Enhancing Land Surface Temperature Imaging Through Deep Learning based Super Resolution

2023. 11.08

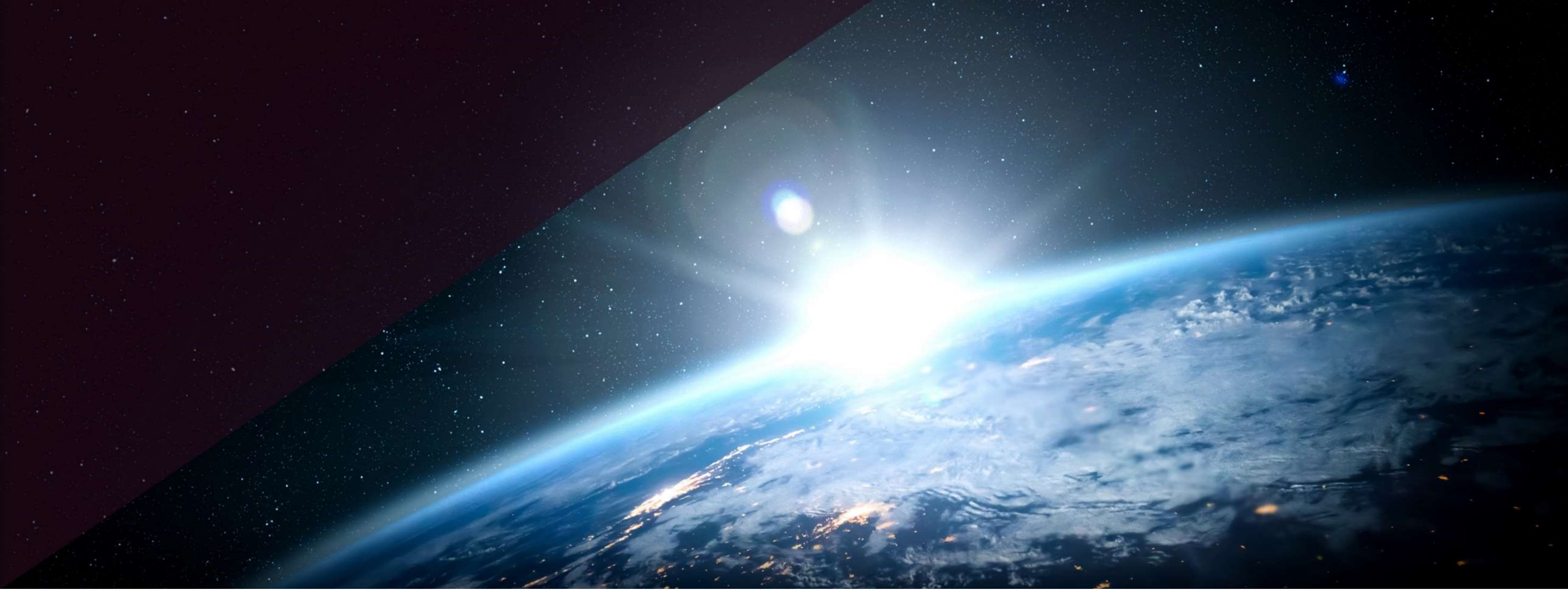
Sunju Lee

Research Scientist, GeoAI Research Center, SI Analytics

SIA SI ANALYTICS

SIA Industry applicable case

SuperX



SIA SuperX enhances the visibility and clarity of optical satellite images, contributing to accurate interpretation of the imagery.

01

PlanetScope

Location: San Francisco
Date: 2018

Original (GSD: 3.0m-4.1m) | SIA SuperX (GSD ≈ 1.6m-2m)

