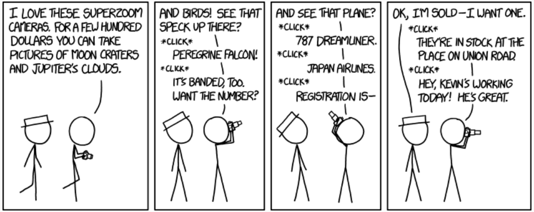
Super-resolution in EO: Best practices for performance and quality metrics

Julien MICHEL

2022.05.26

CESBIO, Université de Toulouse, CNES/CNRS/INRAe/IRD/UPS, Toulouse, FRANCE



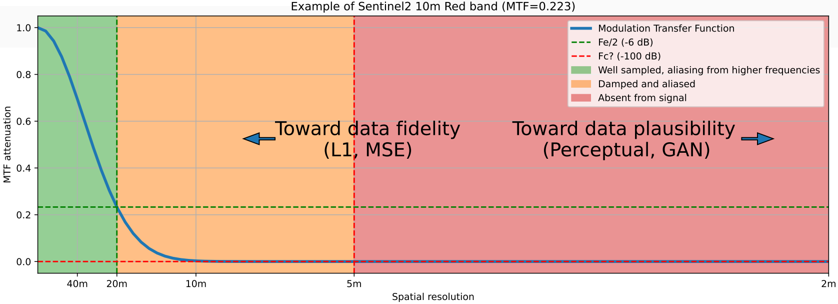
https://xkcd.com/1719

HR / VHR Optical imaging sensors in remote sensing

- Optics, integration time, etc ... $\Rightarrow$ low pass filter with spatial cut-off frequency f_c
- Detectors size $\Rightarrow$ spatial sampling frequency f_e
- System described by frequency damping at $f_e/2$:
  - High MTF : crisp images, but aliasing
  - Low MTF : Well sampled, but blurry

Super-resolution or super-restoration ?

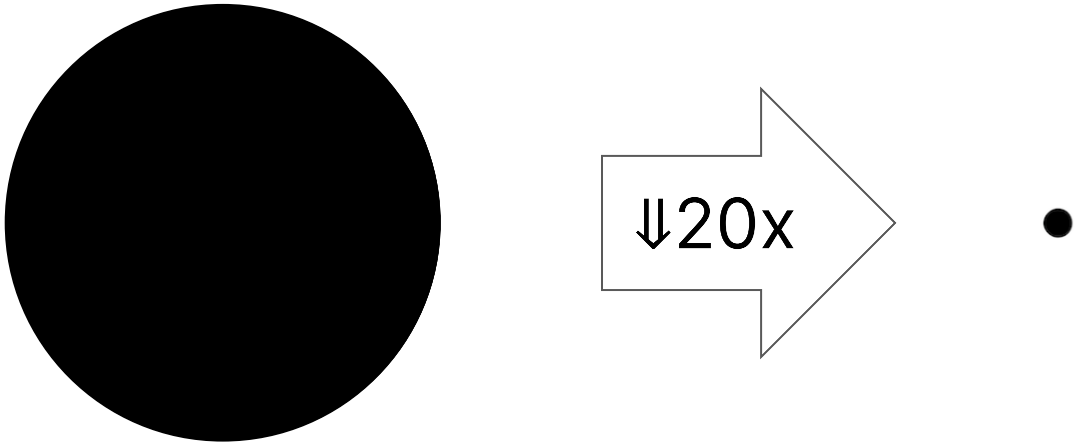
- Efficient restoration techniques use sensor priors. Real data for users are far from the sensor $\Rightarrow$ SISR is more agnostic.
- Networks trained on simulated data often fail IRL $\Rightarrow$ Need for realistic datasets

