THE OHIO STATE UNIVERSITY

Comparison of statistical and deep learning methods for spatio-temporal fusion of satellite images

Rohit Mukherjee, PhD Student, The Ohio State University

Prof. Desheng Liu, The Ohio State University

AAG 2019, Washington DC

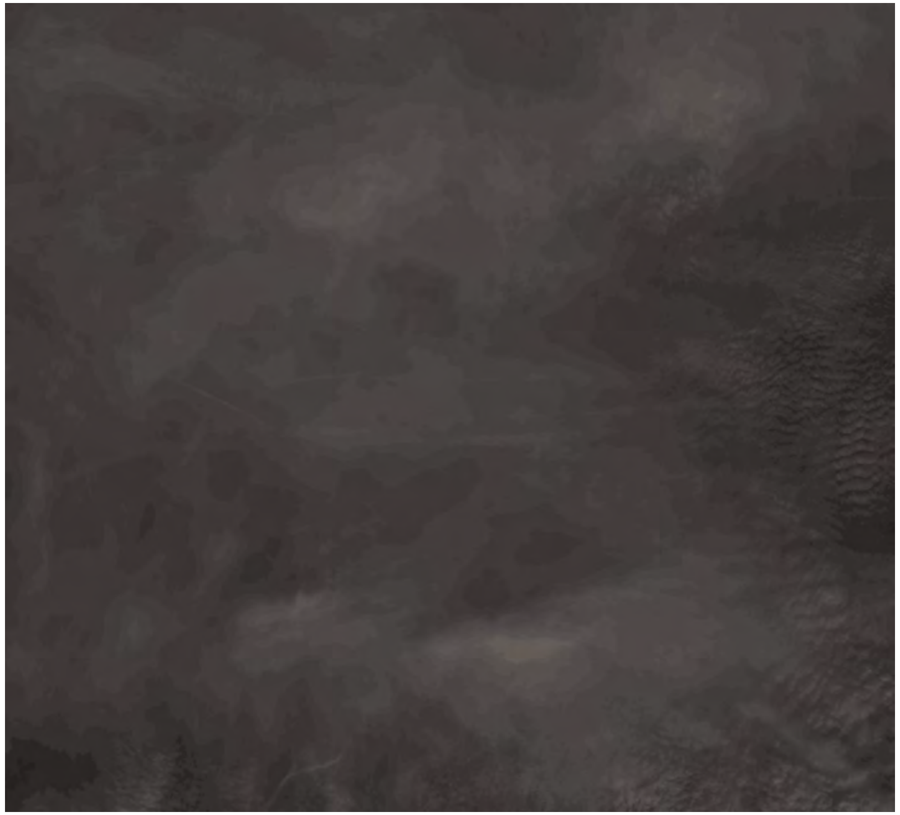
Continuous Land Surface Monitoring

Monitoring land use and land cover change globally is a major challenge for the remote sensing community.

- MODIS Land Cover Products
- Global Land Survey (Landsat)

Issues

- Spatial-temporal resolution trade-off
- Cloud Contamination

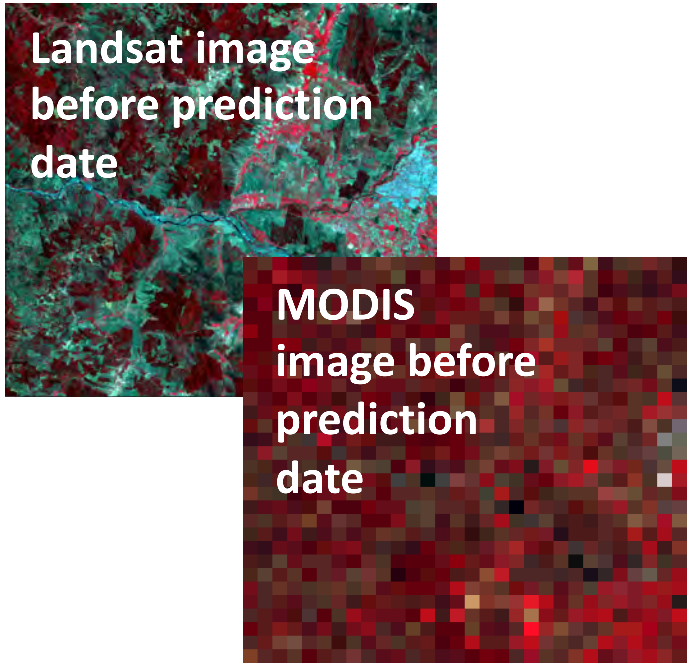


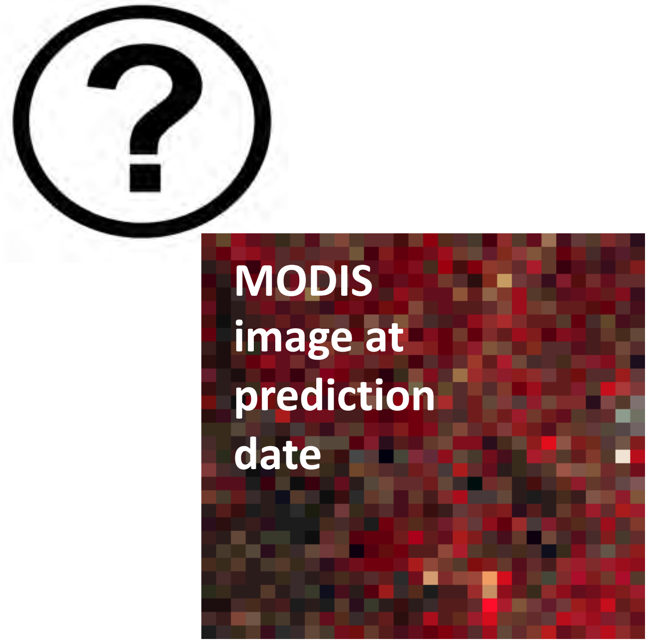
How to solve these issues?

