Internal Learning for Satellite Image Super-Resolution

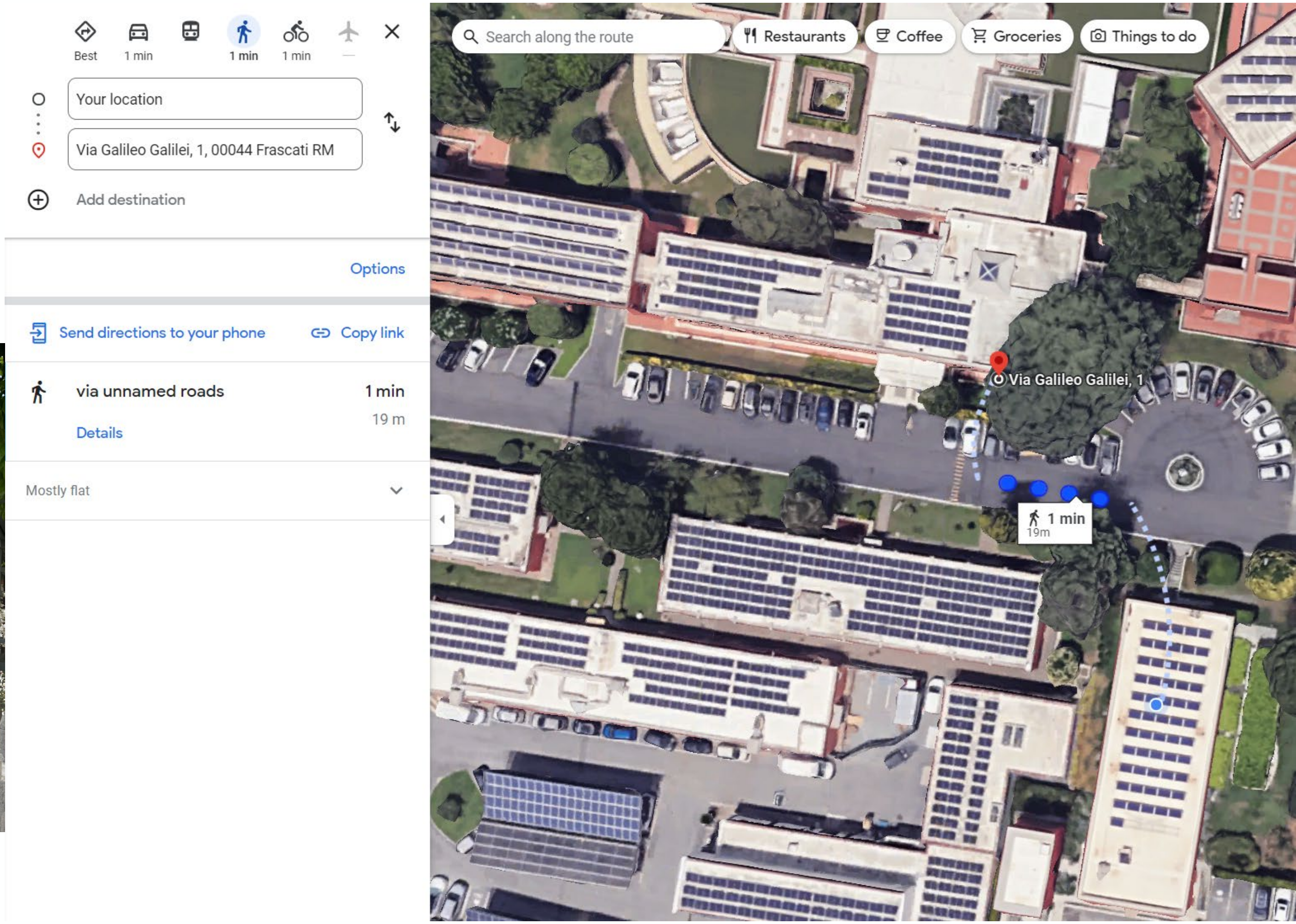
Mikolaj Czerkawski, Φ-lab, European Space Agency

ESA SUREDOS 2024

30/05/2024

Fun Fact

- I came here by foot
- 1 minute
- 19 metres
- Spectacular views

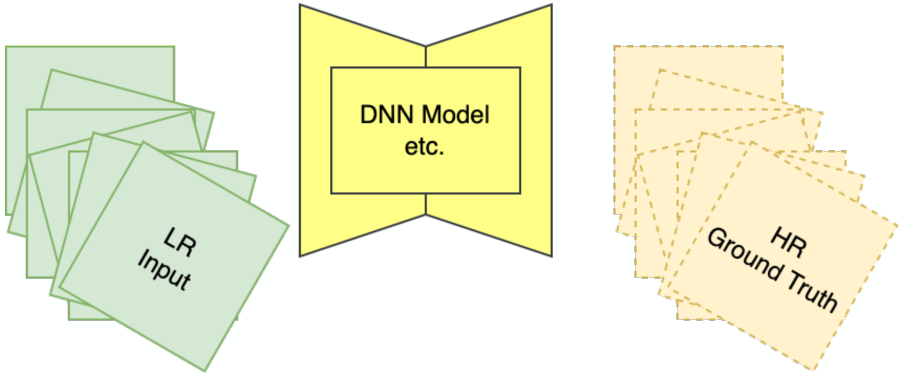


Internal Learning

- Alternative mode of deep learning
- Unlike conventional (external) learning, the priors are extracted from the inference sample
- **Topology can be defined “on the spot”**

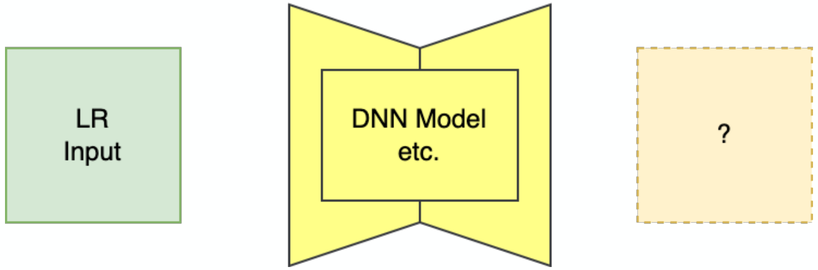
Convention ("External")

Involves optimising a model over a set of example samples, most often with ground truth to maximise aggregate performance.



Internal Learning

Involves optimising a model over a set of example samples, most often with ground truth to maximise aggregate performance.