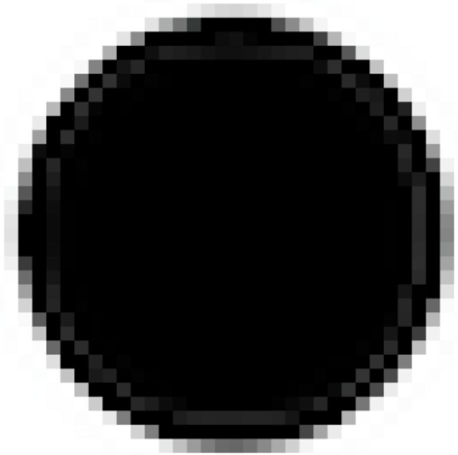


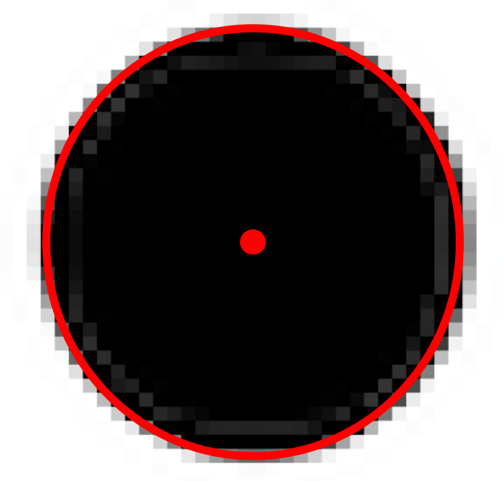
Single image super-resolution explained

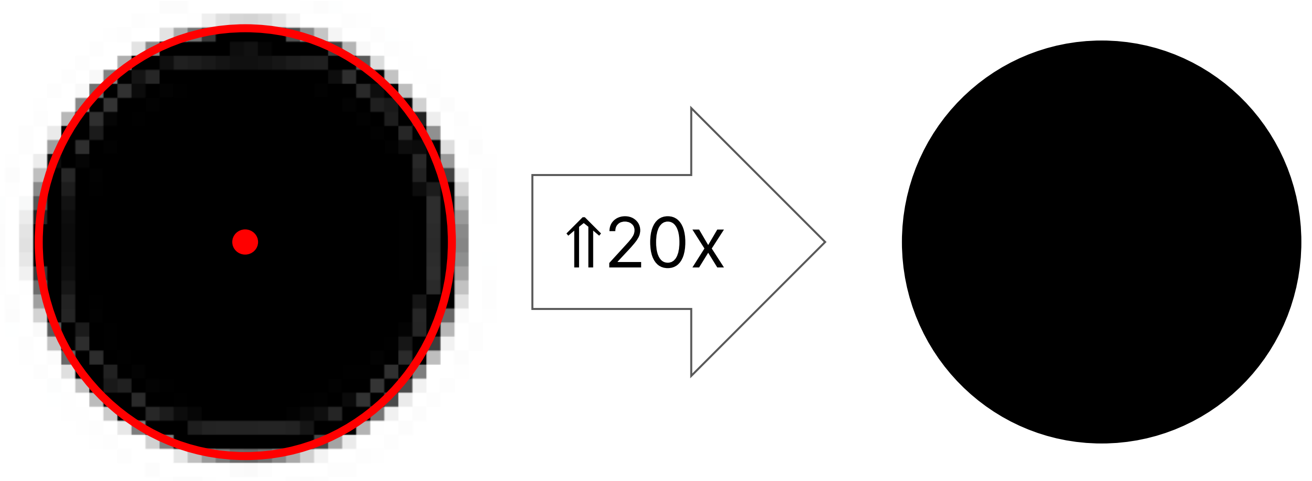
Dedicated to Claude Shannon and Harry Nyquist