02

WorldView2

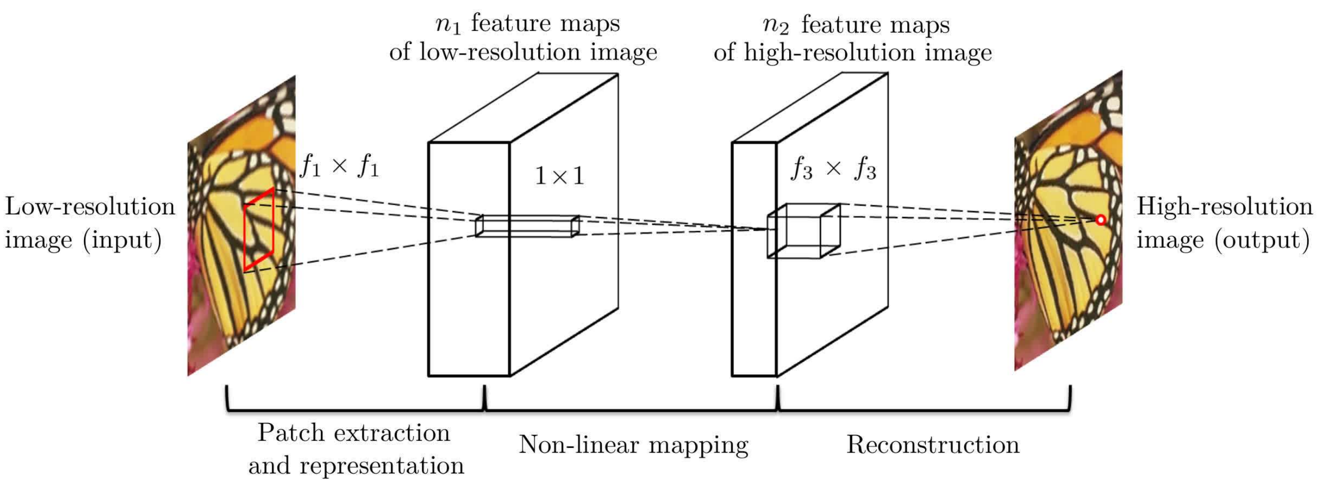
Location: South Sea Fleet LCAC Base
Date: 2022/05/03

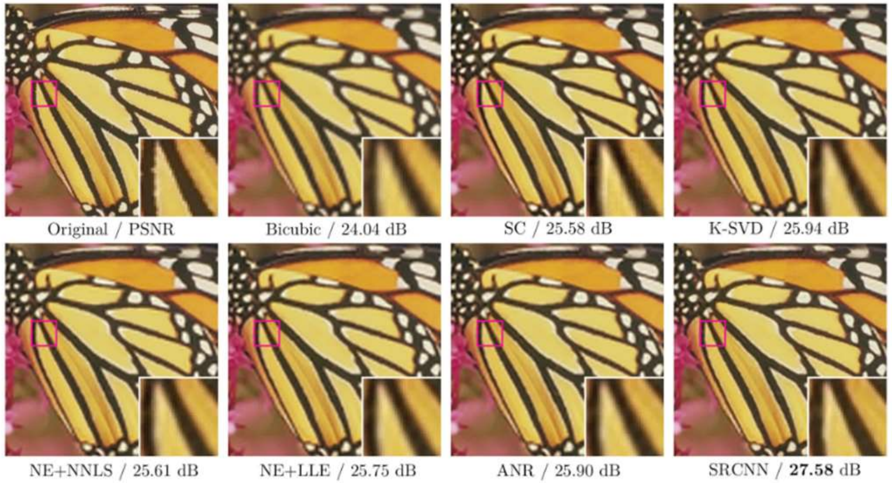
Original (GSD: 0.46m) | SIA SuperX (GSD ≈ 0.23m)

Super-Resolution of Land Surface Temperature Data Using Multi-Source Satellite Imagery