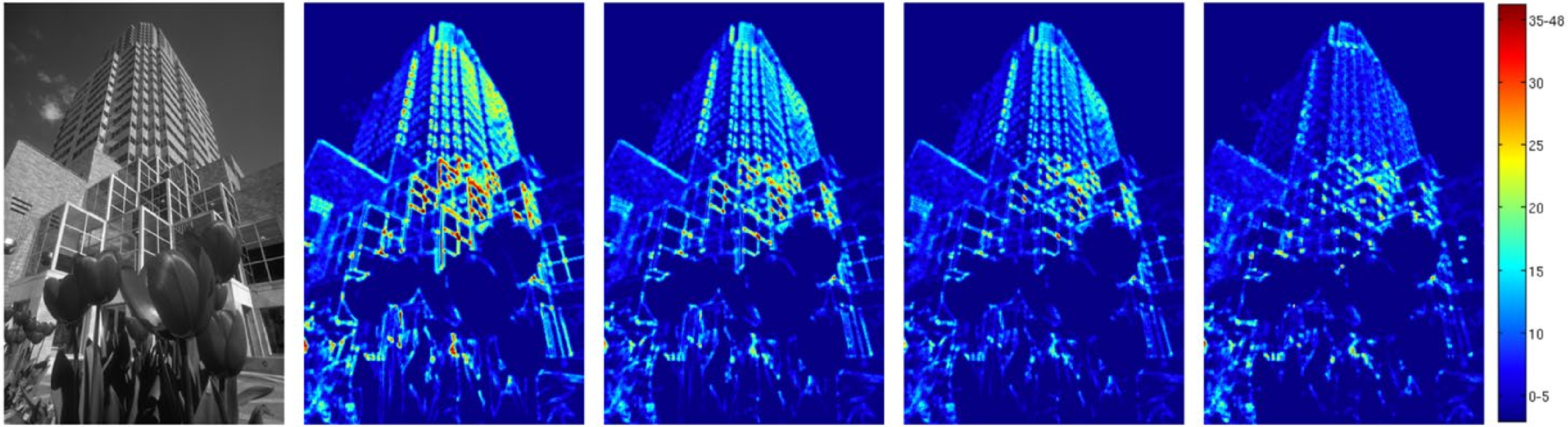
Internal Learning for Super-Resolution

- Why should it work at all for super-resolution?

  recurrence of patches across scales of a single image

  S. B. D. Glasner and M. Irani. “*Super-Resolution from a Single Image*” (ICCV, 2009)

  M. Zontak and M. Irani. “*Internal Statistics of a Single Natural Image*” (CVPR, 2011)



(a) Input Image (b) Ext. DB - 5 imgs. (c) Ext. DB - 40 imgs. (d) Ext. DB - 200 imgs. (e) Internally (Error values)

Figure 4. **External vs. Internal “Expressiveness”.** *Errors induced by replacing each patch from the input image with its most similar patch found in: (b)-(d) an external database of 5, 40, 200 images, vs. (e) internally, within the input image (excluding the patch itself and its immediate local vicinity). Red signifies high errors, blue signifies low errors. Patches obtain lower error internally than externally.*

Internal Learning for Super-Resolution

- **recurrence of patches** across scales of a single image

  “*Blind Super-Resolution Kernel Estimation using an Internal-GAN*” (NeurIPS 2019)

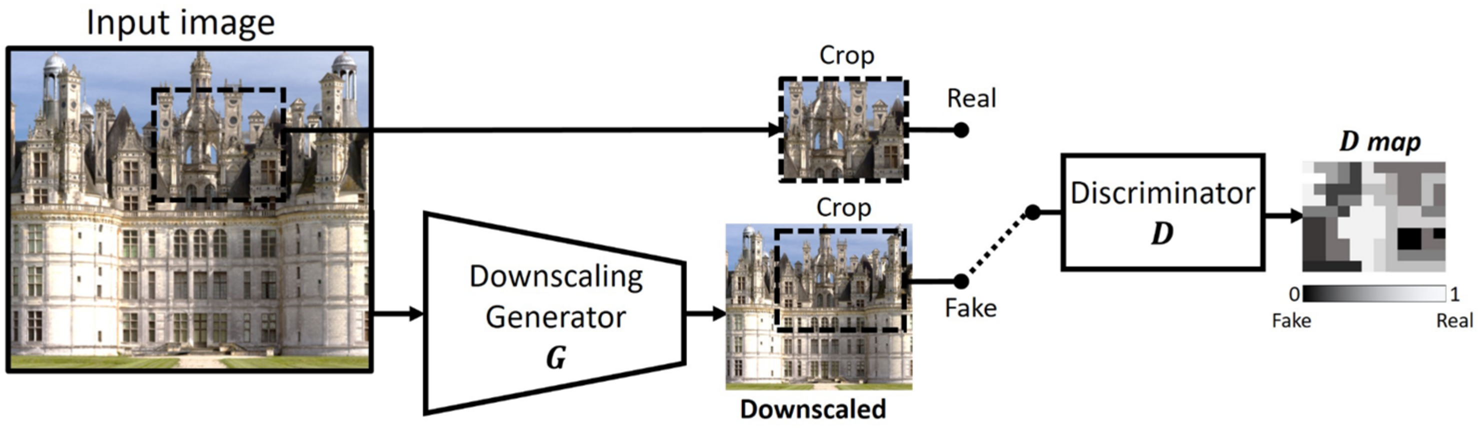
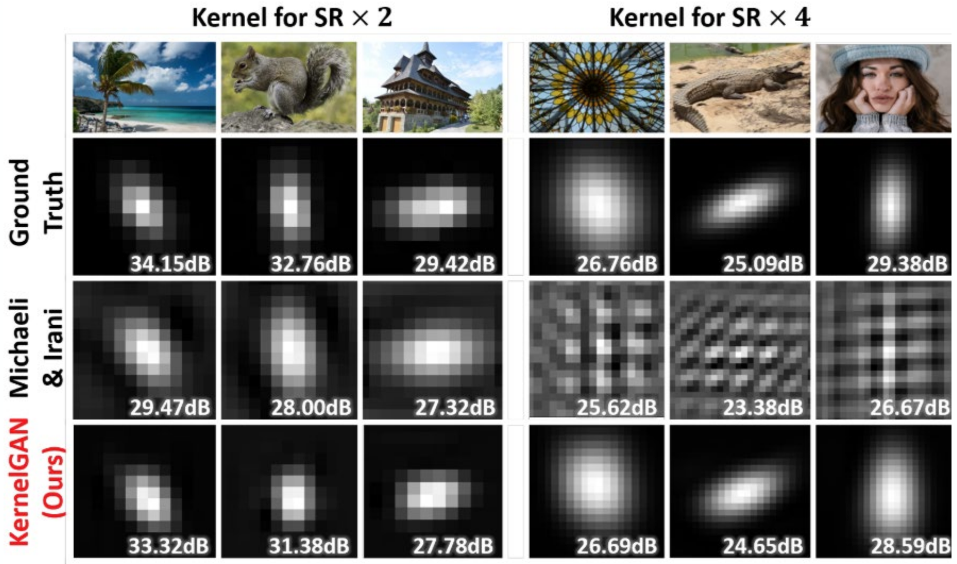