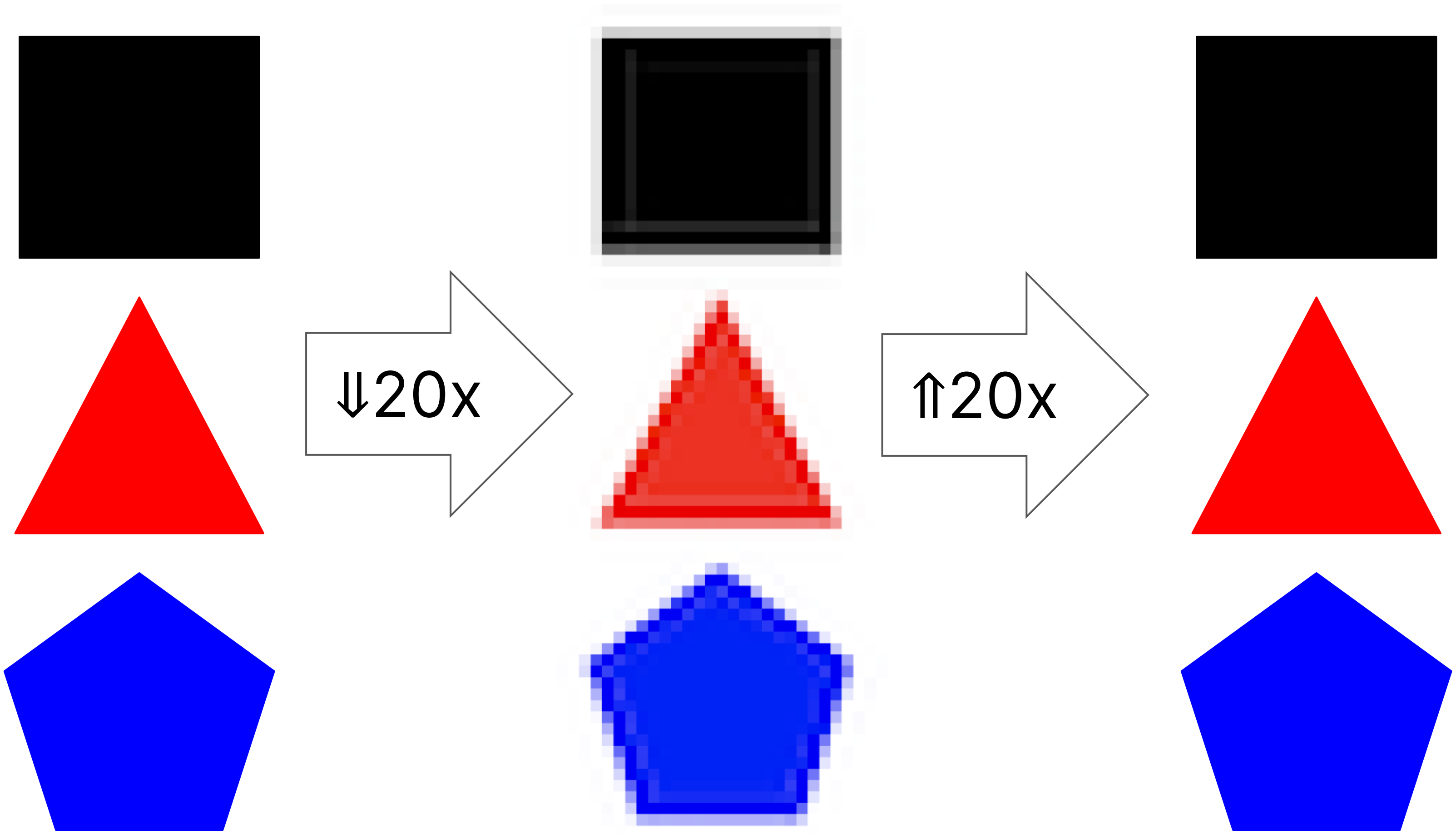
Yosef Akhtman, May 27, 2022

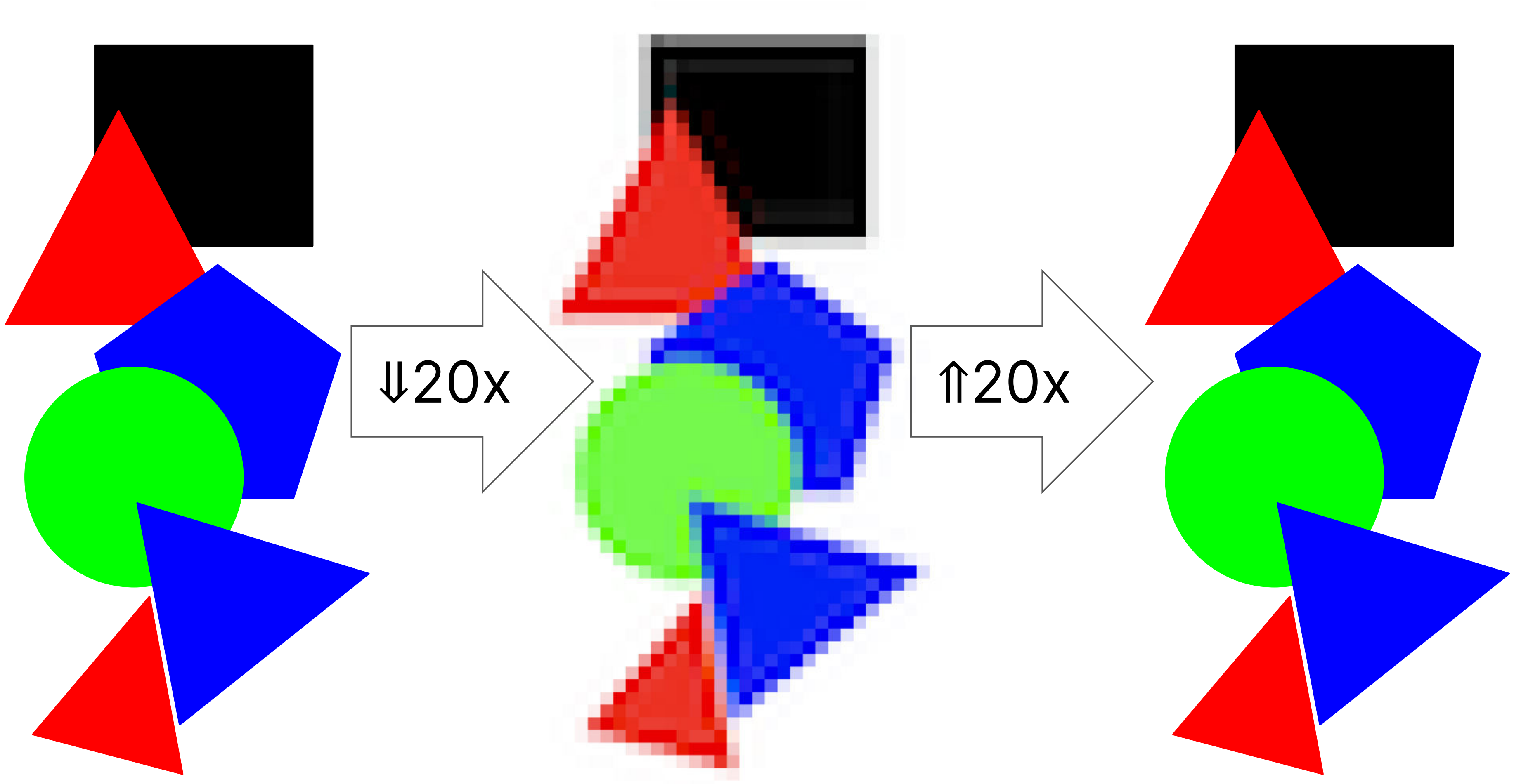

⇓20x