Super Resolution

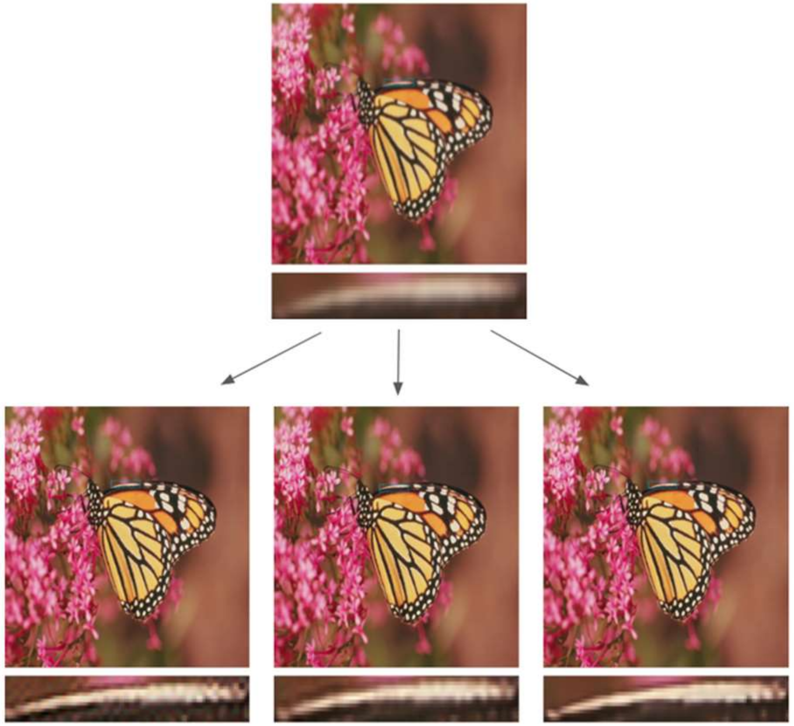
- Learning a deep convolutional network for image super-resolution (2014)
  - Basic Single Image Super Resolution Model using Low Resolution and High Resolution Image

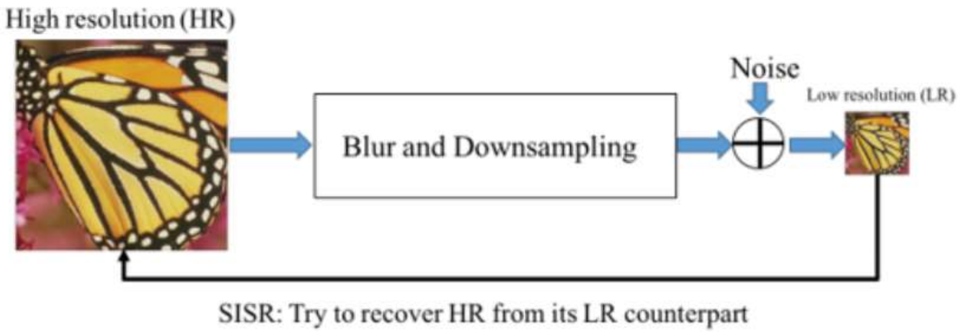




Super Resolution

- How to make real Low-resolution image? (ill-Posed problem)





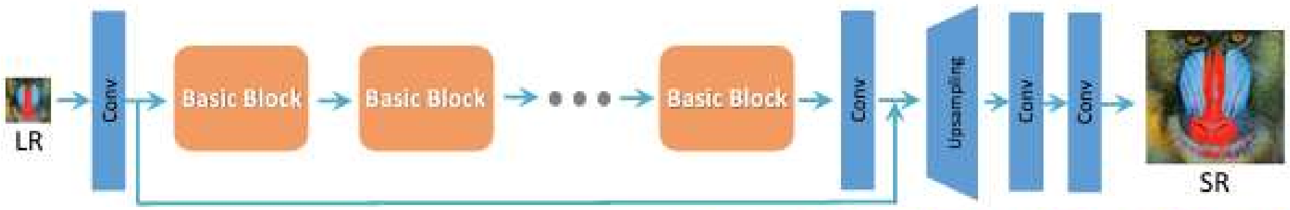
Model ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks

- Model with improved overall perceptual quality for Super Resolution

- Although the SRGAN model improved visually more than the existing PSNR-oriented methods, there was still a difference from the real High Resolution image, and ESRGAN suggested a way to improve it.

- The structure is almost same to SRGAN, changing to a Residual Dense Block (RRDB) and density-connecting to the residual connection while removing the batch normalization

- Before the feature map after the activation function was used to calculate the Perceptual loss, but in this paper, the feature map prior to the application of the activation function was used => because the feature map after the application of the activation function has low performance due to thin information from the activated feature, and the use of the feature after activation has the disadvantage of being inconsistent even in brightness restoration in reconstructed images.