Unfortunately, with passive remote sensing we cannot mitigate cloud contamination.

Spatial Temporal trade-off

- Spatio-temporal Image Fusion or STIF methods
- Single Image Downscaling





Existing Methods

STIF Methods

- Spatio-temporal adaptive reflectance model (STARFM) (Gao et al., 2006)
- *Flexible Spatio-temporal Data Fusion (FSDAF)* (Zhu et al., 2016)

Statistical Downscaling Methods

- Bilinear Interpolation (BI)
- Thin Plate Splines (TPS)

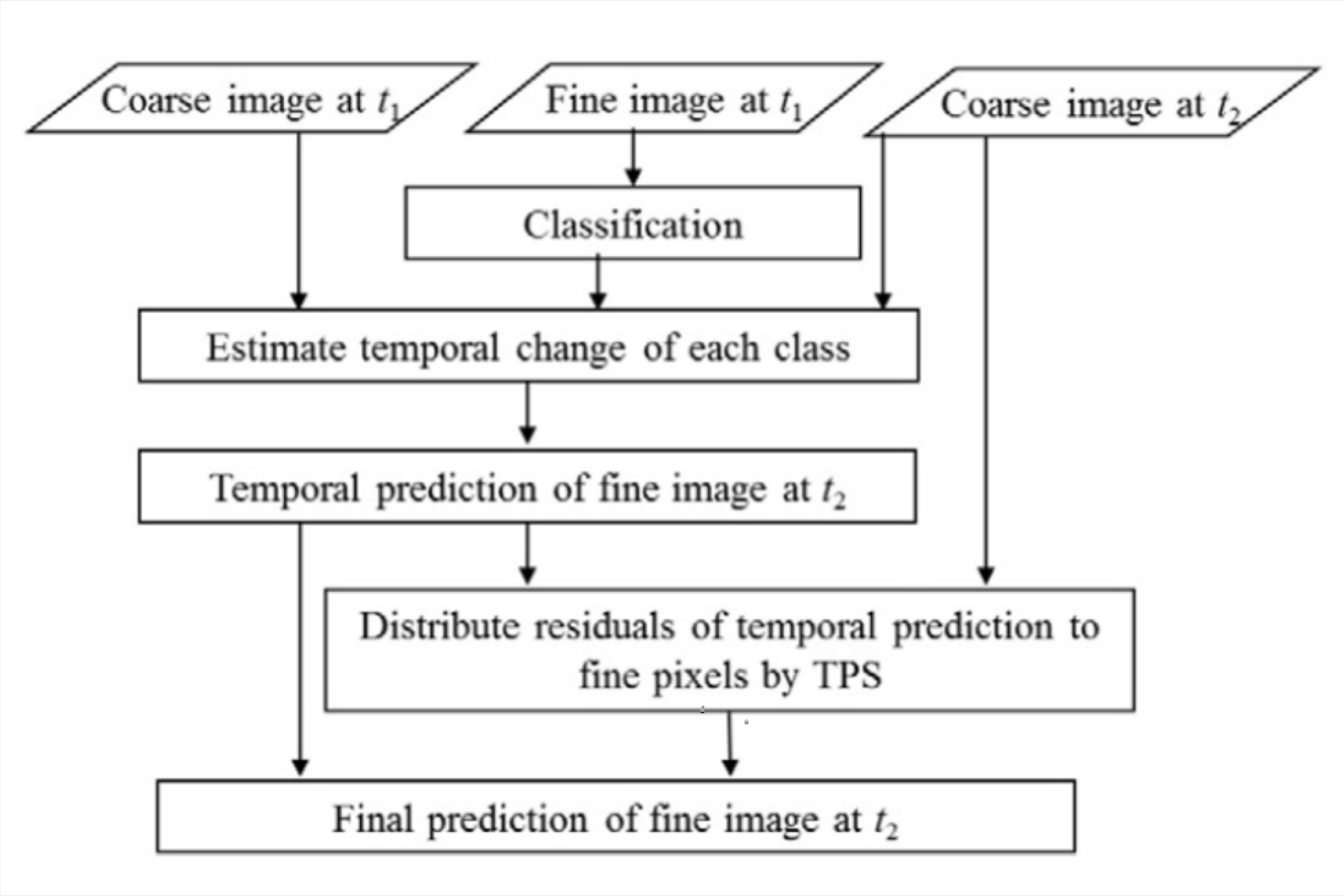
Deep Learning based Superresolution

- *Superresolution Convolutional Neural Network (SRCNN)* (Dong et al., 2014)
- U-Net (Ronneberger et al., 2015)

