2nd NOAA Workshop on Leveraging AI in Environmental Sciences
- Session 26: AI/ML for Information Extraction from Data, Part 2

Kick: Shift-N-Overlap Cascades of Transposed Convolutional Layer for Better Autoencoding Reconstruction on Remote Sensing Imagery

Seungkyun Hong, Ph.D. (candidate)
Korea University of Science and Technology (UST)
Korea Institute of Science and Technology Information (KISTI)

[Contents]

00 | Before we begin…

This presentation

- May discuss contents **previously published in IEEE© Access journal (Hong and Song, 2020):** IEEEXplore: #9110496, doi: [10.1109/ACCESS.2020.3000557](https://doi.org/10.1109/ACCESS.2020.3000557)
- Describe an **impact of deconvolutional layers in unsupervised autoencoder modeling**

IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received April 23, 2020, accepted May 24, 2020, date of publication June 8, 2020, date of current version June 18, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3000557

Kick: Shift-N-Overlap Cascades of Transposed Convolutional Layer for Better Autoencoding Reconstruction on Remote Sensing Imagery

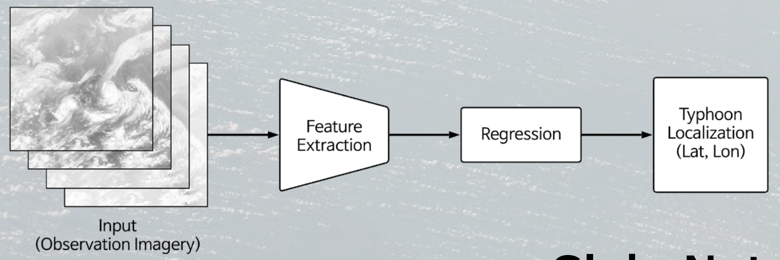
SEUNGKYUN HONG, (Member, IEEE), AND SA-KWANG SONG
Department of Data and HPC Science, Korea University of Science and Technology (UST), Daejeon 34113, South Korea
Research Data Sharing Center, Korea Institute of Science and Technology Information (KISTI), Daejeon 34141, South Korea
Corresponding author: Sa-Kwang Song (esmallj@kisti.re.kr)

This work was supported in part by the Korea Institute of Science and Technology Information (KISTI) through the Project "Establishing a System for Sharing and Disseminating Research Data" under Grant K-20-L01-C04, and in part by the Korea National Supercomputing Center with supercomputing resources under Grant KSC-2019-CRE-0257.

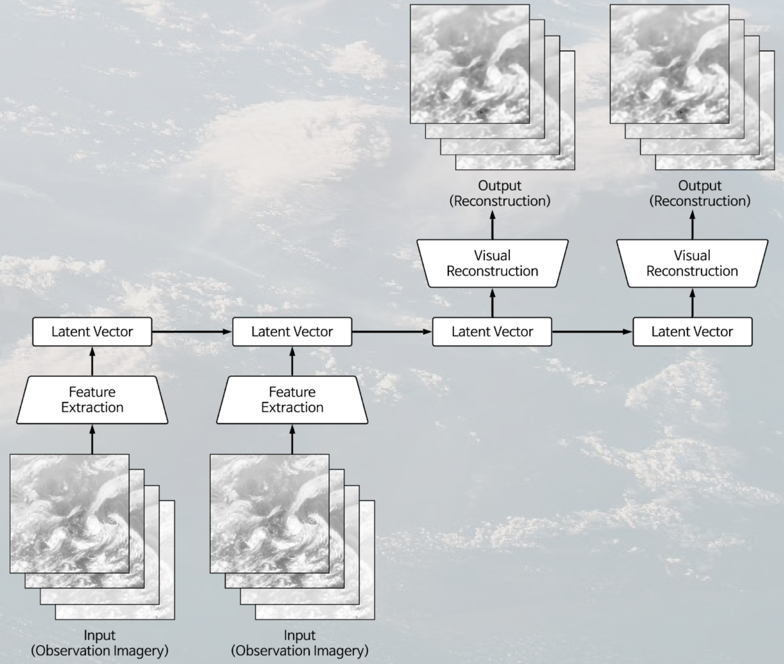
01 | Background

How this research began?

- Fresh attempt to **'extract and predict'** upcoming atmospheric events, from large-scale remote sensing imagery **'rapidly'**
- Preliminary model surveys for **meso-scale weather prediction tasks**
  - **Typhoon Eye Localization**
    (GlobeNet: [arXiv:1708.03417](https://arxiv.org/abs/1708.03417))
  - **Next Sequence of Satellite Images Prediction**
    (PSIque: [arXiv:1711.10644](https://arxiv.org/abs/1711.10644))