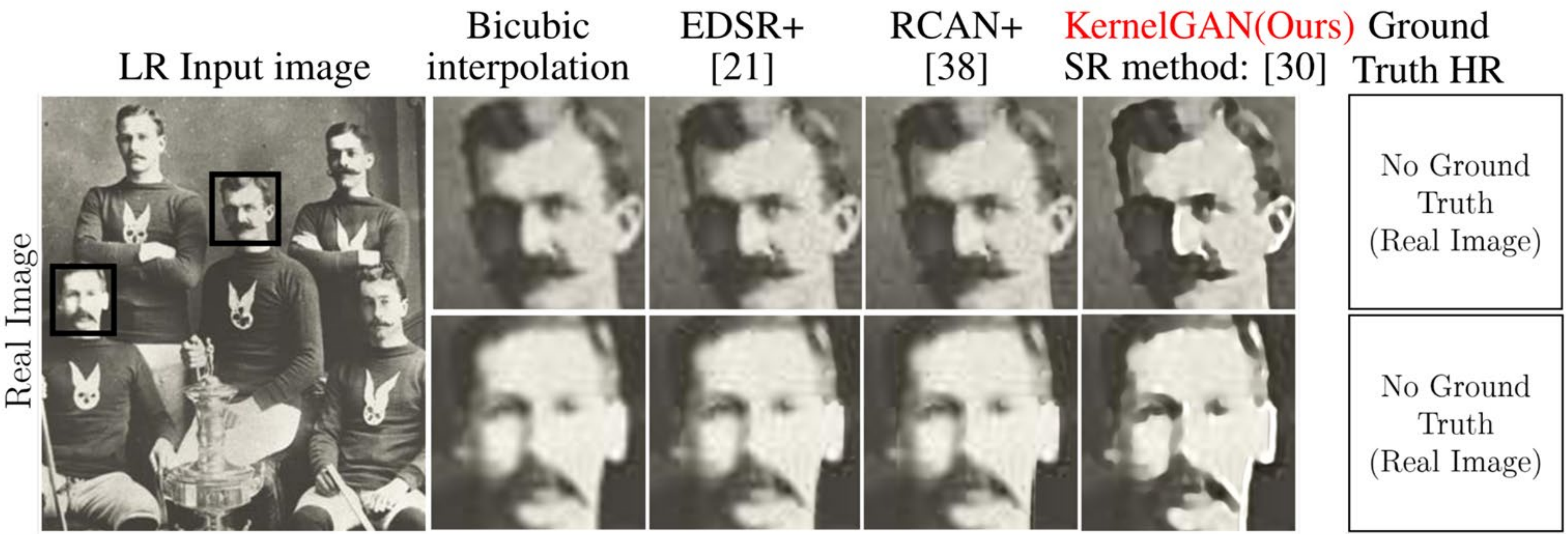
Figure 2: ***KernelGAN***: *The patch GAN trains on patches of a single input image (real). D tries to distinguish real patches from those generated by G (fake). G learns to downscale X2 the image while fooling D i.e. maintaining the same distribution of patches. Both networks are fully convolutional, which in the case of images implies that each pixel in the output is a result of a specific receptive field (i.e. patch) in the input.*



Internal Learning for Super-Resolution

- **recurrence of patches** across scales of a single image

  “Blind Super-Resolution Kernel Estimation using an Internal-GAN” (NeurIPS 2019)

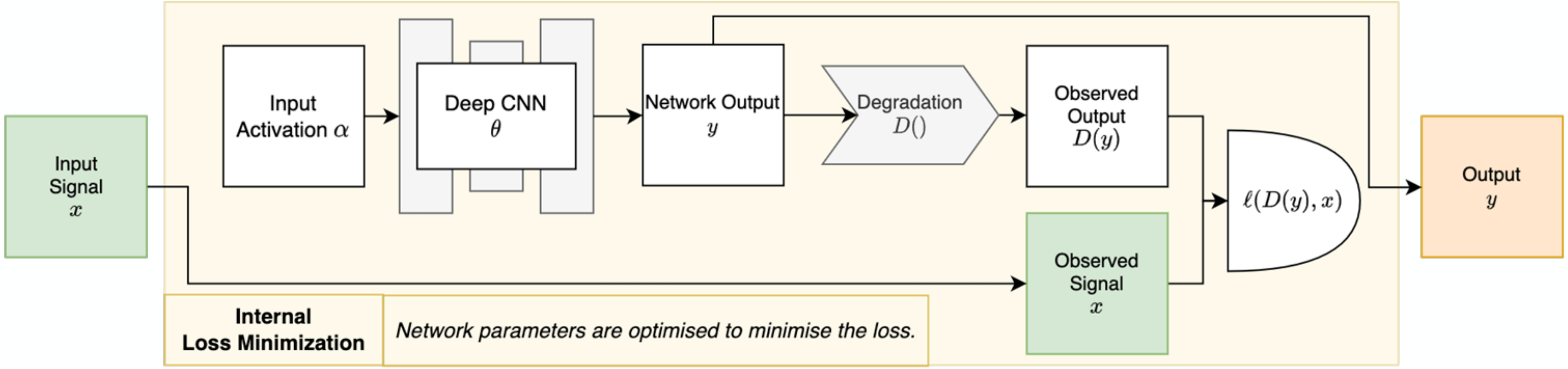


Deep Image Prior

- The power of "deep convolutions" prior

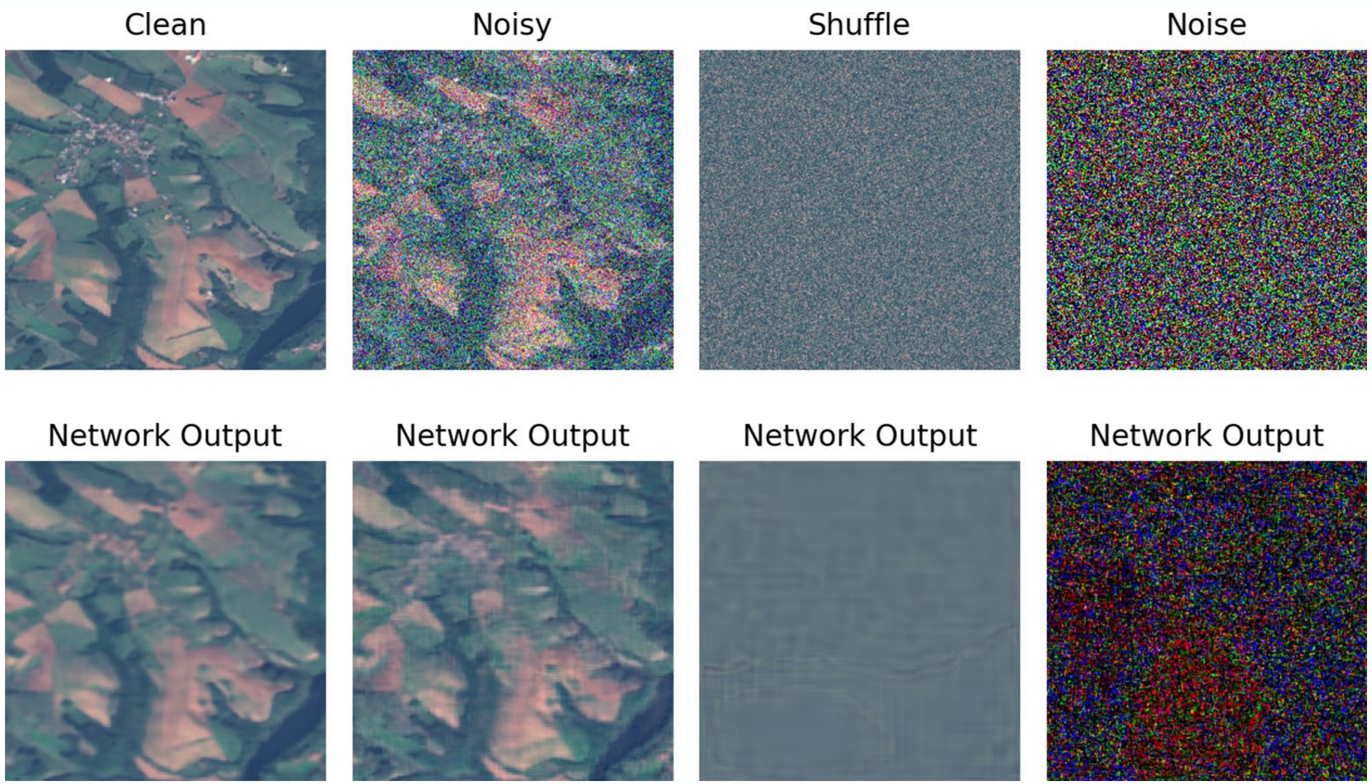
  D. Ulyanov, A. Vedaldi, V. Lempitsky "*Deep Image Prior*" (CVPR, 2018)