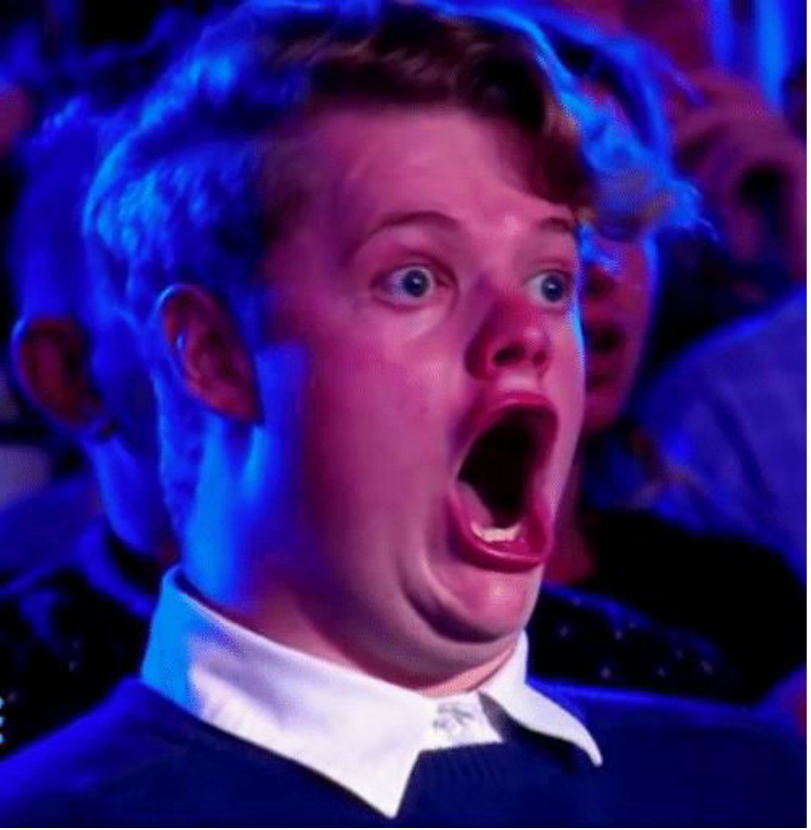
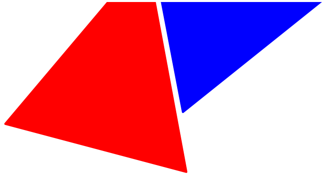
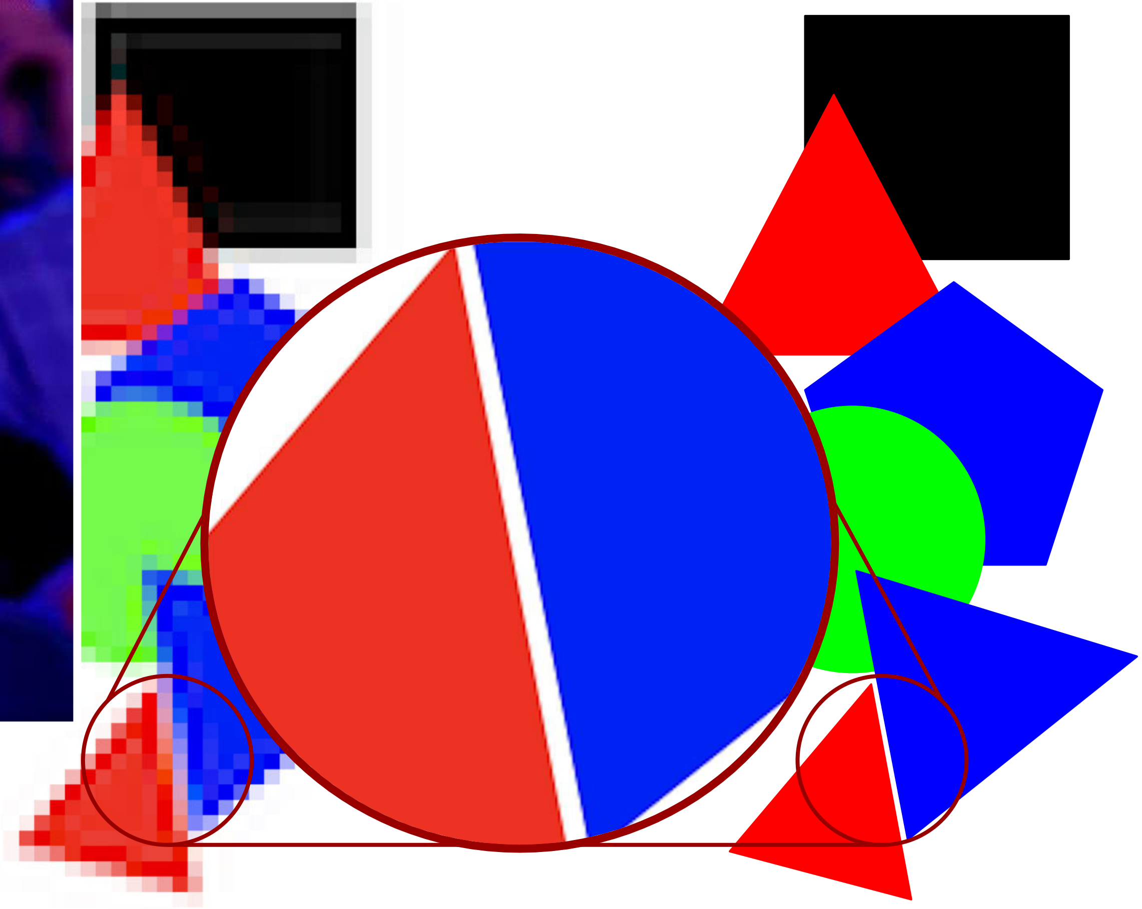


