its accuracy.

Extend the approach to other auxiliary indices or climatic variables (e.g., NDBI, NDWI, AOD, LAI).

Investigate temporal consistency and multi-temporal fusion to enhance prediction stability over time.

Determine **where** on Earth this method is not usable

Source: *Land Surface Temperature Super-Resolution with a Scale-Invariance-Free Neural Approach: Application to MODIS*. Ait-Bachir, Romuald et al.

Thanks for your attention!

Any question?

Results: SC-UNet loss evolution (annexe)

=== FINAL METRICS ON VALIDATION ===
RMSE: 1.0164
MAE: 0.7425
VAL_RMSE: 1.0164
VAL_MAE: 1.0164
TRAIN_LOSS: 0.4