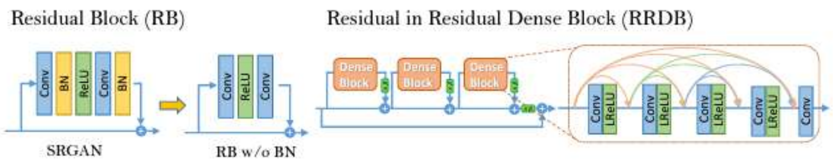
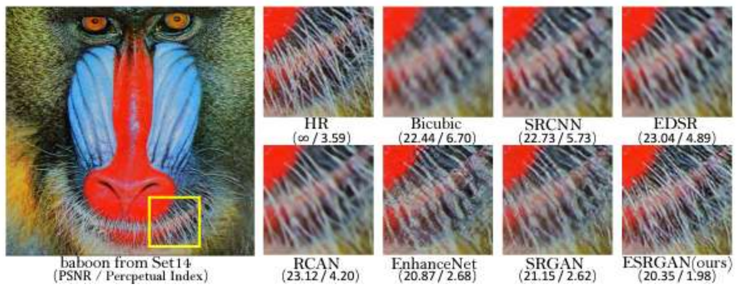


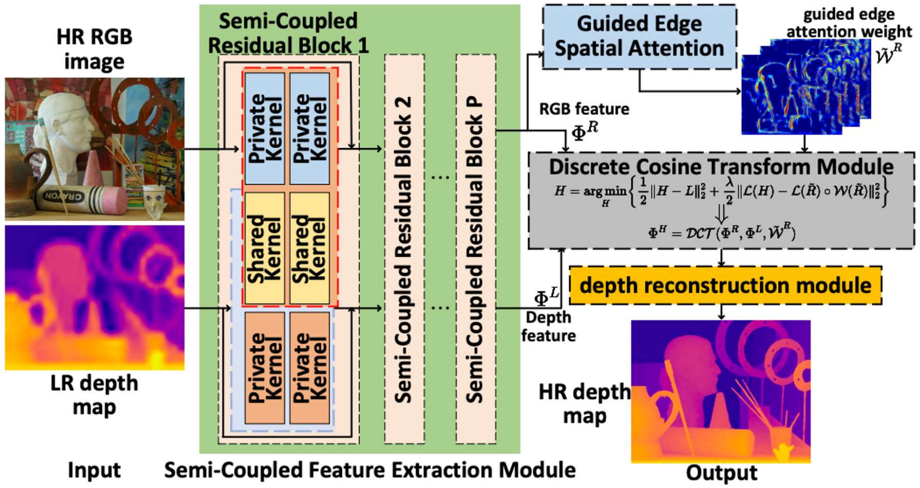
Fig. 4: **Left**: We remove the BN layers in residual block in SRGAN. **Right**: RRDB block is used in our deeper model and β is the residual scaling parameter.



Reference based Super-resolution

- Researching super resolution techniques using reference images using GK2B data

GDSR-DCTNet (Discrete Cosine Transform Network for Guided Depth Map Super-Resolution
(CVPR 2022 Oral)