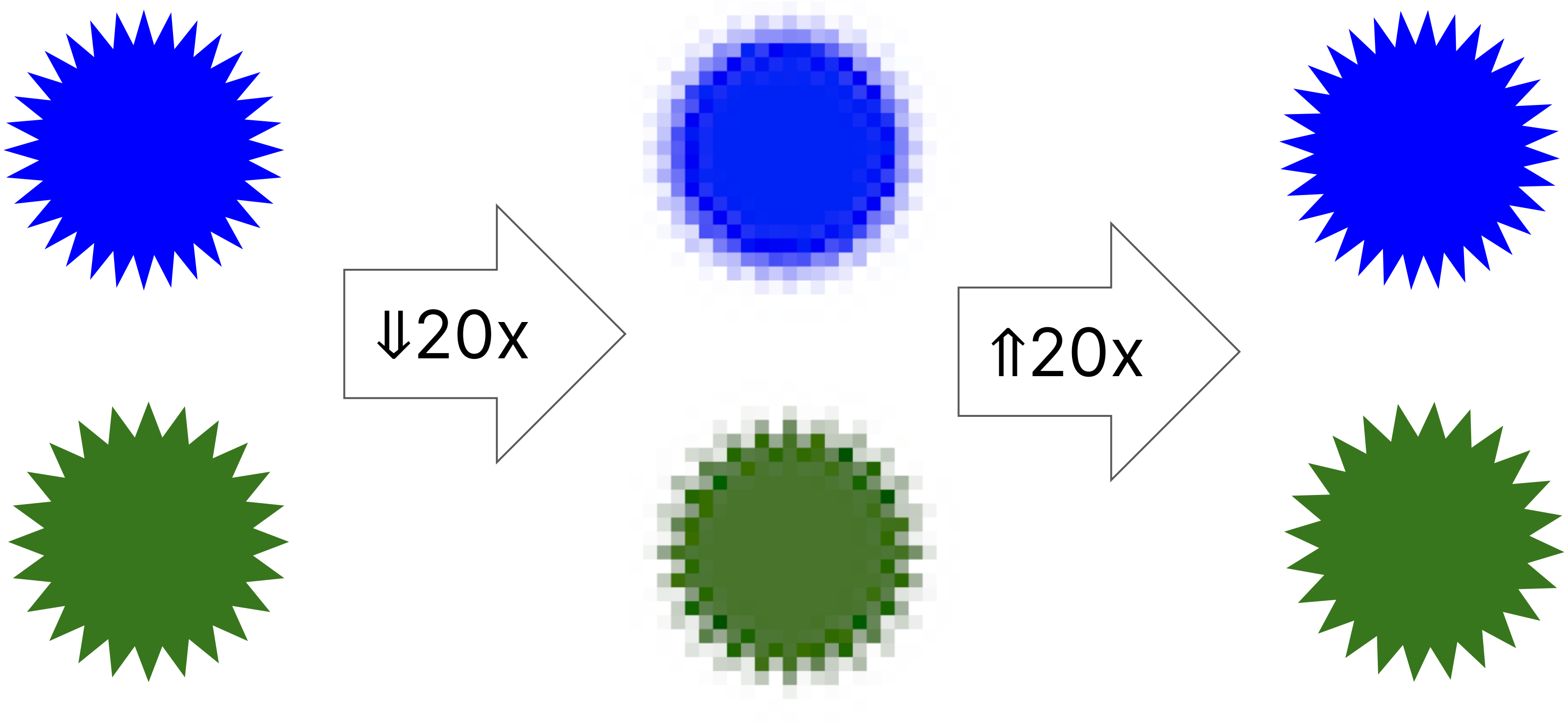



⇑20x


⇓20x
⇑20x


⇓20x
⇑20x

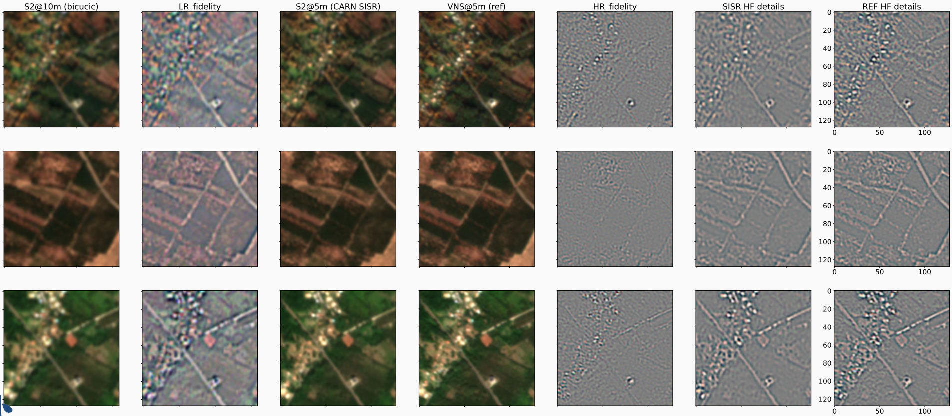

⇓20x
⇑20x

ya@gamma.earth
http://demo-dr.gamma.earth/

Quality Metrics, option A : physically informed losses and metrics

$$LR_{fidelity}(LR, SR) = |LR - \underbrace{LPF_{PSF}(SR)\downarrow_2}_{\textbf{Coarse SR}}|_1 \quad HR_{fidelity}(HR, SR) = |\underbrace{(HR - LPF_{PSF}(HR))}_{\textbf{High freq. details of HR}} - \underbrace{(SR - LPF_{PSF}(SR))}_{\textbf{High freq. details of SR}}|_1$$

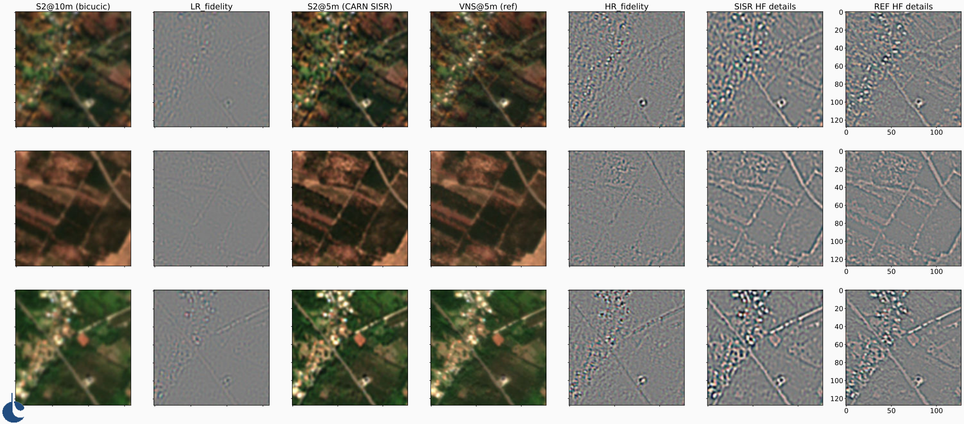
CARN trained with Smooth L1 Loss



Quality Metrics, option A : physically informed losses and metrics

$$LR_{fidelity}(LR, SR) = |LR - \underbrace{LPF_{PSF}(SR)\downarrow_2}_{\textbf{Coarse SR}}|_1 \qquad HR_{fidelity}(HR, SR) = |\underbrace{(HR - LPF_{PSF}(HR))}_{\textbf{High freq. details of HR}} - \underbrace{(SR - LPF_{PSF}(SR))}_{\textbf{High freq. details of SR}}|_1$$

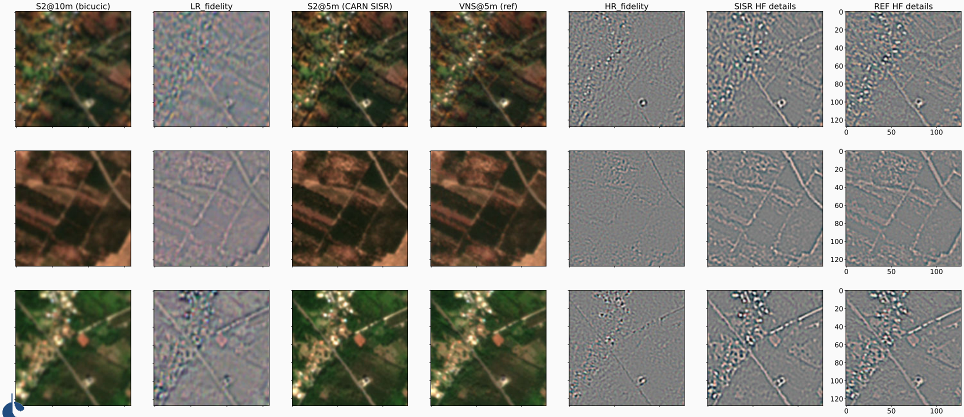
CARN trained with $LR_{fidelity}$ and $HR_{fidelity}$