Existing Methods

STIF Methods

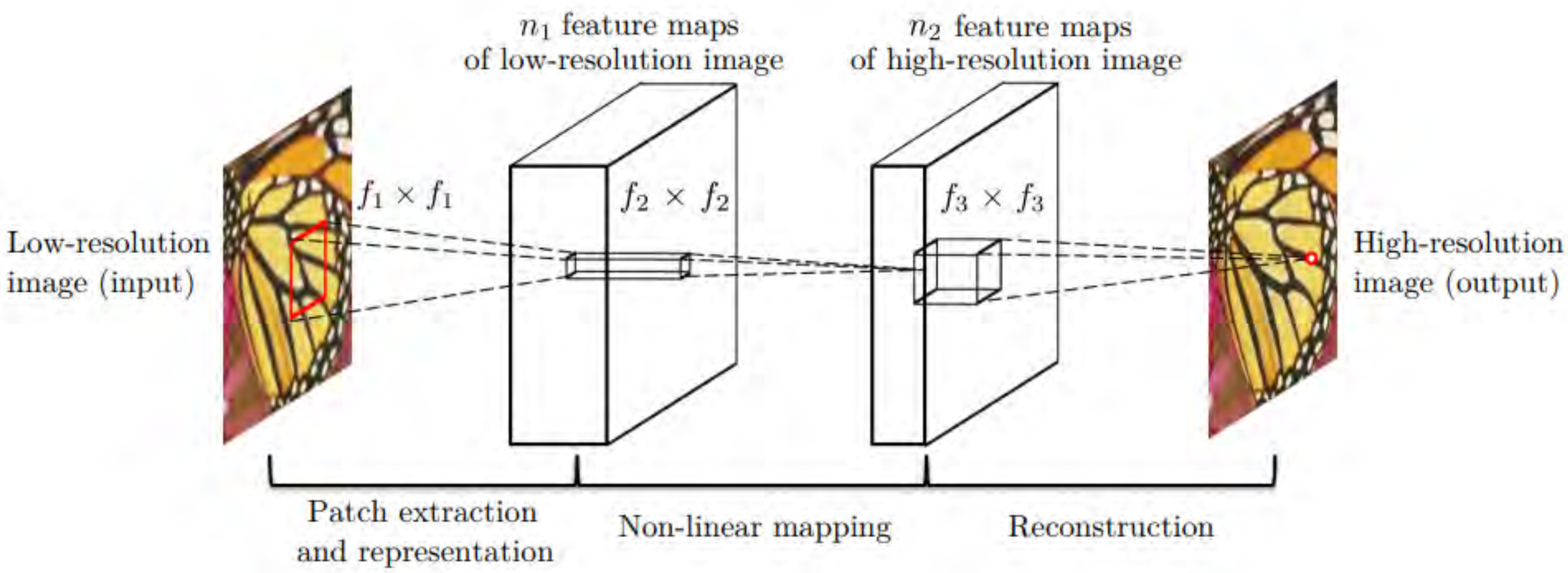
Flexible Spatio-temporal Data Fusion (FSDAF)



Existing Methods

Deep Learning based Superresolution Method

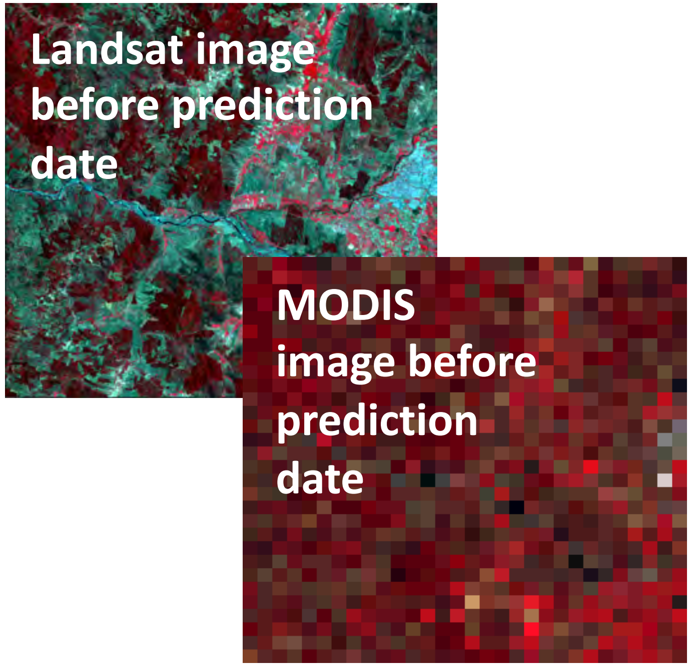
Superresolution Convolutional Neural Network (SRCNN)

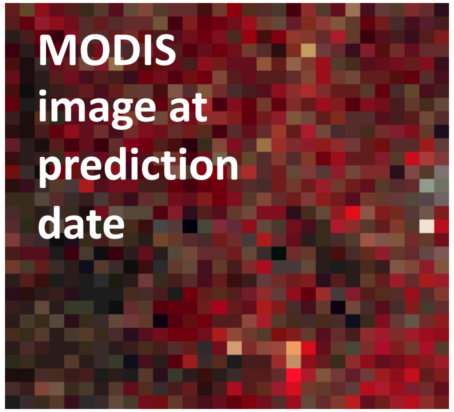


Temporal Information

- **Incorporates temporal information –** *Spatio-temporal image fusion methods (FSDAF)*
- **Does NOT incorporate temporal information –** *Statistical Downscaling and Deep learning based superresolution methods*