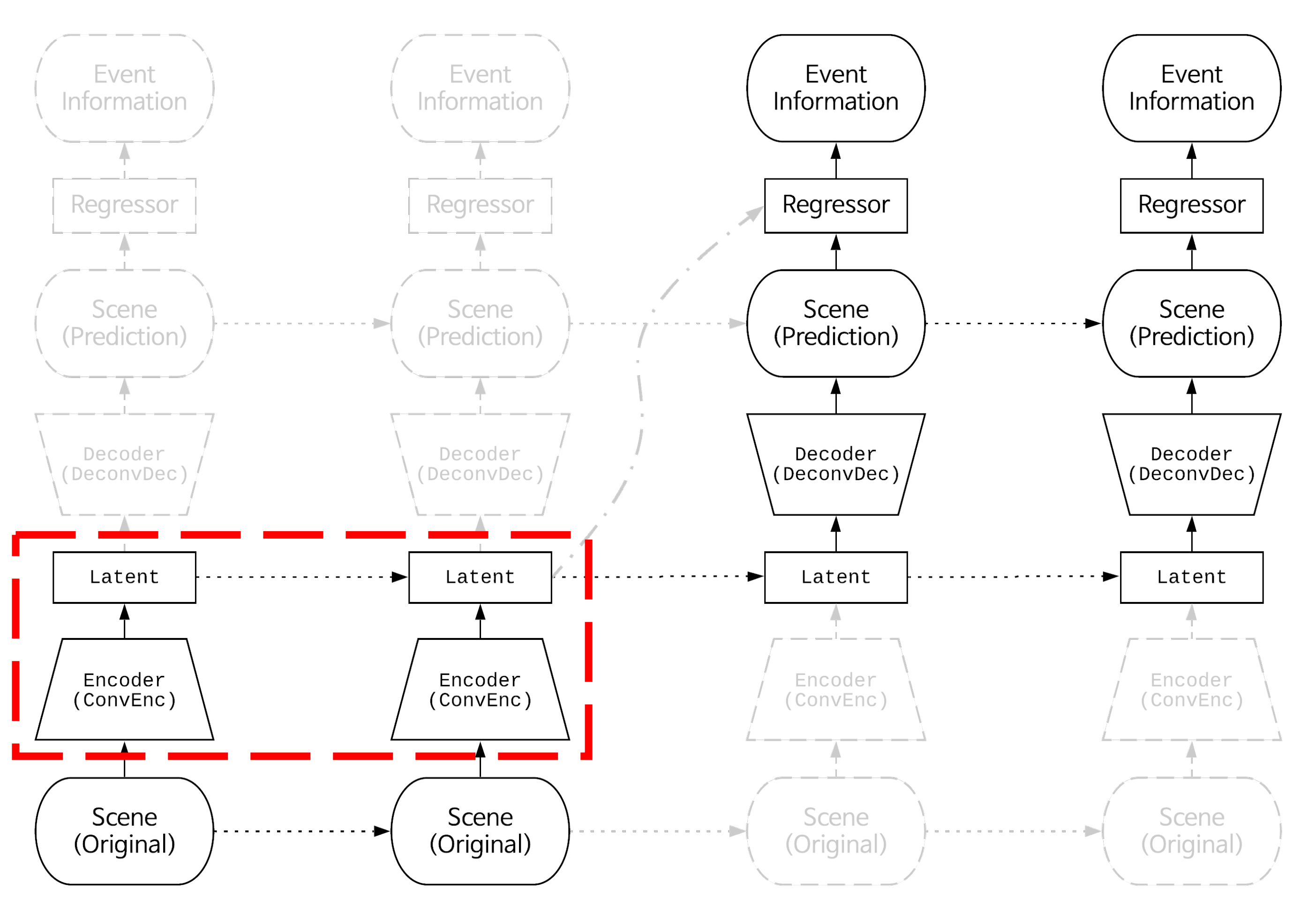
GlobeNet



PSIque

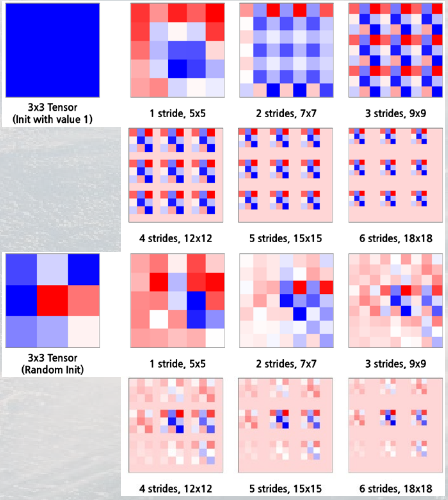
01 | Background

- **Issue #1**: Correlated models **share common part** which trained **separately**

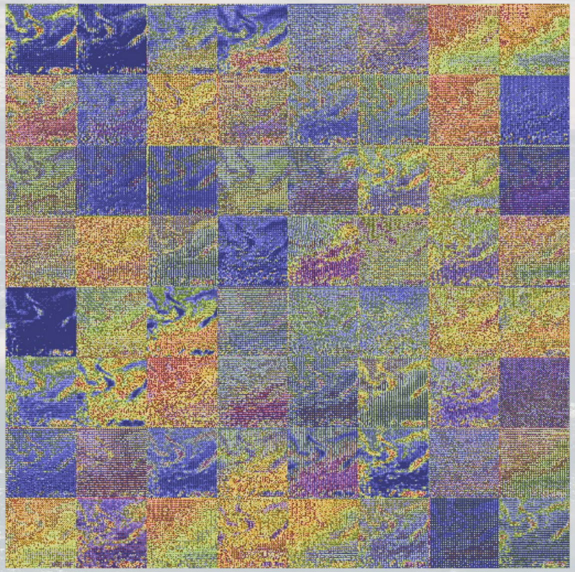


01 | Background

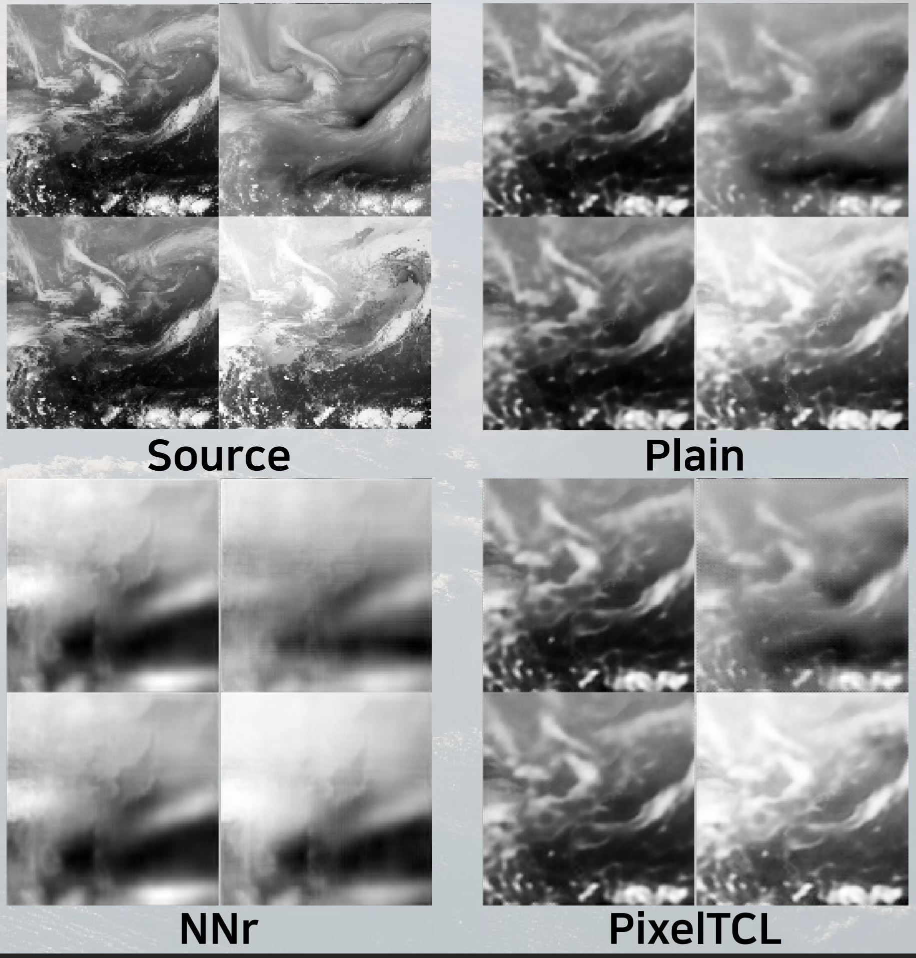
- **Issue #2**: A **deconvolution** process **matters** on **reconstruction quality**



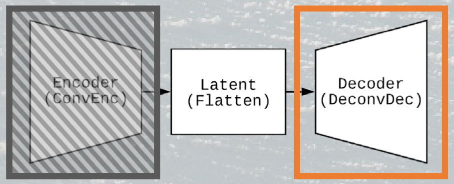
Checkerboard artifacts from deconv.



Reconstruction results from deconv.