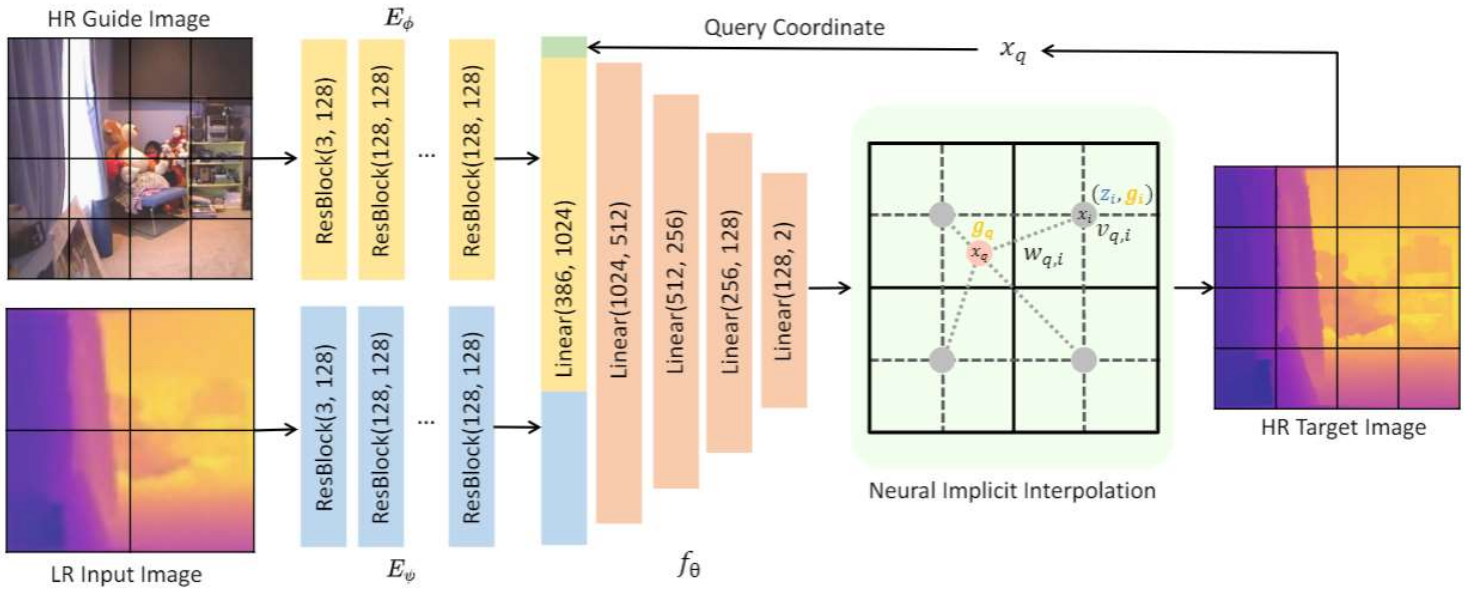


- Designed to extract shared and cross modality of each feature using <u>Semi-coupled feature extraction (SCFE).</u>
- <u>Guided edge spatial attachment (GESA)</u> is a module that prevents irrelevant textures from being transmitted to the super-resolution depth map when the guided RGB image contains rich textures. The purpose is to effectively emphasize the outline of the object and smooth the texture information inside the object.
- <u>Discrete cosine transform module</u> enables the Upsampled LR to be compared by separating low frequency and high frequency through Laplacian filter detection. The feature of Depth map is using Laplacian filters (edge detection). DCT computation has been commonly used when conducting compression methods that usually maintain low-frequency components and reduce high-frequency components. In this model, DCT operation was used to separate low and high frequencies, and DCT block allows the value transferred from the Guided Edge Spatial Attachment block to be calculated separately into Low Frequency and High Frequency before DCT operation.
- <u>The Depth reconstruction module</u> aims to predict a high-resolution Depth map from the feature map.

Reference based Super-resolution

- Researching super resolution techniques using reference images using GK2B data

JIIF: Joint Implicit Image Function for Guided Depth Super-Resolution (ACM MM 2021)
Represent target image with local latent codes for both input and guide images.
Using the graph attention mechanism, the learning of interpolation weights and values is integrated into one representation.