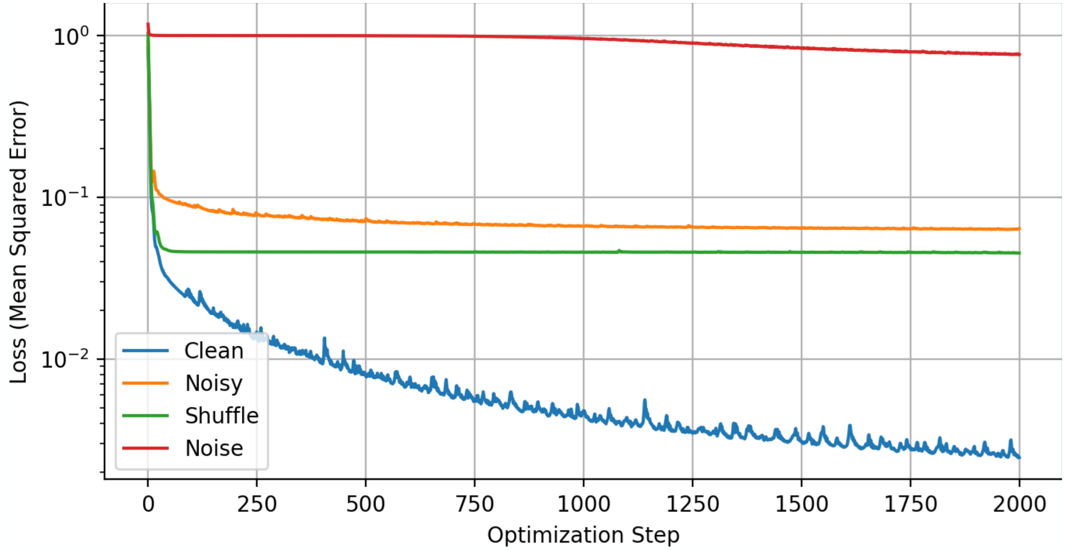
  M. Czerkawski "*Satellite image cloud removal: learning within and beyond the sample*" (2023)



Deep Image Prior

- Natural impedance of DNNs to unstructured data can be observed
    - D. Ulyanov, A. Vedaldi, V. Lempitsky “*Deep Image Prior*” (CVPR, 2018)
    - M. Czerkawski “*Satellite image cloud removal: learning within and beyond the sample*” (2023)



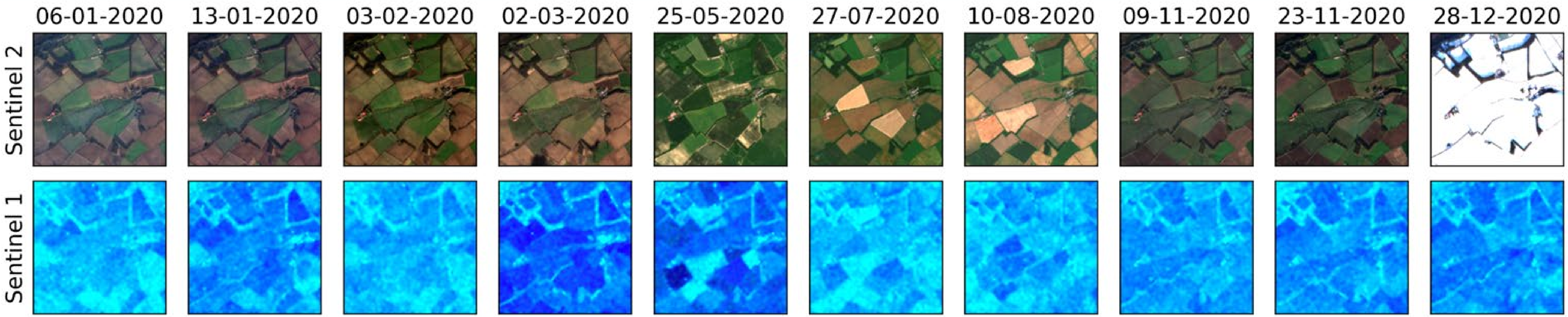


Internal Super-Resolution

- Inherent multi-modality of the SR problem
- Depending on the **downsampling kernel**, **SR factor,** and **underlying data distribution**, each LR can correspond to a large number of images
- Internal learning does not model the **data distribution**

  M. Czerkawski "*Satellite image cloud removal: learning within and beyond the sample*" (2023)

  M. Czerkawski et al. "*Deep Internal Learning for Inpainting of Cloud-Affected Regions in Satellite Imagery*" (MDPI Remote Sensing, 2022)