Low Temporal Change

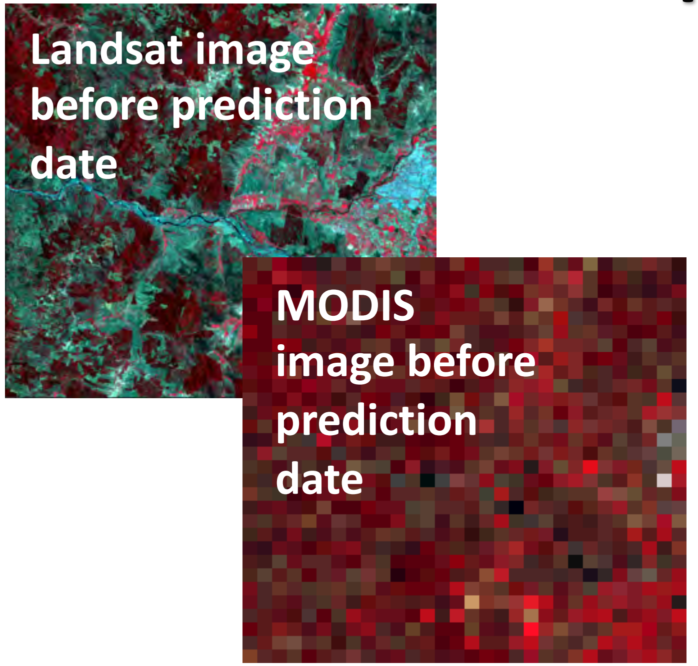


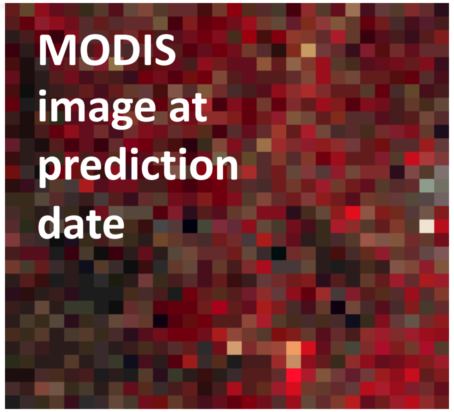


Temporal Information

- **Incorporates temporal information –** *Spatio-temporal image fusion methods (FSDAF)*
- **Does NOT incorporate temporal information –** *Statistical Downscaling and Deep learning based superresolution methods*