Quality Metrics, option A : physically informed losses and metrics

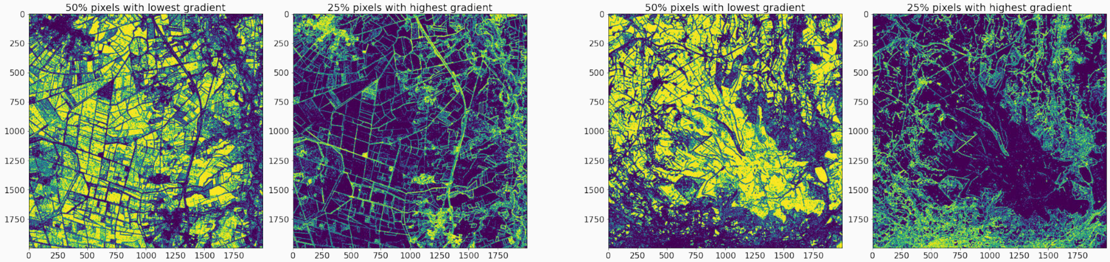
$$LR_{fidelity}(LR, SR) = |LR - \underbrace{LPF_{PSF}(SR)\downarrow_2}_{\text{Coarse SR}}|_1 \qquad HR_{fidelity}(HR, SR) = |\underbrace{(HR - LPF_{PSF}(HR))}_{\text{High freq. details of HR}} - \underbrace{(SR - LPF_{PSF}(SR))}_{\text{High freq. details of SR}}|_1$$

CARN trained with Perceptual Loss (VGG) + Smooth L1 loss on simulated S2



Quality Metrics, option B : using standard metrics and stratifying gradients

- Low gradient magnitude pixels (homogeneous, low texture) outnumber high gradient magnitude pixels (edges, high texture)
- $\Rightarrow$ statistical measures (RMSE, MAE, PSNR . . .) are biased toward low gradient magnitude pixels accuracy
- $\Rightarrow$ poor measures of super-resolution performances (will favor low resolution fidelity)
- Idea : apply metrics separately on higest / lowest gradient magnitude stratas



Super Resolution in EO: Best practices for performance and quality metrics

Data for training: current practices, challenges, future pathways

Michal Kawulok[1,2]

(✉) michal.kawulok@ieee.org

[1]KP Labs, Gliwice, Poland
[2]Silesian University of Technology, Gliwice, Poland

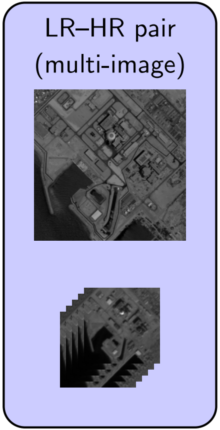
Data for training super-resolution networks

- Data required to train super-resolution (SR) networks
  - pairs of high-resolution reference coupled with low-resolution observation
  - single-image SR: 1 LR image coupled with 1 HR image
  - multi-image SR: N LR images coupled with 1 HR image
  - an image composed of a single or multiple bands (for multi- or hyperspectral data)
- Data source
  - real HR and LR image(s) showing the same region of interest
  - LR image(s) simulated from the HR image

LR–HR pair

Data for training super-resolution networks

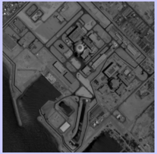
- Data required to train super-resolution (SR) networks
  - pairs of high-resolution reference coupled with low-resolution observation
  - single-image SR: 1 LR image coupled with 1 HR image
  - multi-image SR: N LR images coupled with 1 HR image
  - an image composed of a single or multiple bands (for multi- or hyperspectral data)
- Data source
  - real HR and LR image(s) showing the same region of interest
  - LR image(s) simulated from the HR image

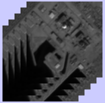


Data for training super-resolution networks

- Data required to train super-resolution (SR) networks
  - pairs of high-resolution reference coupled with low-resolution observation
  - single-image SR: 1 LR image coupled with 1 HR image
  - multi-image SR: N LR images coupled with 1 HR image
  - an image composed of a single or multiple bands (for multi- or hyperspectral data)
- Data source
  - real HR and LR image(s) showing the same region of interest
  - LR image(s) simulated from the HR image

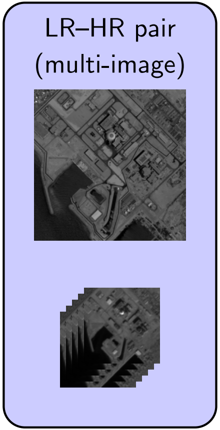
LR–HR pair (multi-image)





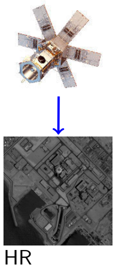
Data for training super-resolution networks

- Data required to train super-resolution (SR) networks
  - pairs of high-resolution reference coupled with low-resolution observation
  - single-image SR: 1 LR image coupled with 1 HR image
  - multi-image SR: N LR images coupled with 1 HR image
  - an image composed of a single or multiple bands (for multi- or hyperspectral data)
- Data source
  - real HR and LR image(s) showing the same region of interest
  - LR image(s) simulated from the HR image



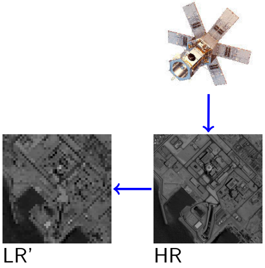
Data for training super-resolution networks

- Simulated LR data
  - easy to obtain
  - may not reflect the real LR–HR relation
  - trained models produce artifacts when applied to real images
- Real LR–HR image pairs
  - acquired using a different sensor of lower resolution
  - difficult and costly to collect
  - models suitable for operating conditions
  - example: Proba-V dataset
- Potential solution: realistic LR image simulation