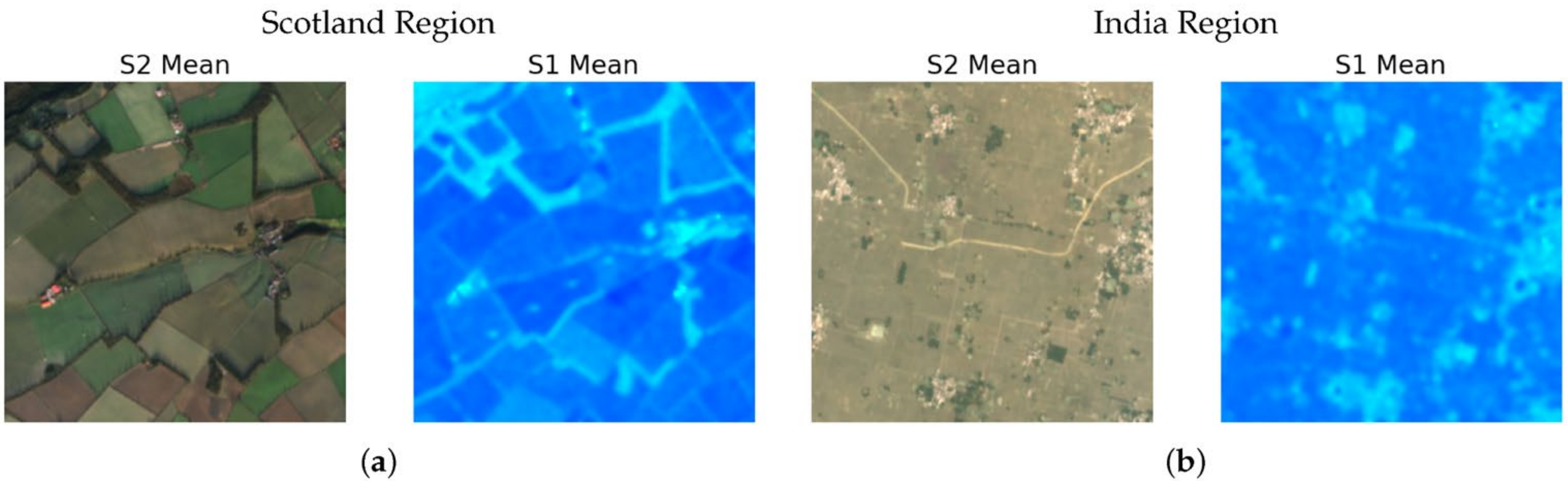


Internal Super-Resolution

- Inherent multi-modality of the SR problem
- Depending on the **downsampling kernel**, **SR factor,** and **underlying data distribution**, each LR can correspond to a large number of images
- Internal learning cannot model the **data distribution**

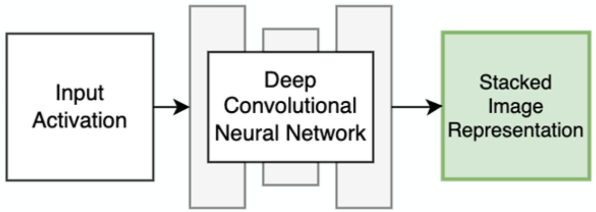
    M. Czerkawski "*Satellite image cloud removal: learning within and beyond the sample*" (2023)

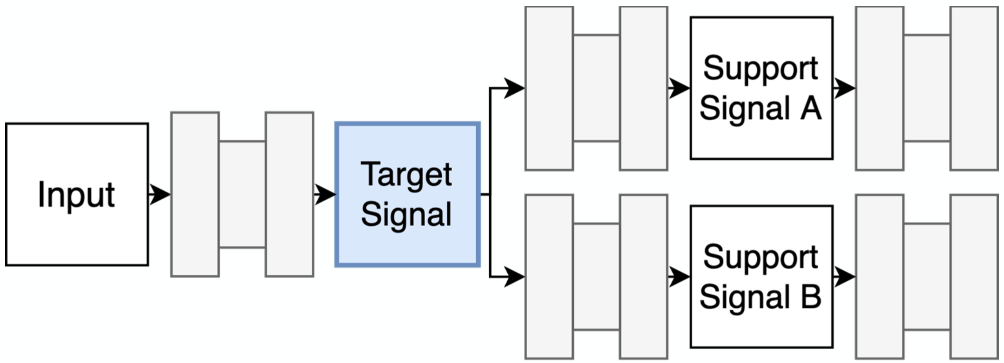
    M. Czerkawski et al. "*Deep Internal Learning for Inpainting of Cloud-Affected Regions in Satellite Imagery*" (MDPI Remote Sensing, 2022)

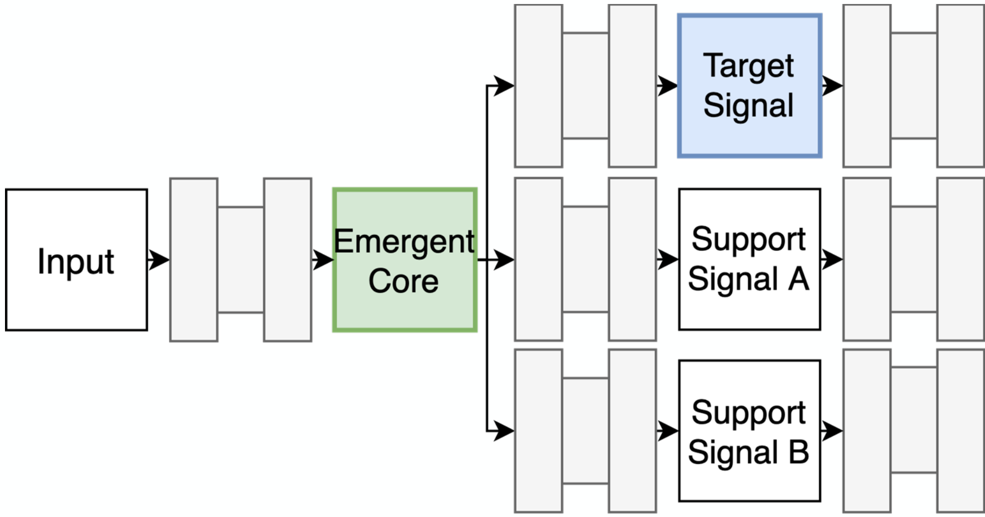


Multi-modal Convolutional Parameterisation Network

- Expansion to Deep Image Prior to handle spatially aligned multi-modal samples
- Includes additional local adjustment heads to handle disparities between domains