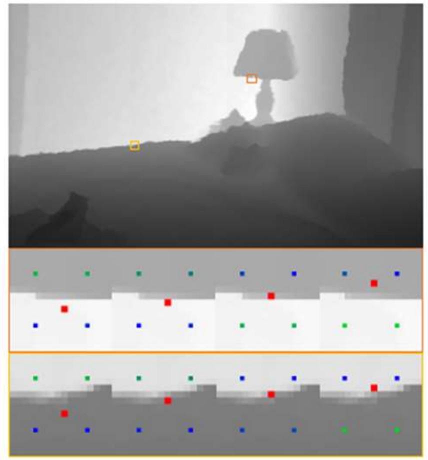
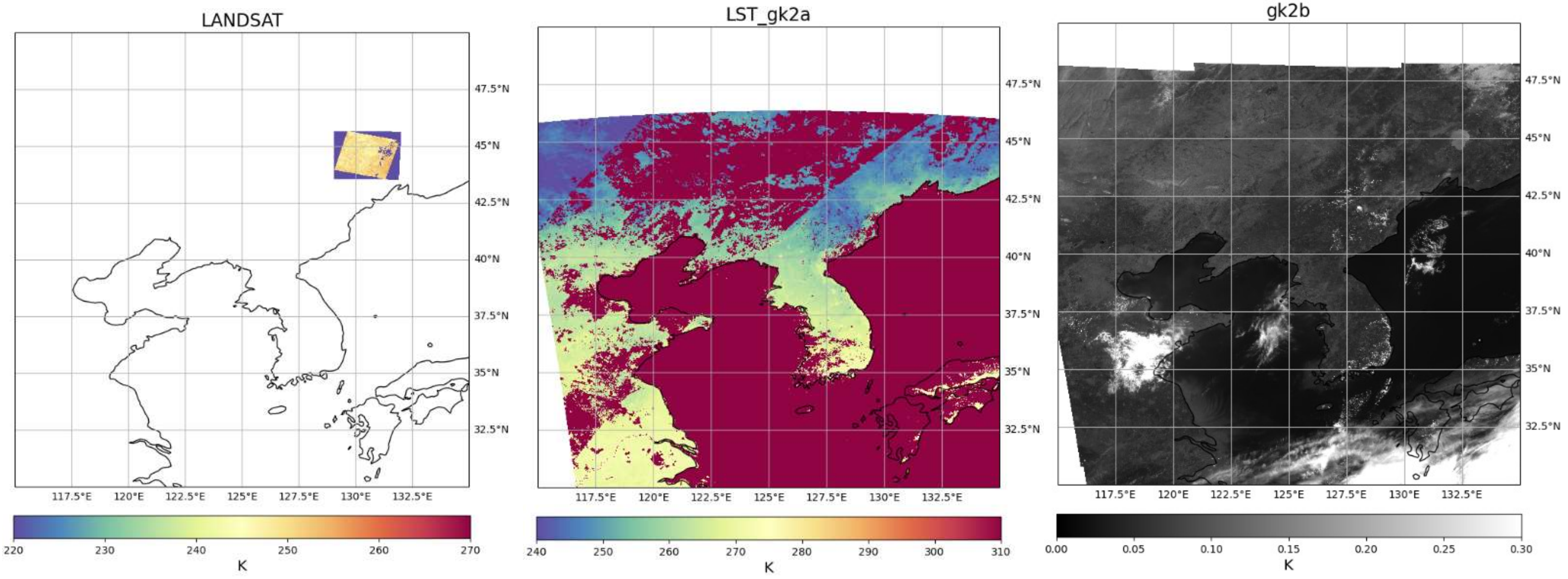


Figure 6: Visualization of the learned interpolation weights. We show two examples of the query pixel crossing an edge in the HR target image. The query pixel is in red, and the four corner pixels' color indicates the learned interpolation weights. Higher weights are in bluer color, while lower weights are in greener color.

Reference based Super-resolution

- Dataset

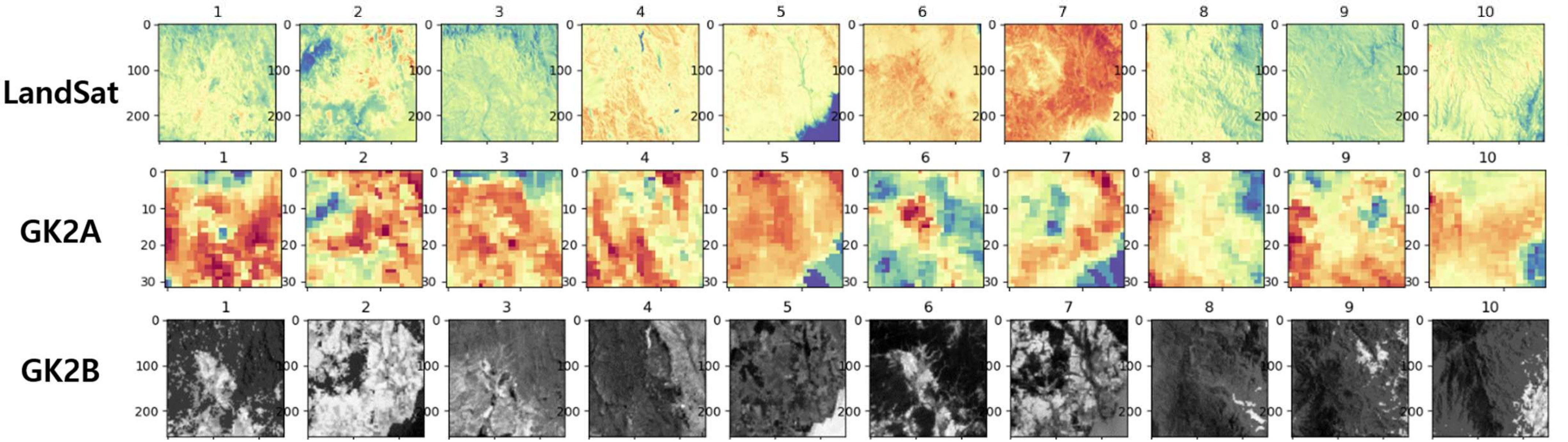
- Image registration using GK2B.
HR (Target image) – LANDSAT, LR (input image) – GK2A, HR (Reference Image) – GK2B
- LANDSAT : GSD 30m -> 250m, GK2A -> 2km, GK2B -> 250m