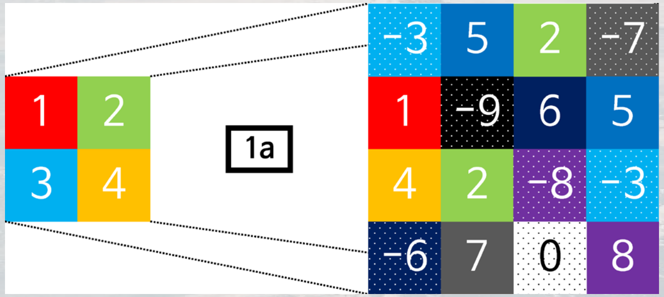


- **Every deconvolutional process MUST produce checkerboard artifacts**
- Furthermore, **variations of deconvolutional** process on decoder provide **considerable differences with fixed encoder operation** during End-to-End learning
- Therefore, **reconstruction quality** in **convolutional-based autoencoder** is **highly dependent on decoder**

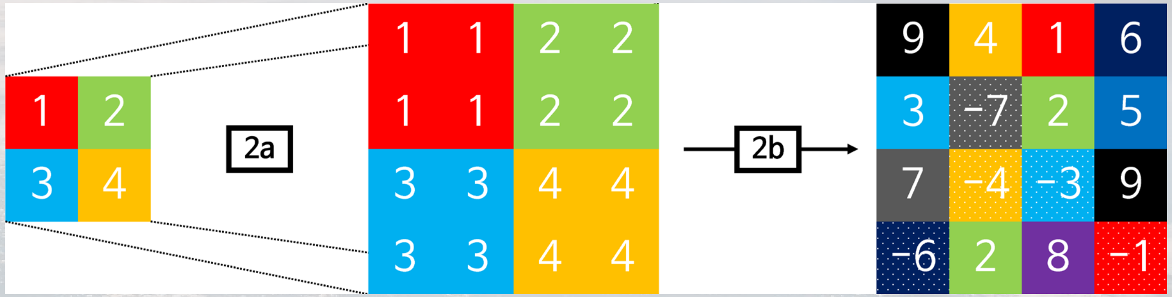


02 | Methods | Related works

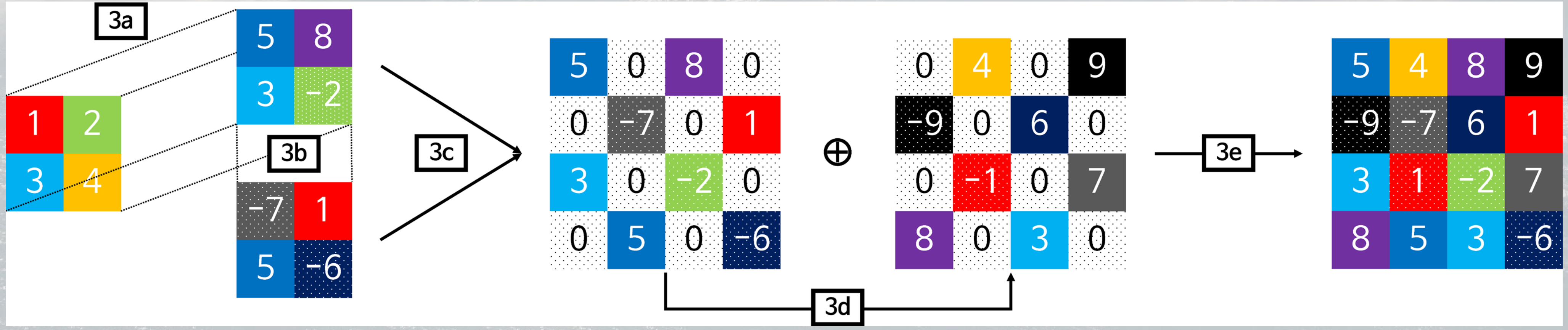
Plain Deconvolution
(Zeiler et al., 2010)



Nearest-Neighbor Resize Deconvolution
(Odena et al., 2016)

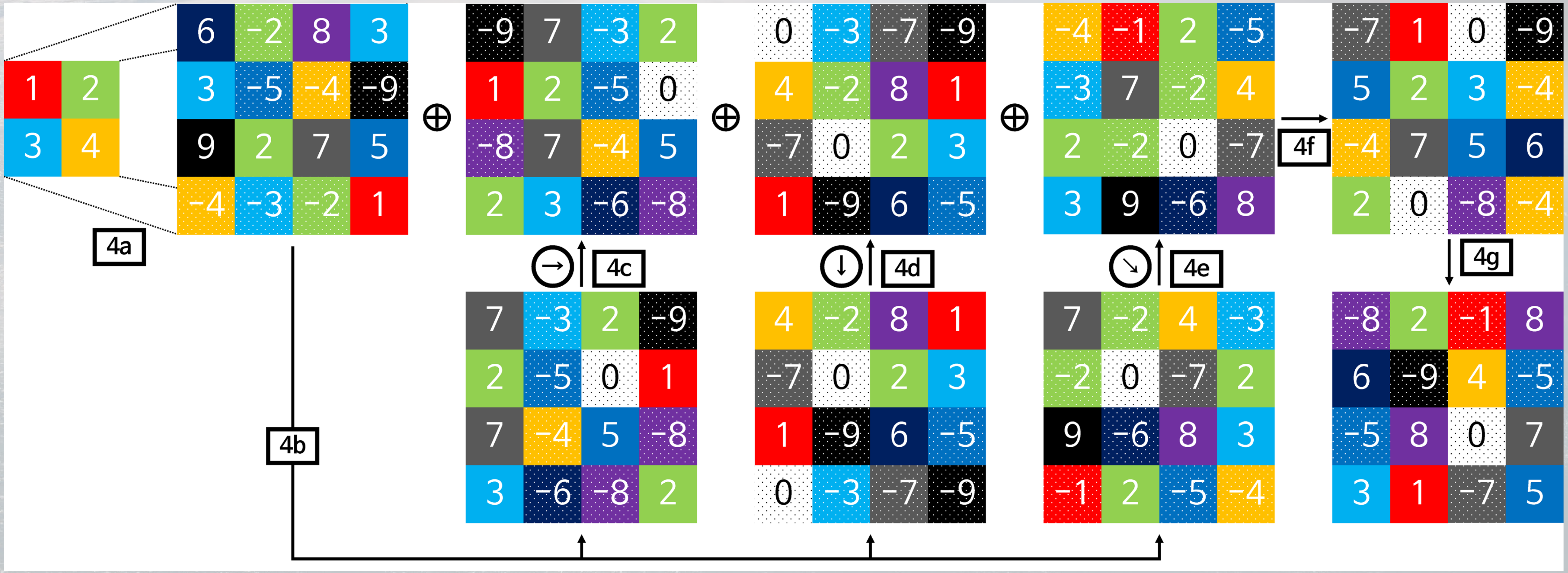


Pixel Transposed Convolution
(Gao et al., 2020)