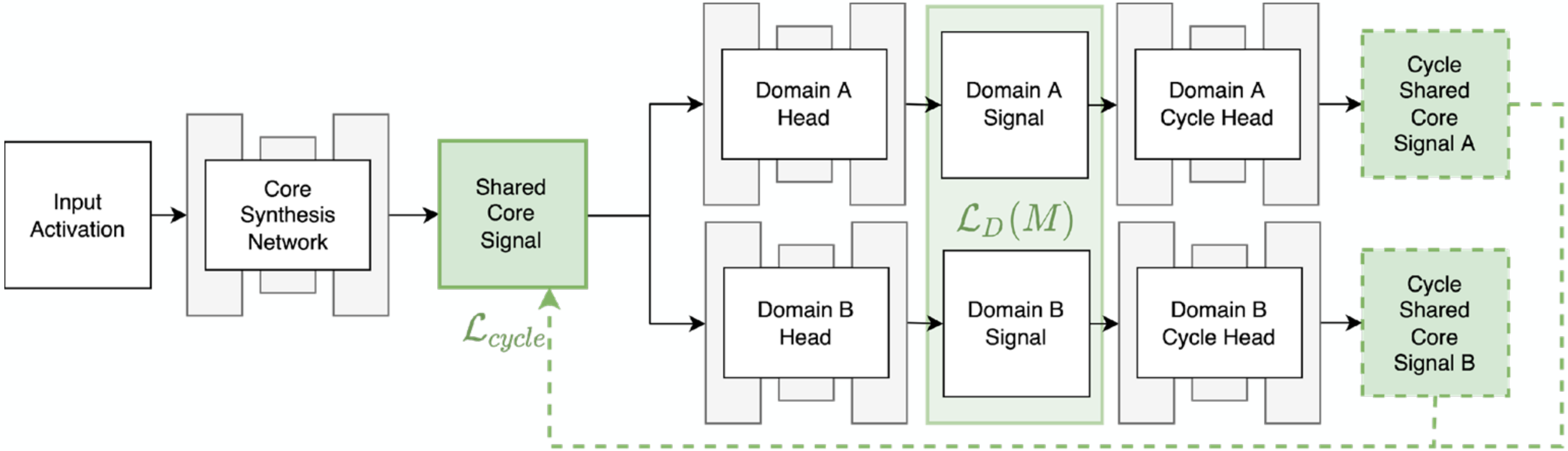






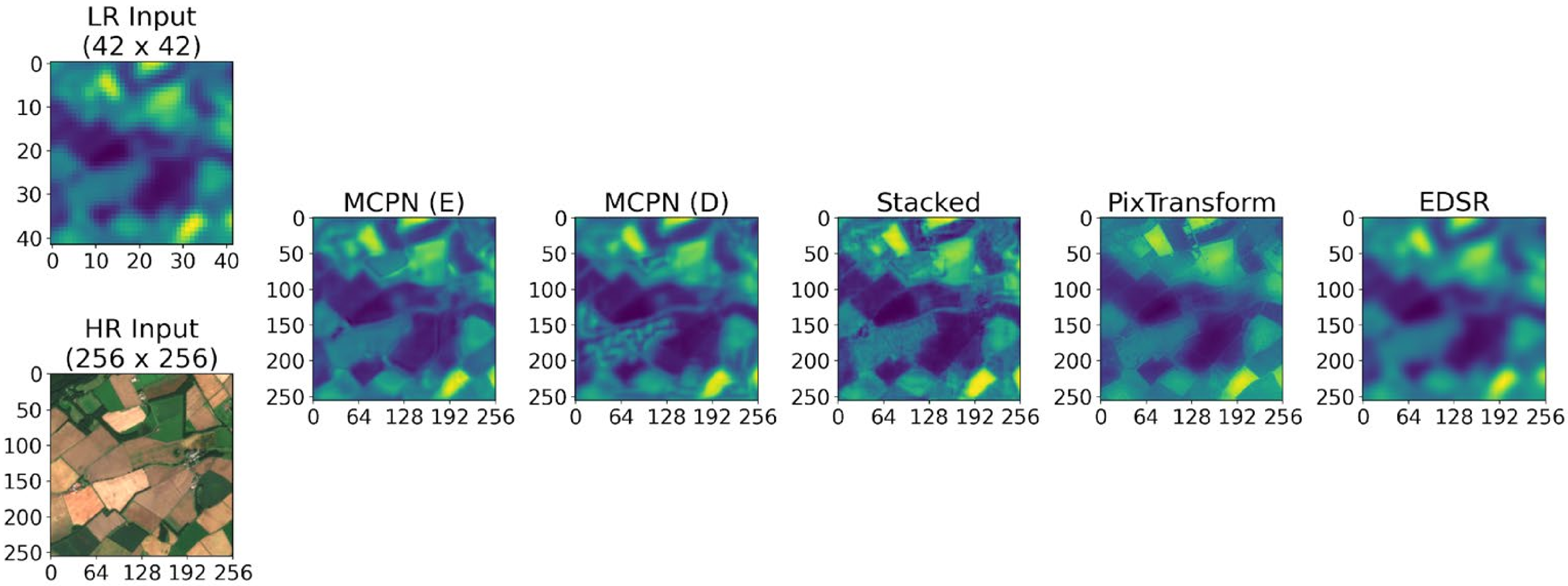
Multi-modal Convolutional Parameterisation Network

- Loss include **domain-specific losses** as well as **cyclic constraints**



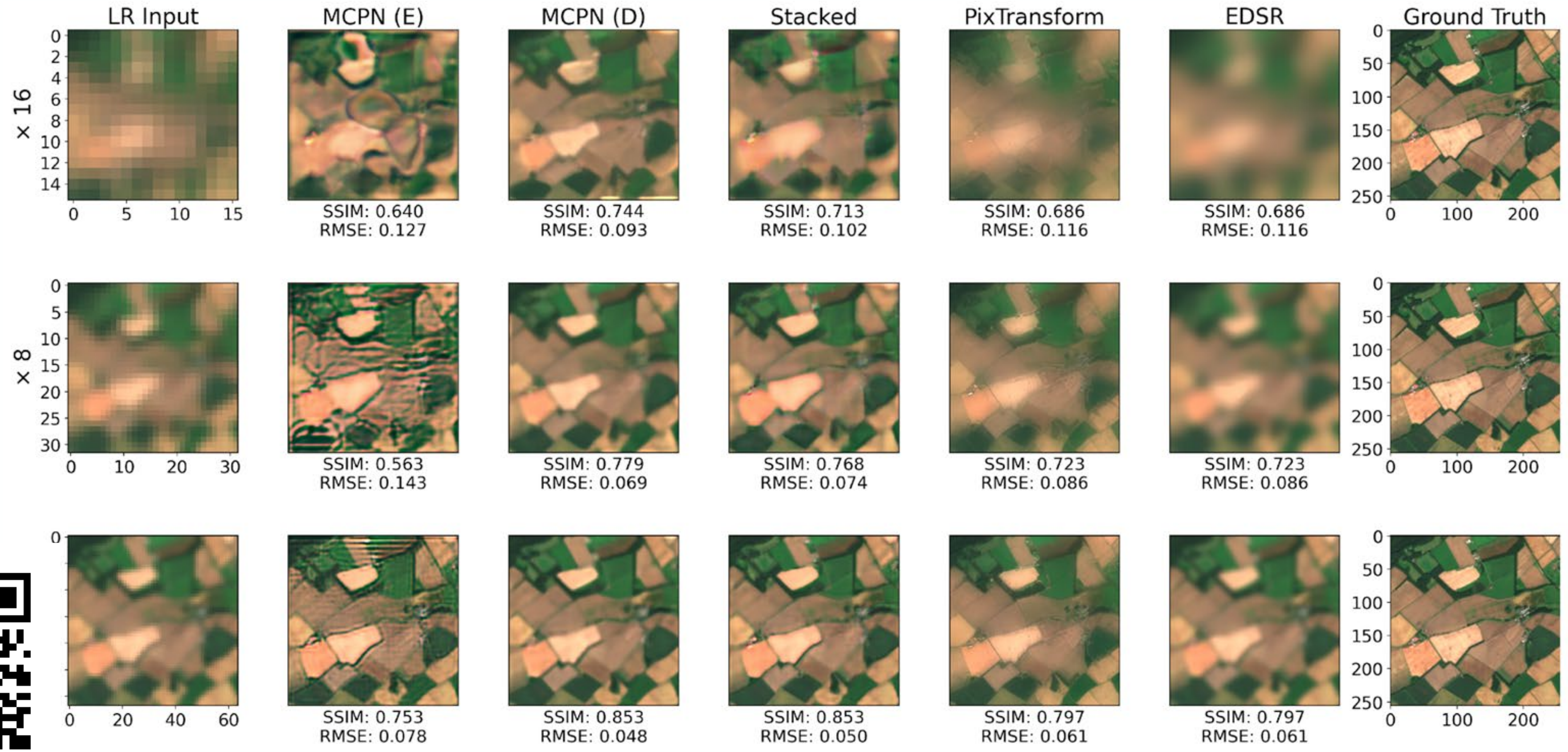
Multi-modal Convolutional Parameterisation Network