Medium Temporal Change

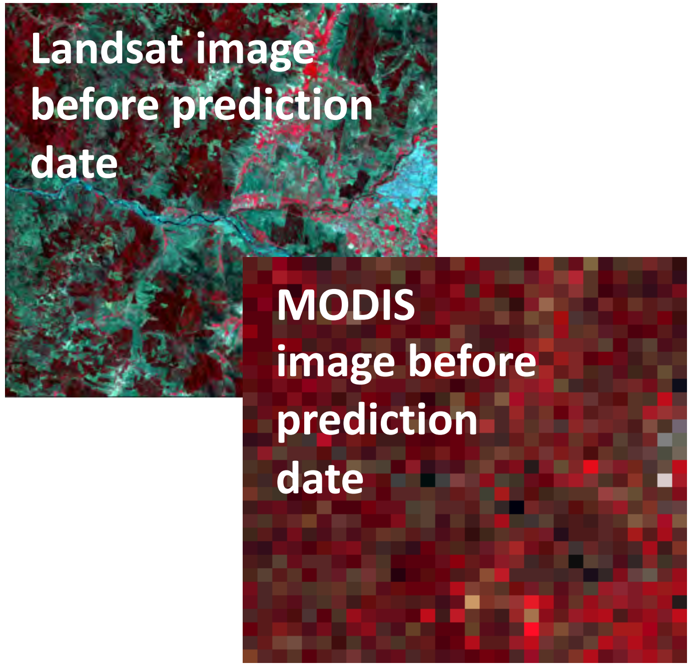


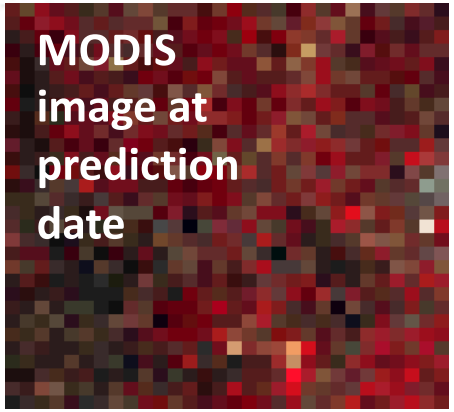


Temporal Information

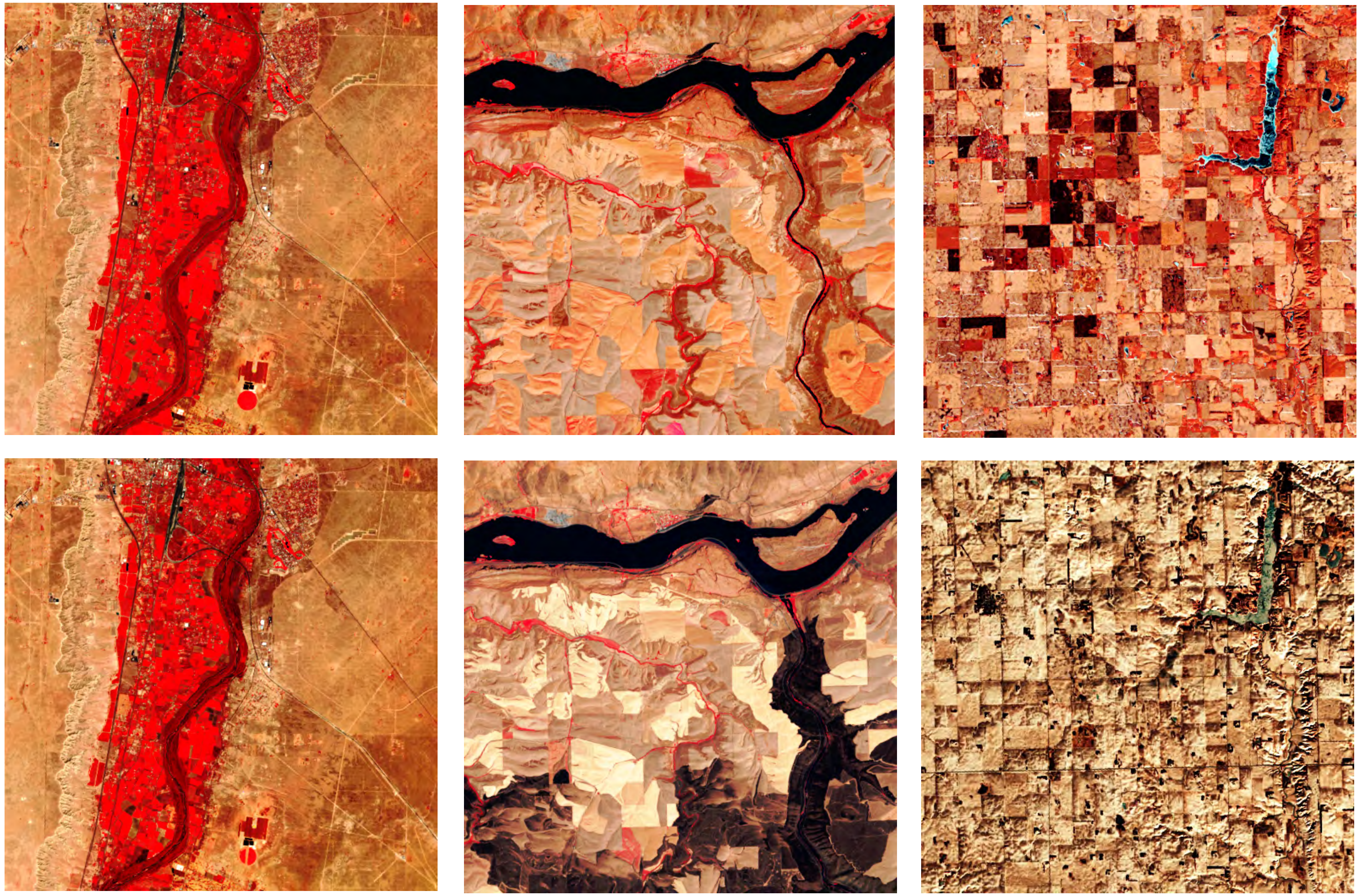
- **Incorporates temporal information –** *Spatio-temporal image fusion methods (FSDAF)*
- **Does NOT incorporate temporal information –** *Statistical Downscaling and Deep learning based superresolution methods*