Reference based Super-resolution

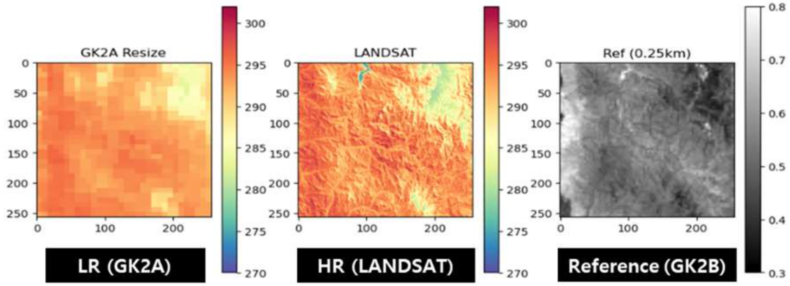
- Dataset

- Image registration using GK2B.
HR (Target image) – LANDSAT, LR (input image) – GK2A, HR (Reference Image) – GK2B
- LANDSAT : GSD 30m -> 250m, GK2A -> 2km, GK2B -> 250m



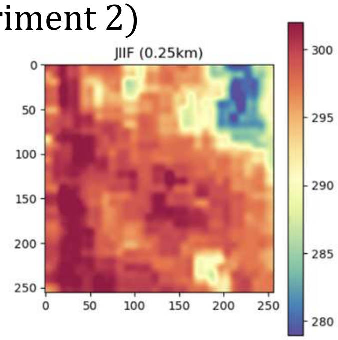
Reference based Super-resolution

- Results (x8)



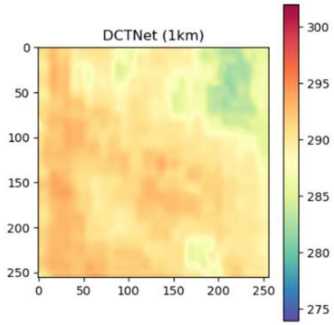
Comparison of experimental results with JIIF and DCTNet in a methodology using Reference Image to compare the resulting images according to the structure of the model
(The dataset and loss function of the experiment are all the same, and the input is the same as experiment 2)

JIIF



Advantage: Model using Implicit Neural Representation
LST values for pixels have various range values
Cons: What appears to be Block Noise occurs