- Example: upsample the 60 m SWIR band of Sentinel-2 with RGB bands used as reference



Multi-modal Convolutional Parameterisation Network

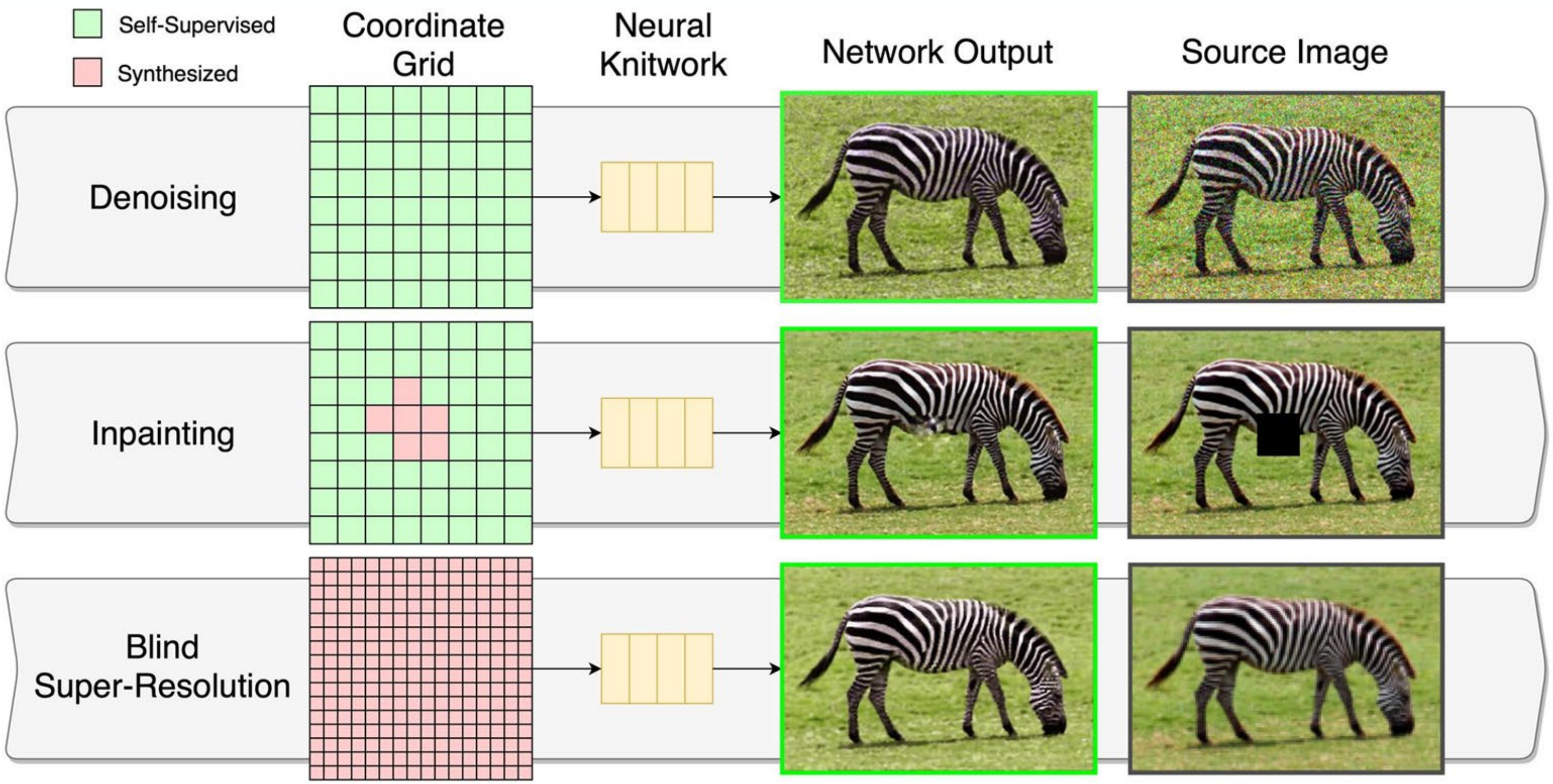
- RGB bands based on historical optical mean



Neural Knitworks

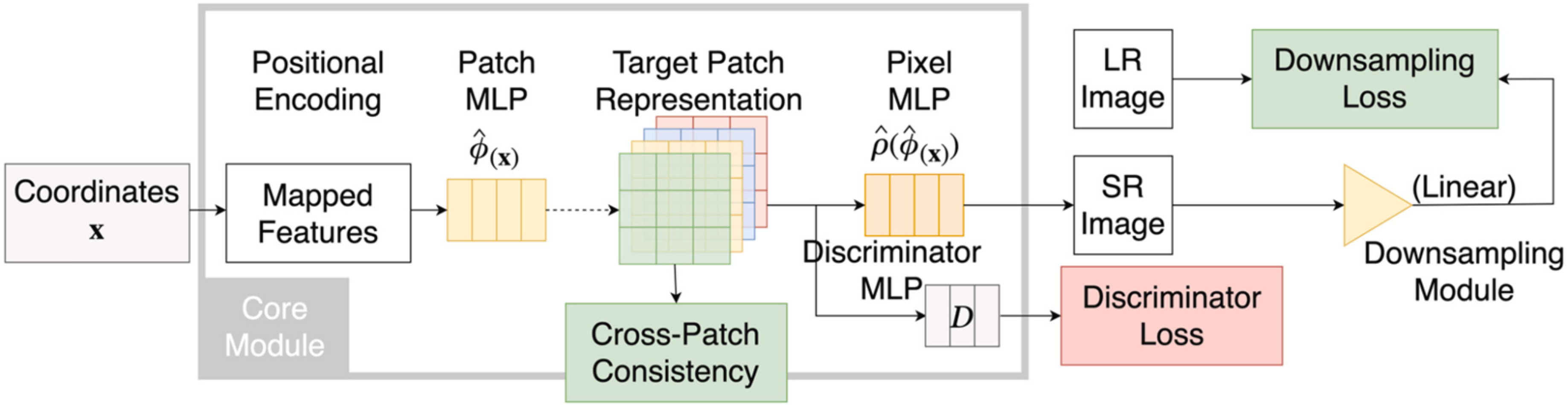
- Flexible model for internal learning based on **Neural Implicit Representations**
- A tiny network is used to map from coordinate space to colour space