High Temporal Change





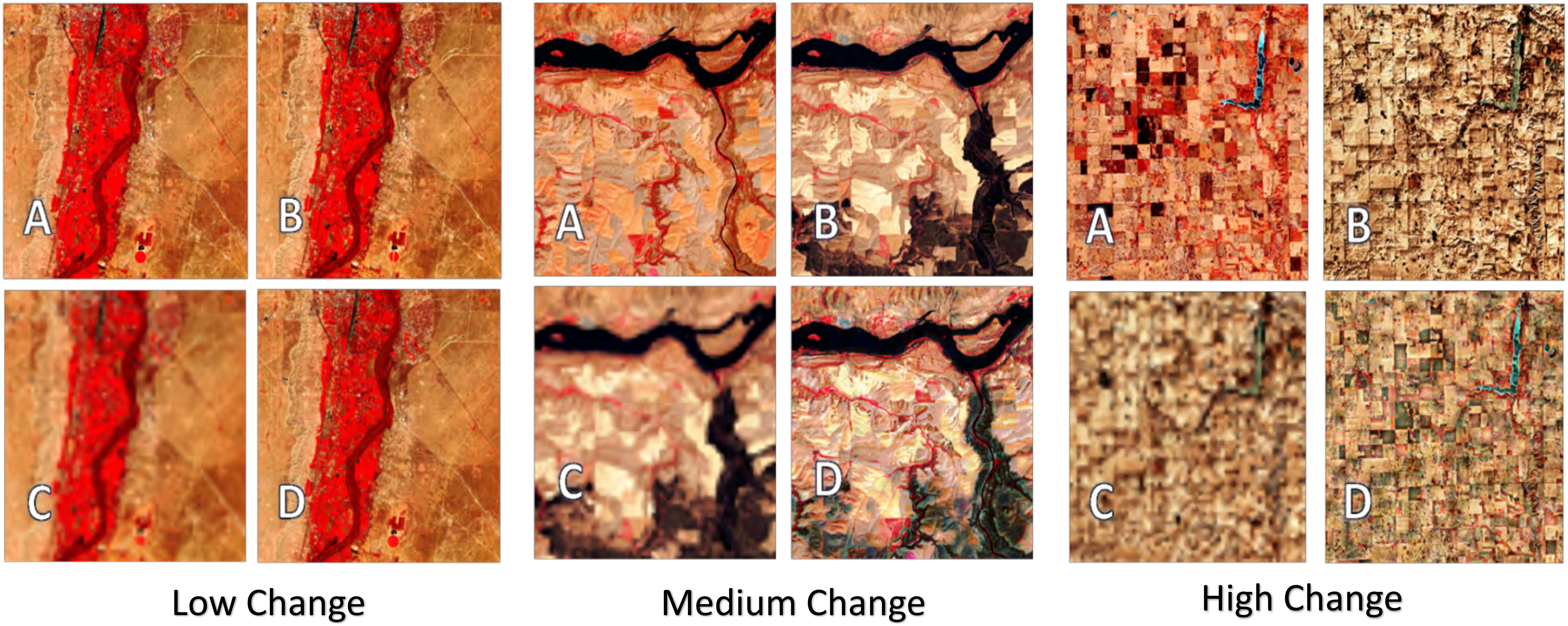
Data



Low Change — Medium Change — High Change

	Scene 1	Scene 2	Scene 3
CC	0.9869	0.7055	0.5364
SSIM	0.9856	0.7441	0.5866
Change	Low	Medium	High

Results



A – Original Landsat 8 image prior change, **B** – Original Landsat 8 image to be predicted, **C** – SRCNN, **D** - FSDAF

Results

Low Change

	TPS	FSDAF	SRCNN	BI
CC	0.9470	**0.9886**	0.9626	**0.9449**
PSNR	20.25	**26.69**	21.74	**20.41**
RMSE	175.33	**152.24**	181.42	**178.21**
SSIM	0.9390	**0.9859**	0.9502	**0.9413**

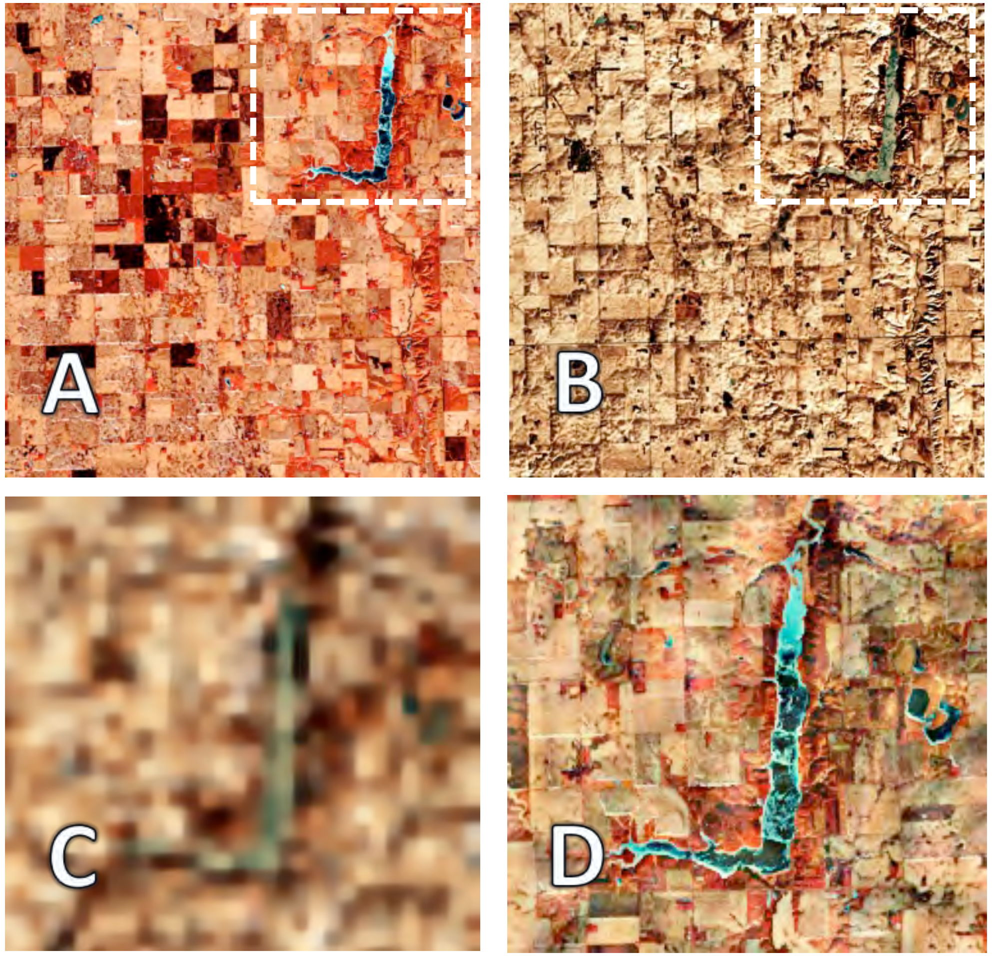
	TPS	FSDAF	SRCNN	BI
CC	0.9357	**0.9024**	**0.9578**	0.9427
PSNR	19.37	**17.58**	**21.25**	19.97
RMSE	185.82	**179.26**	**174.46**	174.82
SSIM	0.7846	**0.7312**	**0.8295**	0.8150