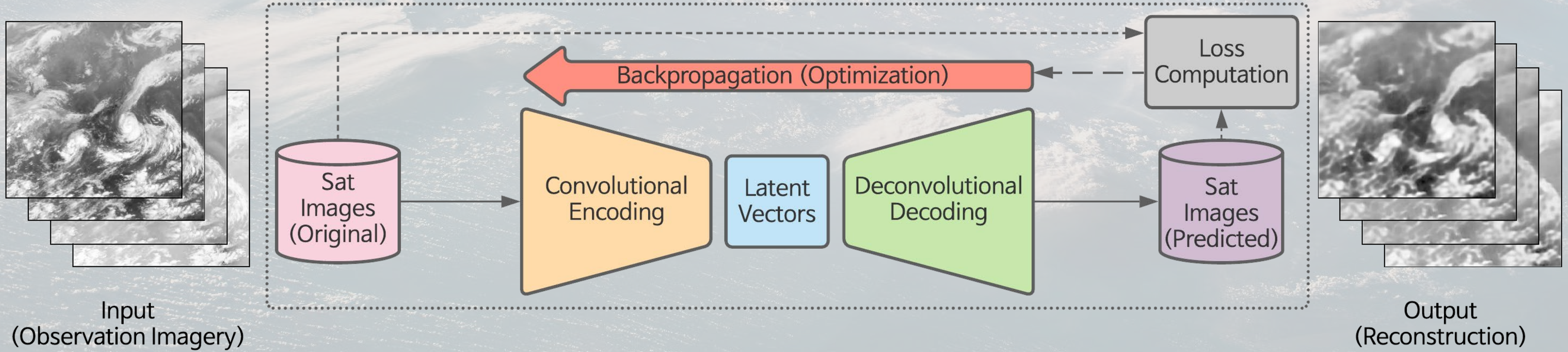
02 | Methods | Our solution

Shift-N-Overlap Cascaded Deconvolution



02 | Methods | Overall

- An **unsupervised autoencoder** model can **extract principal factors** and **learn** both **key visual features** and **desired data distribution** by **itself**
- A **universal and versatile encoder-decoder** model is demanded for **multi-purpose prediction tasks** from **wide observation area**



03 | Target Data | MTSAT1R/2,COMS

- Preparation of homogeneous observation dataset

COMS-1:128.2° E(COMS-1/OSCAR)

MTSAT-1R: 140° E (Himawari-6/OSCAR)

MTSAT-2: 145° E (Himawari-7/OSCAR)

Homogeneous Satellite Instruments (MI, JAMI, IMAGER)

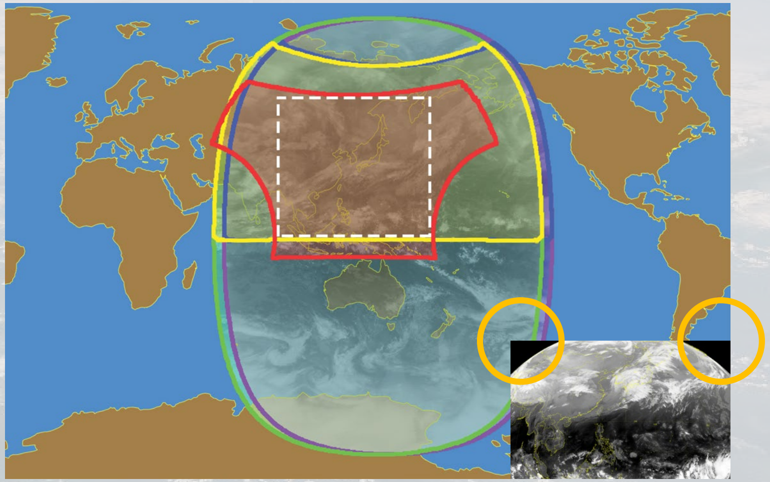
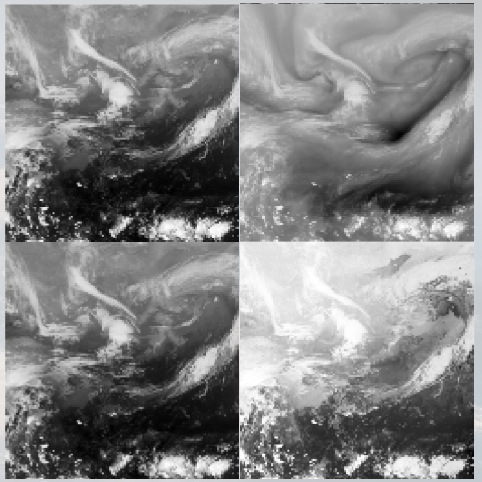