M. Czerkawski "*Neural Knitworks: Patched neural implicit representation networks*" (Pattern Recognition, 2024)



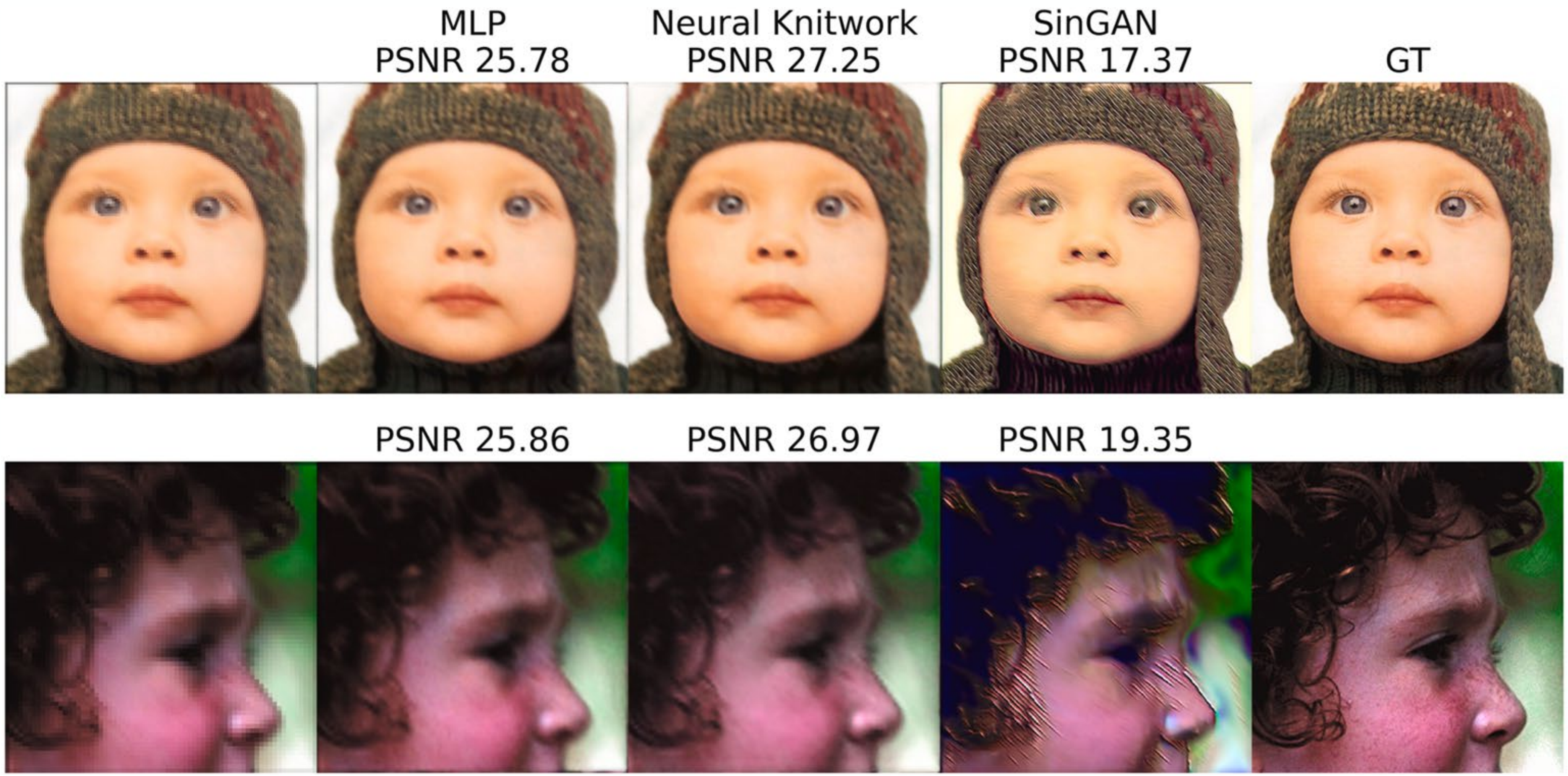
Neural Knitworks

- Combination of:
    - NIR (Neural Implicit Representation)
    - Adversarial Patch Loss
    - Deep Linear Approximation of the Kernel
- M. Czerkawski “*Neural Knitworks: Patched neural implicit representation networks*” (Pattern Recognition, 2024)



Neural Knitworks

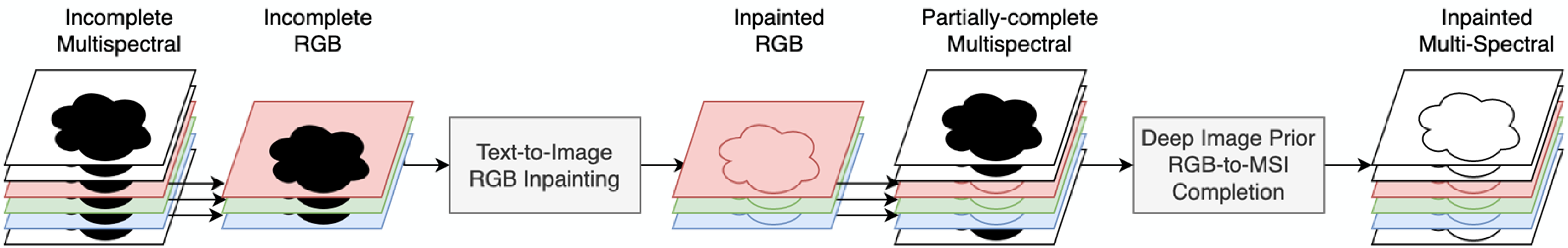
- Internal learning with adversarial losses can often lead to artefacts - neural implicit representation can provide the right foundation for internal adversarial losses as in Neural Knitworks
  M. Czerkawski "*Neural Knitworks: Patched neural implicit representation networks*" (Pattern Recognition, 2024)



Internal Learning for Modality Transfer (ZS RGB-2-MSI)

- Internal learning for zero-shot model prediction transfer
- Transfer RGB predictions to MSI in a zero-shot manner

Czerkawski & Tachtatzis "*Exploring the Capability of Text-to-Image Diffusion Models With Structural Edge Guidance for Multispectral Satellite Image Inpainting*" (IEEE GRSL 2024)



Conclusions

- Alternative learning paradigm
- Advantages
  - Flexible problem definition (topology, conditions etc.)
  - No large-scale training costs
  - No large-scale dataset required
- Disadvantages
  - Slower inference (mostly a few mins per image)
  - No data prior
- Future prospects
  - *Can we push the conventional models to make better use of the information that already exists in the sample?*