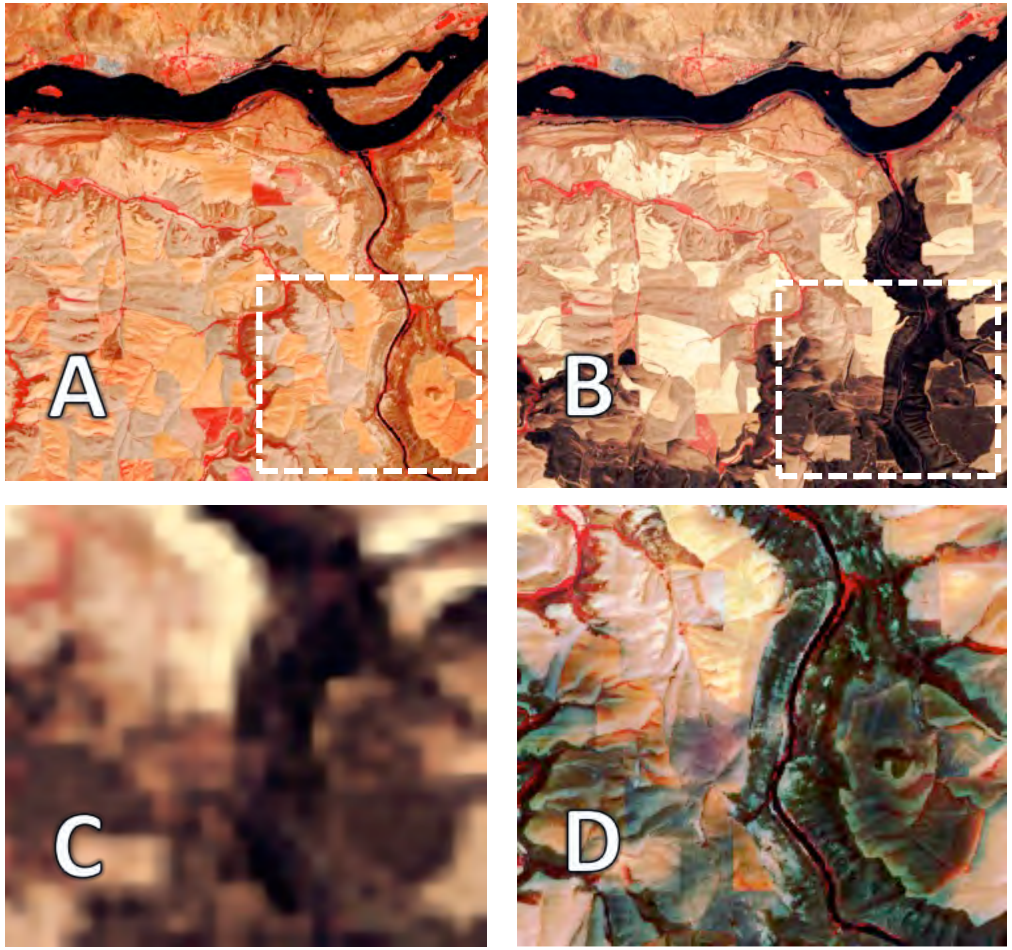
Medium Change

	TPS	FSDAF	SRCNN	BI
CC	0.6502	**0.6264**	**0.7563**	0.6615
PSNR	12.94	**12.67**	**14.44**	13.08
RMSE	175.61	**176.47**	**176.31**	176.88
SSIM	0.6064	**0.5917**	**0.6372**	0.6143

High Change

Sharpness vs Accuracy





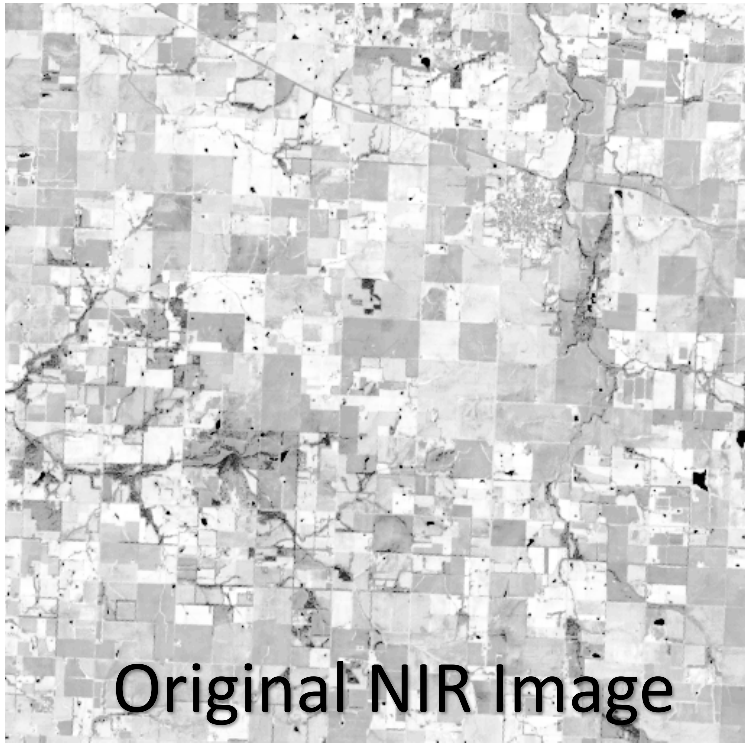
A – Original Landsat 8 image prior change, **B** – Original Landsat 8 image to be predicted, **C** – SRCNN, **D** - FSDAF