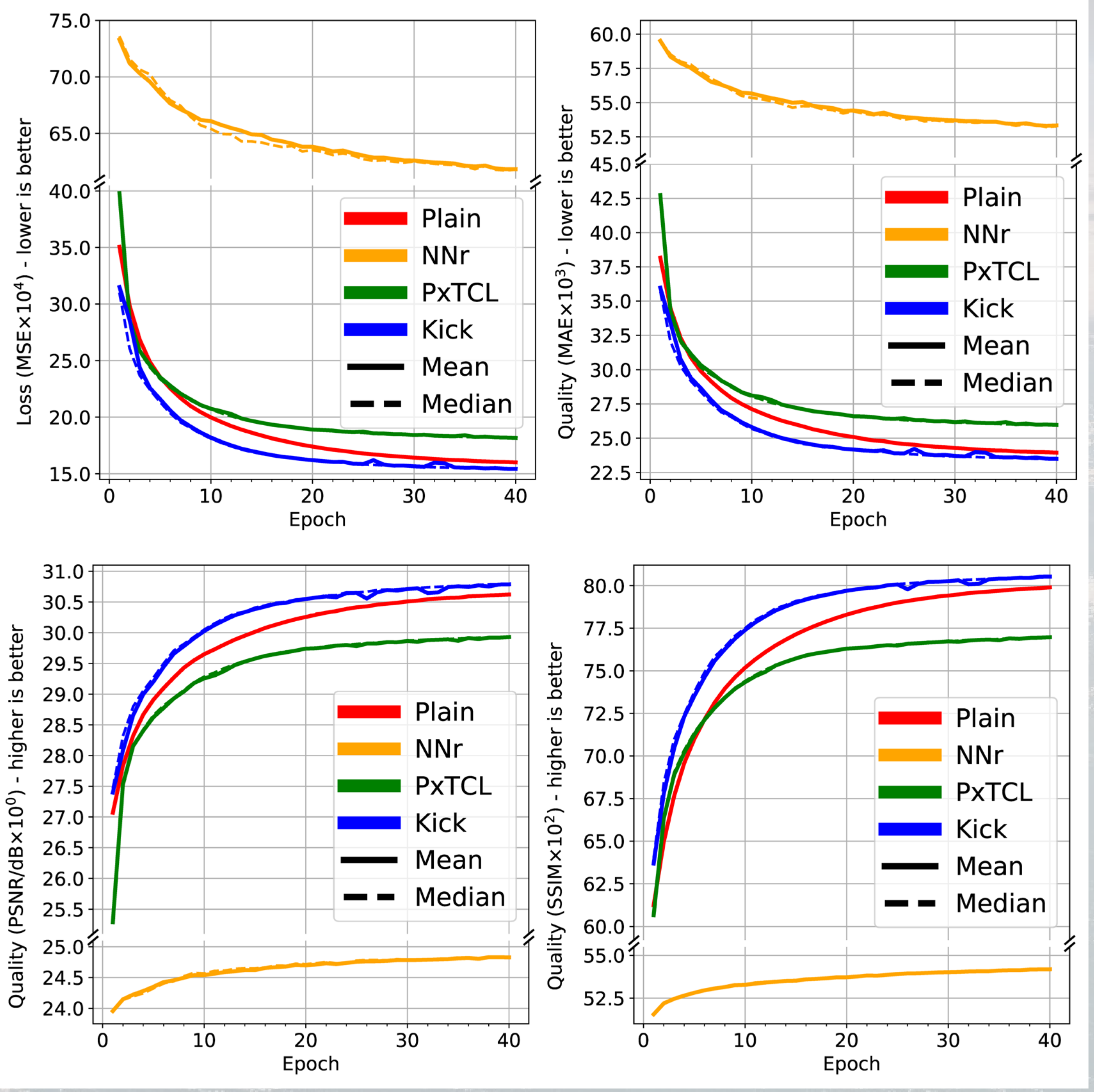
Image rearrangement and reprojection



Final image arrangement

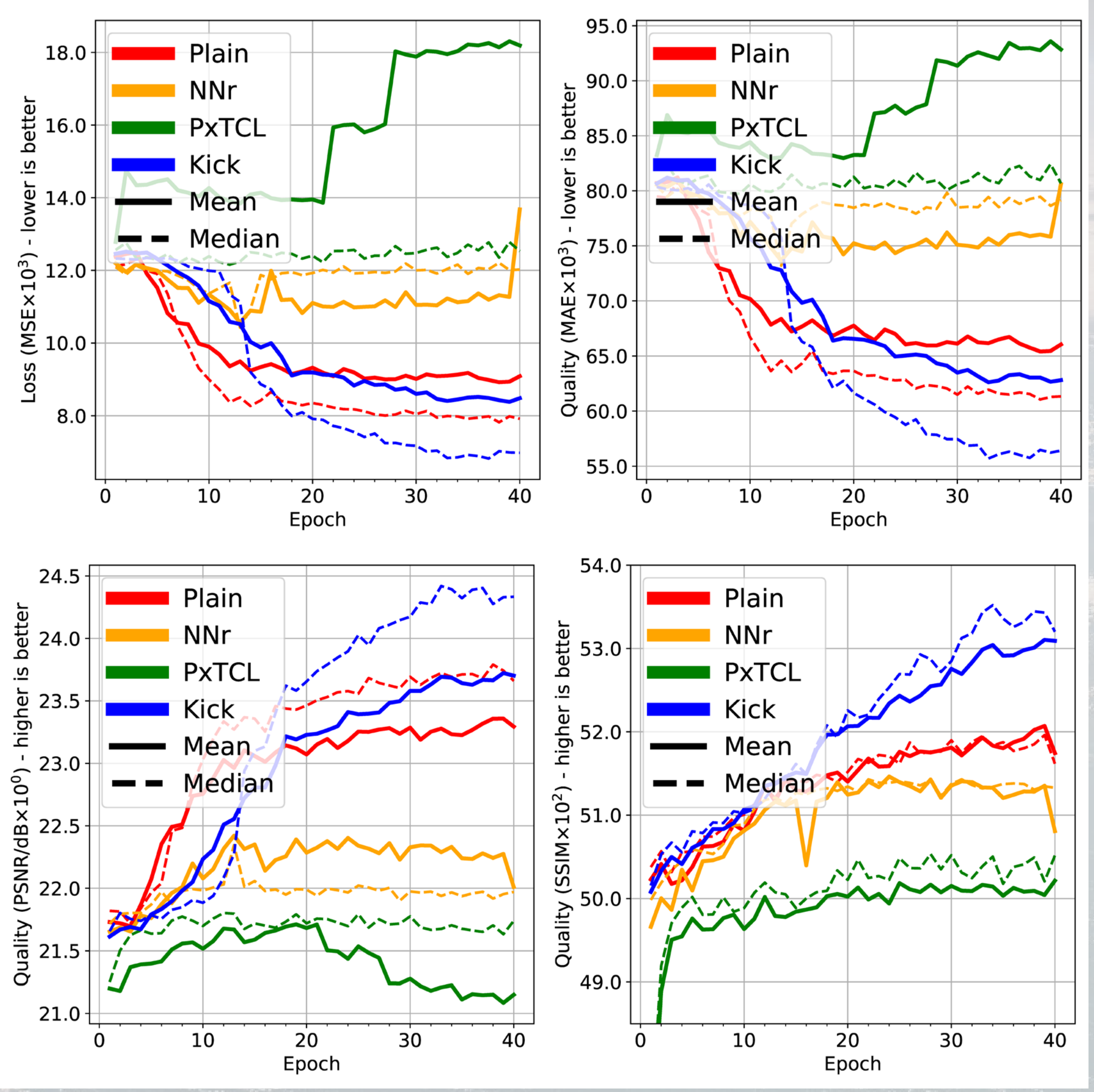
- Each satellite program **operates on limited operational plan** (due to a durability)
- **Observation area reprojection is mandatory,** due to the **reconstruction tendency of void space** in autoencoders
- **Additional image rearrange is required due to the discrete mission orbit**
- **Entire Duration: 2006~2017Y (307,808 images) / Valid, Test: 2014~2017Y**

04 | Results | Convolutional AE



- Model Training Configurations
  - Model:
    Convolutional Autoencoder (ConvAE)
  - Loss/Metric: Mean-squared (MSE)
  - Learning rate: 1e-3
  - Epoch: 40
  - Opt.: Adam (Adaptive Moment Estimation)
- Results
  - **Better performance metrics** than any **other deconvolutional models**
  - Fastest model convergence within epochs

04 | Results | Adversarial AE



- Model Training Configurations
  - Model: **Adversarial Autoencoder (AdvAE)**
  - Loss/Metric: Mean-squared (MSE)
  - Learning rate: 1e-3
  - Epoch: 40
  - Opt.: Adam (Adaptive Moment Estimation)
- Results
  - **Better performance metrics** than any **other deconvolutional models**
  - **Significant loss fluctuation** is observed due to the **model complexity**