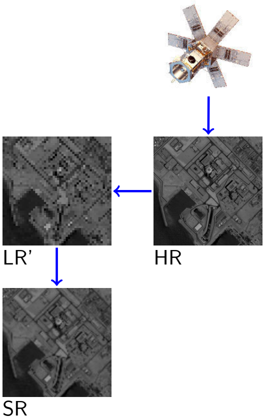
Data for training super-resolution networks

- Simulated LR data
  - easy to obtain
  - may not reflect the real LR–HR relation
  - trained models produce artifacts when applied to real images
- Real LR–HR image pairs
  - acquired using a different sensor of lower resolution
  - difficult and costly to collect
  - models suitable for operating conditions
  - example: Proba-V dataset
- Potential solution: realistic LR image simulation



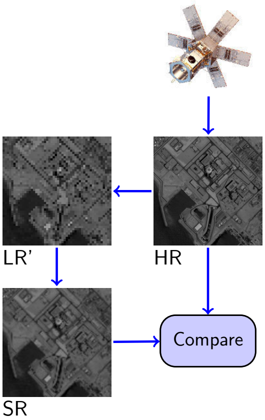
Data for training super-resolution networks

- Simulated LR data
  - easy to obtain
  - may not reflect the real LR–HR relation
  - trained models produce artifacts when applied to real images
- Real LR–HR image pairs
  - acquired using a different sensor of lower resolution
  - difficult and costly to collect
  - models suitable for operating conditions
  - example: Proba-V dataset
- Potential solution: realistic LR image simulation



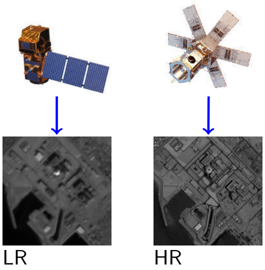
Data for training super-resolution networks

- Simulated LR data
  - easy to obtain
  - may not reflect the real LR–HR relation
  - trained models produce artifacts when applied to real images
- Real LR–HR image pairs
  - acquired using a different sensor of lower resolution
  - difficult and costly to collect
  - models suitable for operating conditions
  - example: Proba-V dataset
- Potential solution: realistic LR image simulation



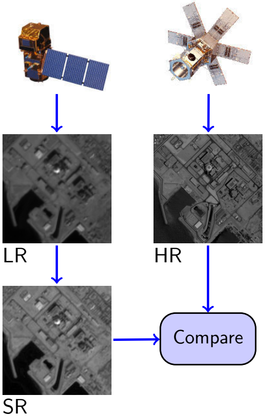
Data for training super-resolution networks

- Simulated LR data
  - easy to obtain
  - may not reflect the real LR–HR relation
  - trained models produce artifacts when applied to real images
- Real LR–HR image pairs
  - acquired using a different sensor of lower resolution
  - difficult and costly to collect
  - models suitable for operating conditions
  - example: Proba-V dataset
- Potential solution: realistic LR image simulation



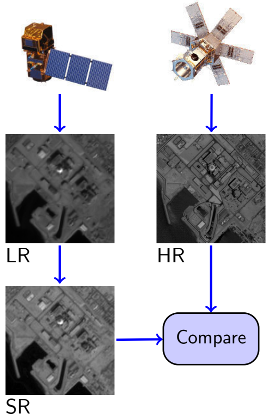
Data for training super-resolution networks

- Simulated LR data
  - easy to obtain
  - may not reflect the real LR–HR relation
  - trained models produce artifacts when applied to real images
- Real LR–HR image pairs
  - acquired using a different sensor of lower resolution
  - difficult and costly to collect
  - models suitable for operating conditions
  - example: Proba-V dataset
- Potential solution: realistic LR image simulation



Data for training super-resolution networks

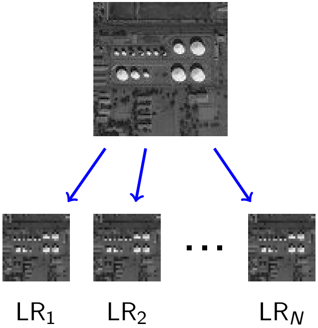
- Simulated LR data
  - easy to obtain
  - may not reflect the real LR–HR relation
  - trained models produce artifacts when applied to real images
- Real LR–HR image pairs
  - acquired using a different sensor of lower resolution
  - difficult and costly to collect
  - models suitable for operating conditions
  - example: Proba-V dataset
- Potential solution: realistic LR image simulation



Semi-simulated data for multi-image SR

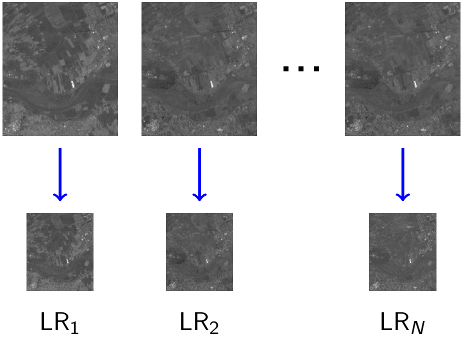
Simulated LR images

Original image



Semi-simulated LR images

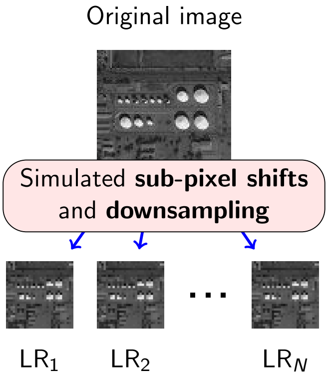
Original images



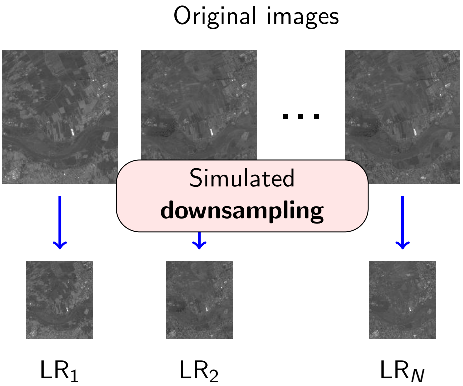
Source: T. Tarasiewicz, J. Nalepa, M. Kawulok: Semi-simulated training data for multi-image super-resolution, IEEE IGARSS 2022