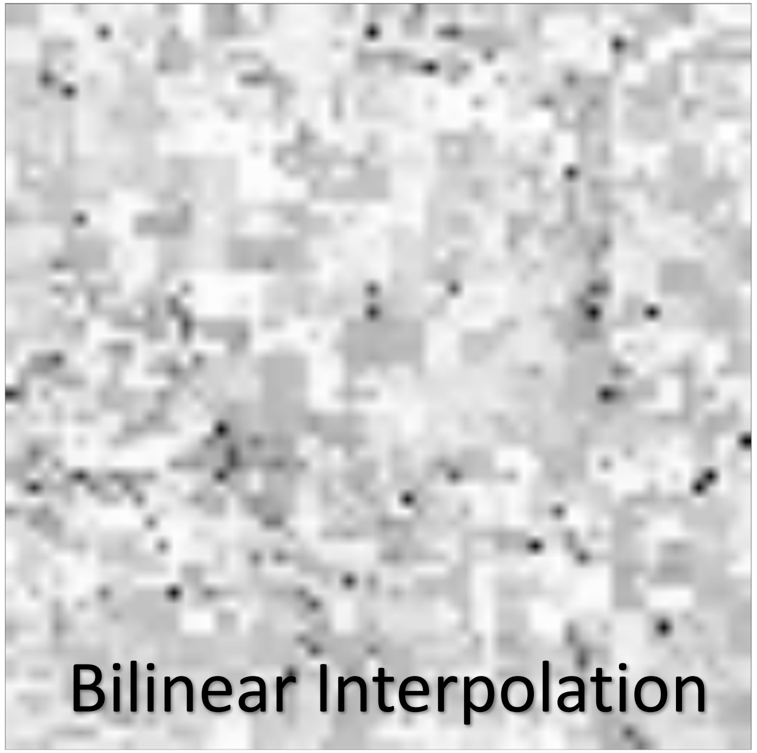
Conclusion: What is your objective?

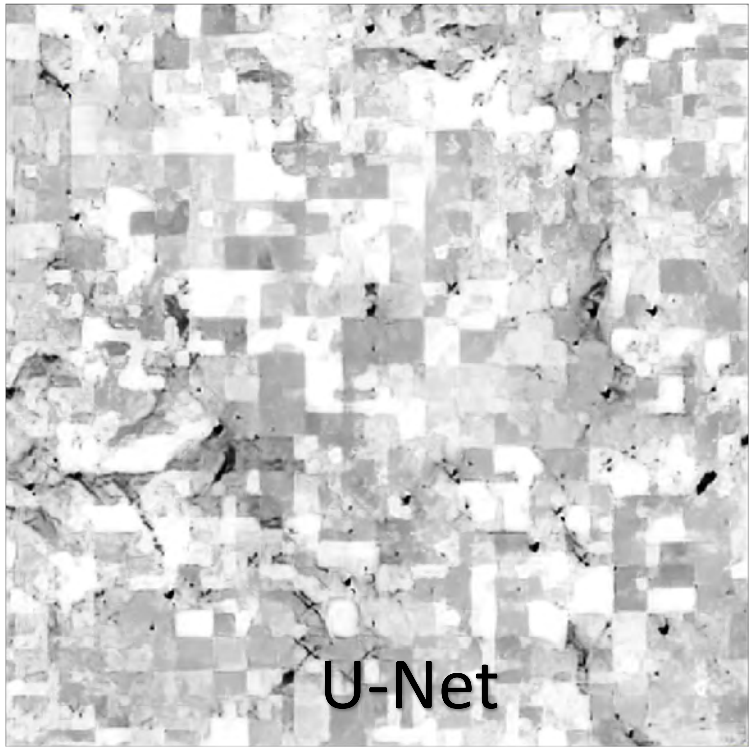
- Incorporating temporal information *reduces* the accuracy for predicting scenes undergoing *moderate to high degree of land cover change*
- In regions experiencing *abrupt changes* (flood/forest fire), a *single image downscaling method* or a *deep learning superresolution method* is more accurate
- *Prior temporal information* carries *erroneous land cover information* due to the *abrupt change*
- Finding an *additional pair of low- and high- resolution images* is challenging due to *cloud cover* and *lower temporal resolution* of the higher spatial resolution satellite image sensor
- Spatio-temporal image fusion (STIF) methods could be used for generating a more accurate time-series data for *recording gradual changes* in a region

Ongoing/Future Work

- **Transfer learning U-Net superresolution model with satellite imagery to predict *sharper* high-resolution satellite imagery**
- Incorporating temporal information to improve prediction accuracy







References

- Zhu, Xiaolin, et al. "An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions." Remote Sensing of Environment 114.11 (2010): 2610-2623.
- Zhu, Xiaolin, et al. "A flexible spatiotemporal method for fusing satellite images with different resolutions." Remote sensing of environment 172 (2016): 165-177.
- Yang, Wenming, et al. "Deep learning for single image super-resolution: A brief review." arXiv preprint arXiv:1808.03344 (2018).
- Wulder, Michael A., et al. "Virtual constellations for global terrestrial monitoring." Remote Sensing of Environment 170 (2015): 62-76.
- Dong, Chao, et al. "Learning a deep convolutional network for image super-resolution." European conference on computer vision. Springer, Cham, 2014.
- FastAI, Practical Deep Learning for Coders
- Ronnerberger, Olaf, et al. “U-Net: Convolutional Networks for Biomedical Image Segmentation” (2015)

Thank You!

Any Questions?