04 | Results | Error Metrics

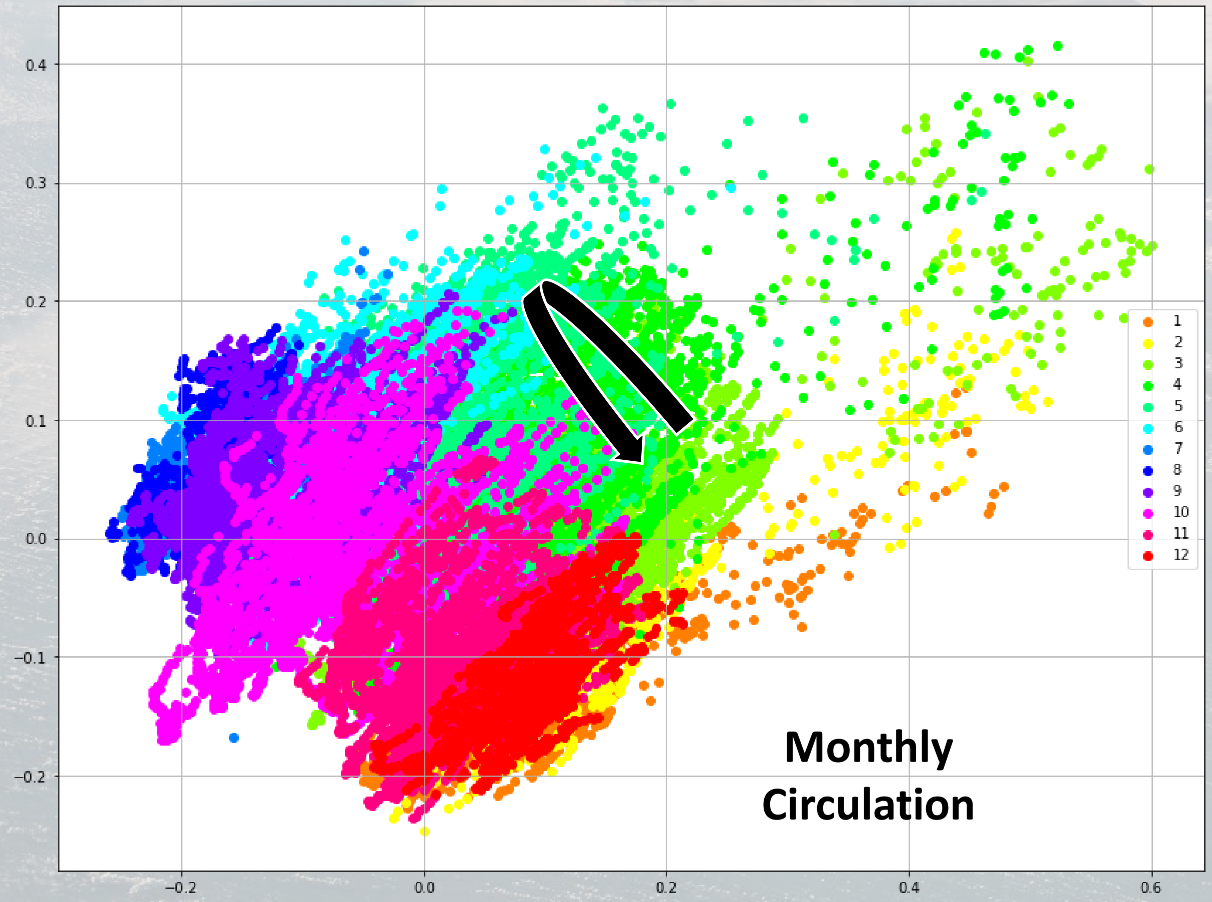
- **Result of Average (50 discrete cold-start)** model training and evaluation
- **NNr** shows weak perf on **ConvAE**, and **PixelTCL** show weak perf on **AdvAE**

	Convolutional Autoencoders (ConvAE)				Adversarial Autoencoders (AdvAE)			
Metric	Plain (Baseline)	NNr	PixelTCL	Kick **(Ours)**	Plain (Baseline)	NNr	PixelTCL	Kick **(Ours)**
MSE	0.00086 (0.00001)	0.00375 (0.00006) [−336.0%]	0.00102 (0.00001) [−18.6%]	**0.00085 (0.00001)** [+1.2%]	0.00510 (0.00101)	0.00821 (0.01564) [−61.0%]	0.01179 (0.02076) [−131.2%]	**0.00493 (0.00109)** [+3.3%]
MAE	0.01726 (0.00012)	0.04024 (0.00049) [−133.1%]	0.01912 (0.00012) [−10.8%]	**0.01708 (0.00015)** [+1.0%]	0.04803 (0.00598)	0.05834 (0.03272) [−21.5%]	0.06929 (0.04973) [−44.3%]	**0.04649 (0.00661)** [+3.2%]
PSNR	33.71803 (0.05058)	27.72582 (0.09519) [−17.8%]	32.77205 (0.03890) [−2.8%]	**33.79470 (0.04102)** [+0.2%]	26.36942 (0.95127)	25.21174 (2.34726) [−4.4%]	24.23747 (3.05105) [−8.1%]	**26.60216 (1.05993)** [+0.9%]
SSIM	0.84087 (0.00051)	0.63357 (0.00135) [−24.7%]	0.81317 (0.00074) [−3.3%]	**0.84415 (0.00055)** [+0.4%]	0.61335 (0.02124)	0.60970 (0.04198) [−0.6%]	0.60392 (0.01087) [−1.5%]	**0.62000 (0.00585)** [+1.1%]