Semi-simulated data for multi-image SR

Simulated LR images

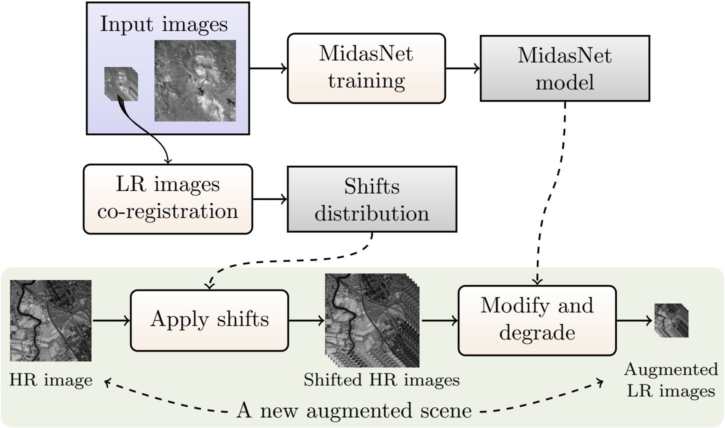


Semi-simulated LR images



Source: T. Tarasiewicz, J. Nalepa, M. Kawulok: Semi-simulated training data for multi-image super-resolution, IEEE IGARSS 2022

Data augmentation for LR image simulation



Source: M. Ziaja, J. Nalepa, M. Kawulok: Data augmentation for multi-image super-resolution, IEEE IGARSS 2022

Thank you for your attention!

Data for training: current practices, challenges, future pathways

Michal Kawulok[1,2]

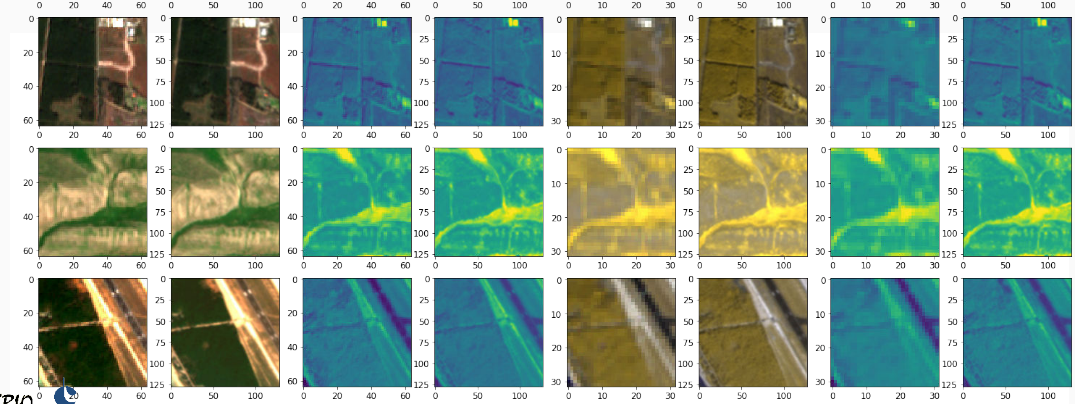
(✉) michal.kawulok@ieee.org

[1]KP Labs, Gliwice, Poland
[2]Silesian University of Technology, Gliwice, Poland

Shameless advertisement : the SEN2VENµS Dataset (132k patches)

- Open Dataset on Zenodo : `https://zenodo.org/record/6514159`
- 132 955 patches, 29 locations, 8 Sentinel-2 bands, from 10m/20m to 5m (VENµS resolution)

Michel, J. ; Vinasco-Salinas, J. ; Inglada, J. ; Hagolle, O. SEN2VENµS, a Dataset for the Training of Sentinel-2 Super-Resolution Algorithms.
Preprints 2022, 2022050230 (doi : 10.20944/preprints202205.0230.v1). *Submitted to MDPI Data.*



Quality indices & xAI for trusted and accountable SR

Several ways to assess and viz SR quality

1. **Uncertainty maps** for xAI [Kendall 2017]
   - show what the model doesn't know → natural fit to ill-posedness of SR
2. **Perceptual metrics** [Johnson 2016, Ledig 2017, Laparra 2010, Talens 2010]
   - traditionally applied on SR (or PAN) products [Wu 2017]
     - in space or spectrum, metric-wise or perceptual-guided
3. **Saliency** - rooted on information & uncertainty [Zhang 2017]
   - Now: perceptual on output → a statistical approach could scrutinize any layer
4. **Distortions** & **hallucinations**
   - Now: tendency evaluated at the model level → Future: layer level
   - study latent representations; what their filters do; their distribution; activation under different stimuli
5. **Attention mechanisms** [Vaswani 2017, Dosovitskiy 2020]
   - alternatives to standard extraction and metric analysis
   - they study the input-latent-output chain jointly

References

[Arrieta 2020]
Arrieta et al., **Explainable Artificial Intelligence (xAI): Concepts, taxonomies, opportunities and challenges toward responsible AI**, Information Fusion, 58, 82-115, (2020).

[Campos-Taberner 2020]
Campos-Taberner et al., **Understanding deep learning in land use classification from Sentinel-2 time series**. Scientific Reports 12 (2020).

[Tuia 2021]
Tuia, et al., **Towards a Collective Agenda on AI for Earth Science Data Analysis**. IEEE Geoscience and Remote Sensing Magazine (2021).

[Samek 2021]
Samek, et al., **Explaining deep neural networks and beyond: A review of methods and applications**. Proceedings of the IEEE, 109(3), 247-278 (2021).

[Zhang 2017]
Zhang, et al., **Learning uncertain convolutional features for accurate saliency detection**. ICCV 2017.