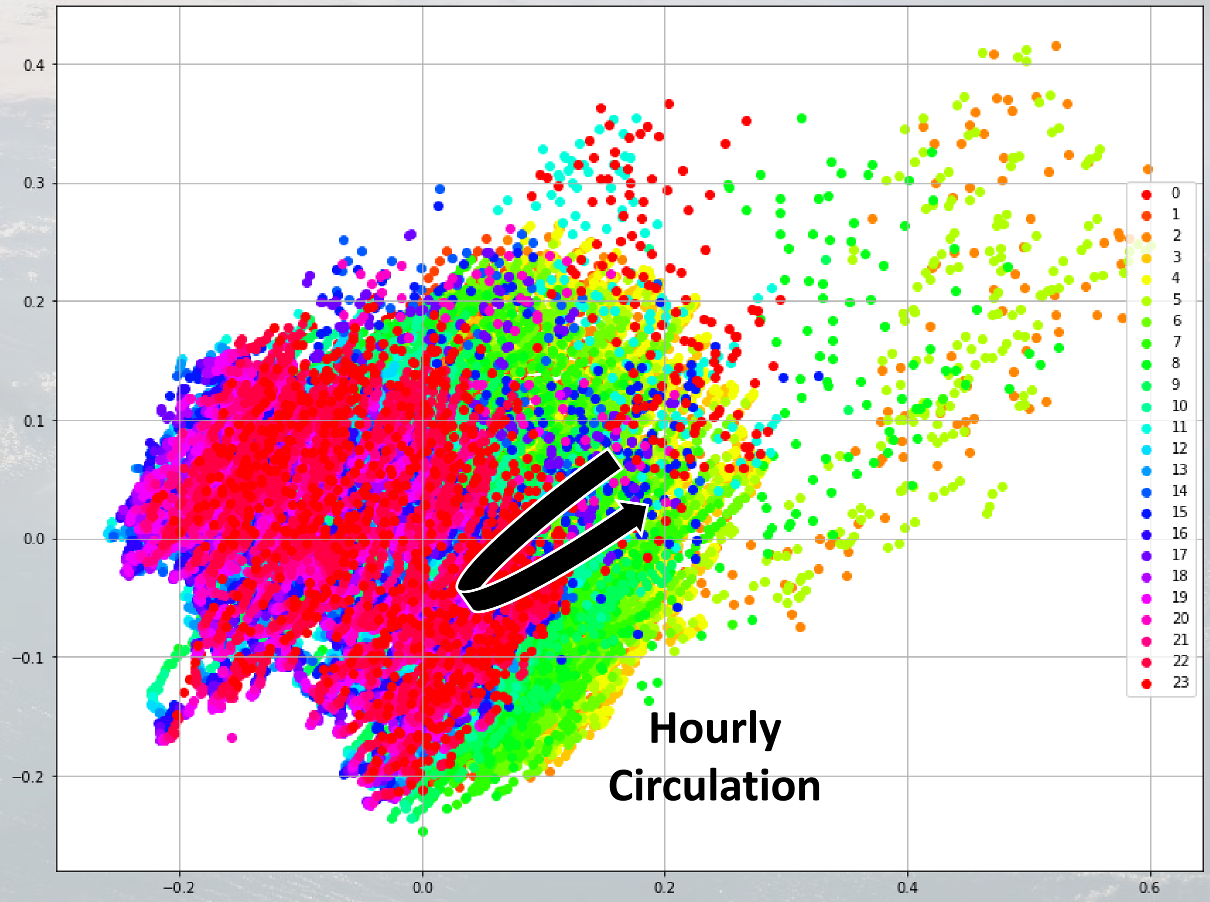
04 | Results | Yearly Circulation

- **Latent Information** (below is PCA-ed) can capture **hourly and monthly circulation** very well

Test (Y2015/Y2017)
Data: Hourly (UTC 00H~23H) | Label: Monthly (M01~M12)



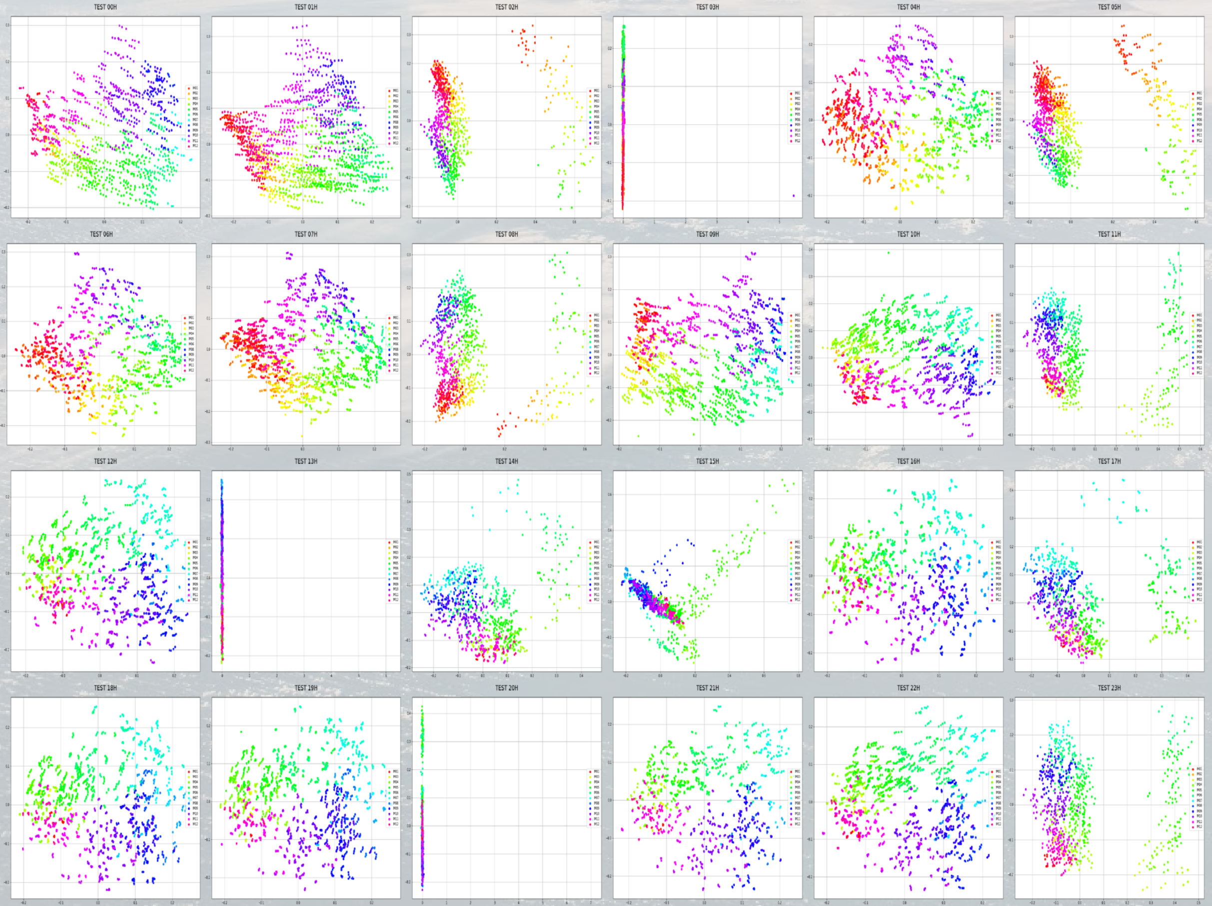
Test (Y2015/Y2017)
Data: Monthly (M01~M12) | Label: Hourly (UTC 00H~23H)



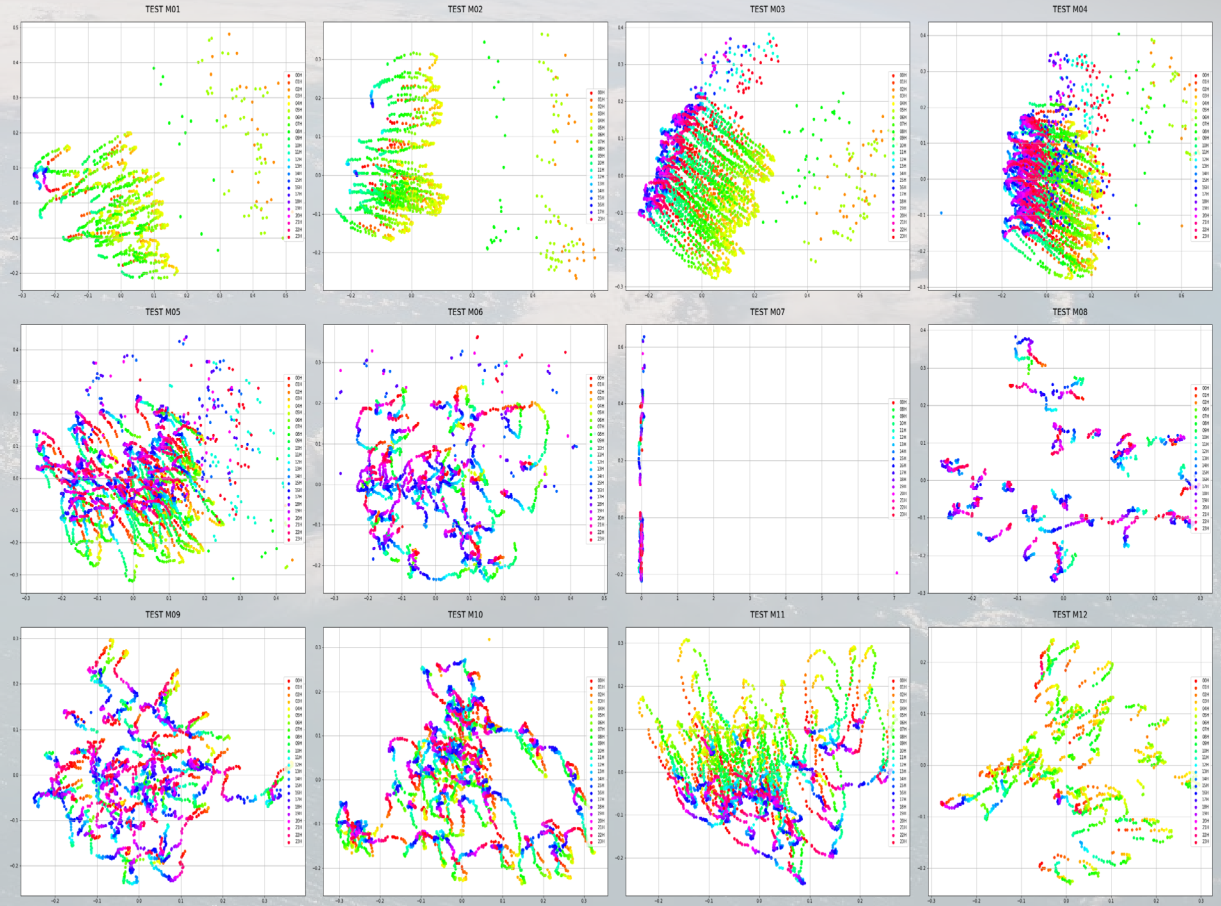
04 | Results | Hourly and Monthly Circulation

- **Latent Information** (below is PCA-ed) can capture **hourly and monthly circulation** very well

Test (Y2015/Y2017)
Data: Hourly (UTC 00H~23H) | Label: Monthly (M01~M12)



Test (Y2015/Y2017)
Data: Monthly (M01~M12) | Label: Hourly (UTC 00H~23H)



05 | Conclusion | Contributions

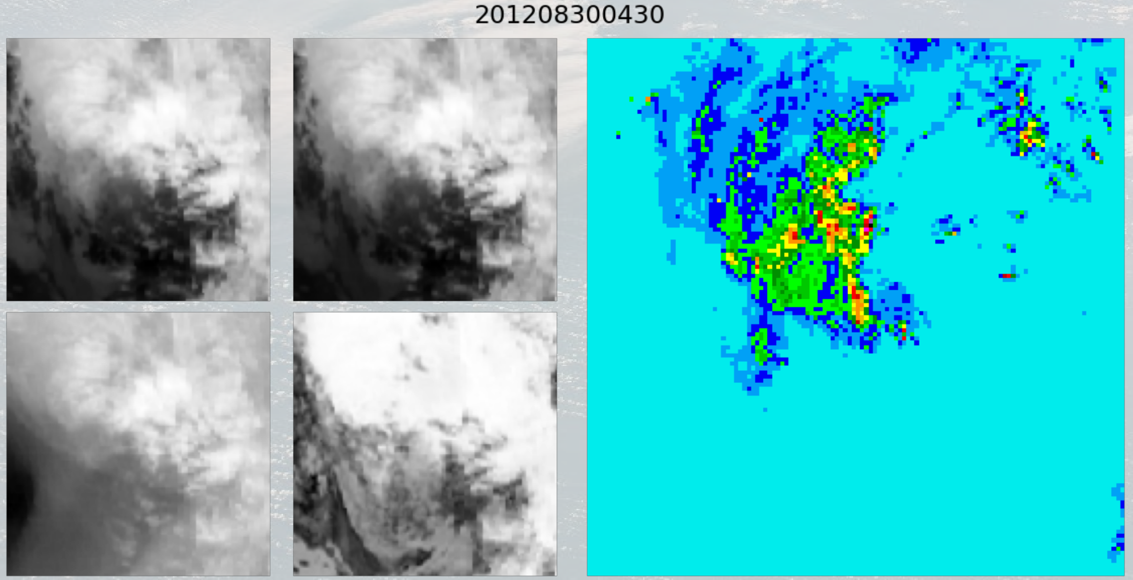
- This work revealed that **a change in the decoder part** of autoencoder could **vary the entire learning of weather representations** without **changing the encoder part**.

- This work is providing a previously-trained convolutional autoencoder model for **understanding 4-channel geostationary satellite images** using autoencoders.

- This work enhanced an **image reconstruction performance** by very unique deconvolutional layer named 'Kick' **to minimize checkerboard artifacts** issue.

05 | Conclusion | Future works

- **Application on Real-world Prediction** using **Pre-trained Convolutional AE** (currently working on radar echo image generation task)
- **Large-scale Image Research Dataset Processing** with **Distributed-Parallel Computing Acceleration** on Open Research Collaboration Platforms



Geostationary Satellite Images   MAPLE Rain Echo Estimation

RAPIDS → data·on → kubernetes

References

- Gao, H., H. Yuan, Z. Wang, and S. Ji (2020, May). Pixel Transposed Convolutional Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence 42 (5), 1218-1227. doi: 10.1109/TPAMI.2019.2893965
- Hong, S. and S.-K. Song (2020, June). Kick: Shift-N-Overlap Cascades of Transposed Convol utional Layer for Better Autoencoding Reconstruction on Remote Sensing Imagery. IEEE Access 8, 107244-107259. doi: 10.1109/ACCESS.2020.3000557
- Odena, A., V. Dumoulin, and C. Olah (2016, October). Deconvolution and Checkerboard Artifacts. Distill, http://distill.pub/2016/deconv-checkerboard.
- Zeiler, M. D., D. Krishnan, G. W. Taylor, and R. Fergus (2010, June). Deconvolutional networks. In 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Proceedings, pp. 2528-2535. doi: 10.1109/CVPR.2010.5539957

Acknowledgments

❖**Data Provision**

MTSAT-1R/2 observation data were provided by the **Center for Environmental Remote Sensing (CEReS)**, Chiba University, Japan. COMS-1 observation data were provided by the **National Meteorological Satellite Center (NMSC)** of Korea Meteorological Administration (KMA).

❖**Computing Resource & Technical Support**

Computing resources supporting this work were mainly provided by the **High-Performance Computing (HPC) Support Program** for 2019 (xo) and 2020 (xo4wj) through the **National IT Industry Promotion Agency (NIPA)**, Republic of Korea.

I would also like to acknowledge the support of **NVIDIA Corporation** with the **donation of multiple GPUs used for the initial survey of this research**, and the **Korea National Super-computing Center (KISTI KSC)** for technical support with **supercomputing resources**.

Thank you for listening! Any questions?

Seungkyun Hong

xo@kisti.re.kr | ontheklaud@gmail.com

Korea University of Science and Technology (UST)
Korea Institute of Science and Technology Information (KISTI)