DCTNet

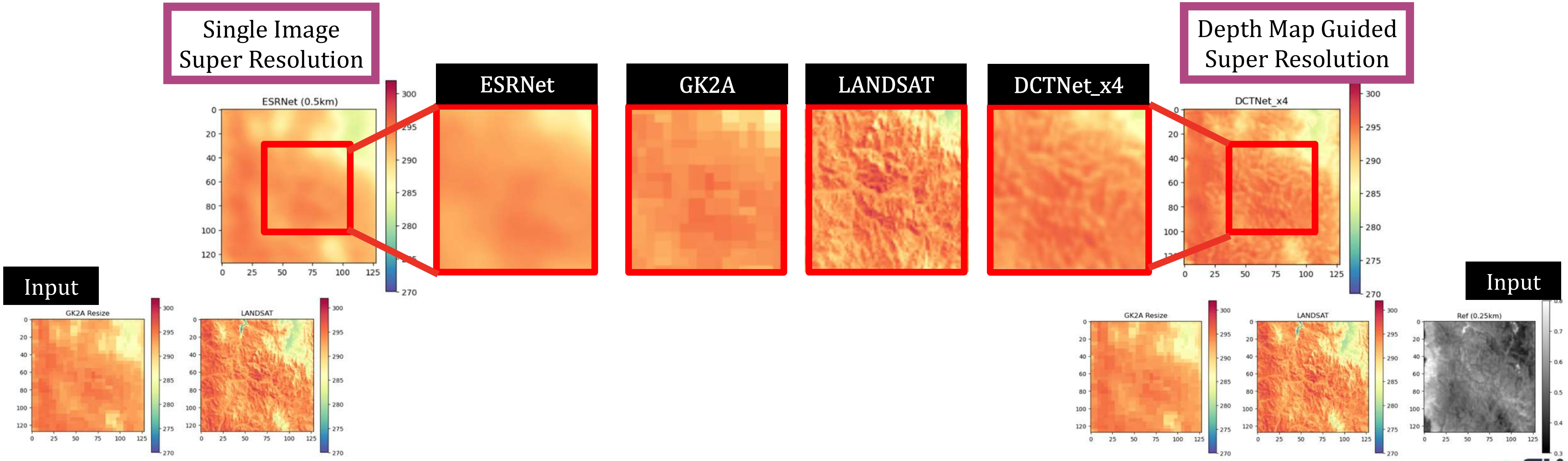


Advantage: Model using Guided Edge Attachment operation Properly addressing Low and High Frequency
Cons: SRised images have a fairly narrow range of values

Reference based Super-resolution

- Results (x4)

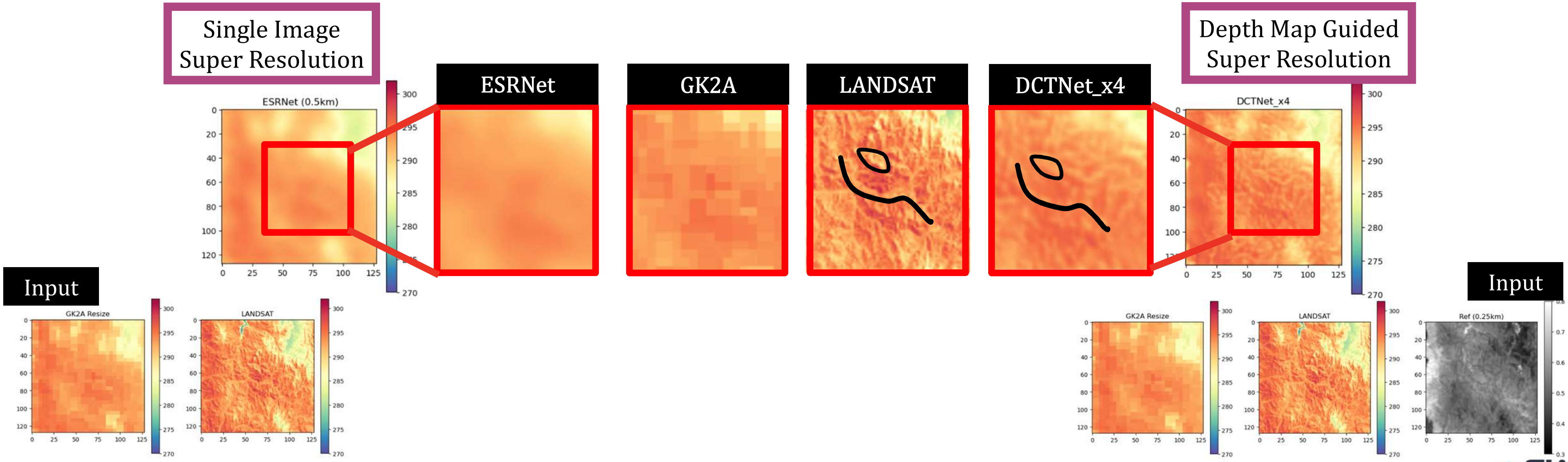
Comparison of Reference Image-based models (DCTNet) and Single Image-based models (ESRNet) to verify the effectiveness of the delivery of Reference Images.
(DataNormalization and Loss Function of the experiment are all the same)



Reference based Super-resolution

- Results (x4)

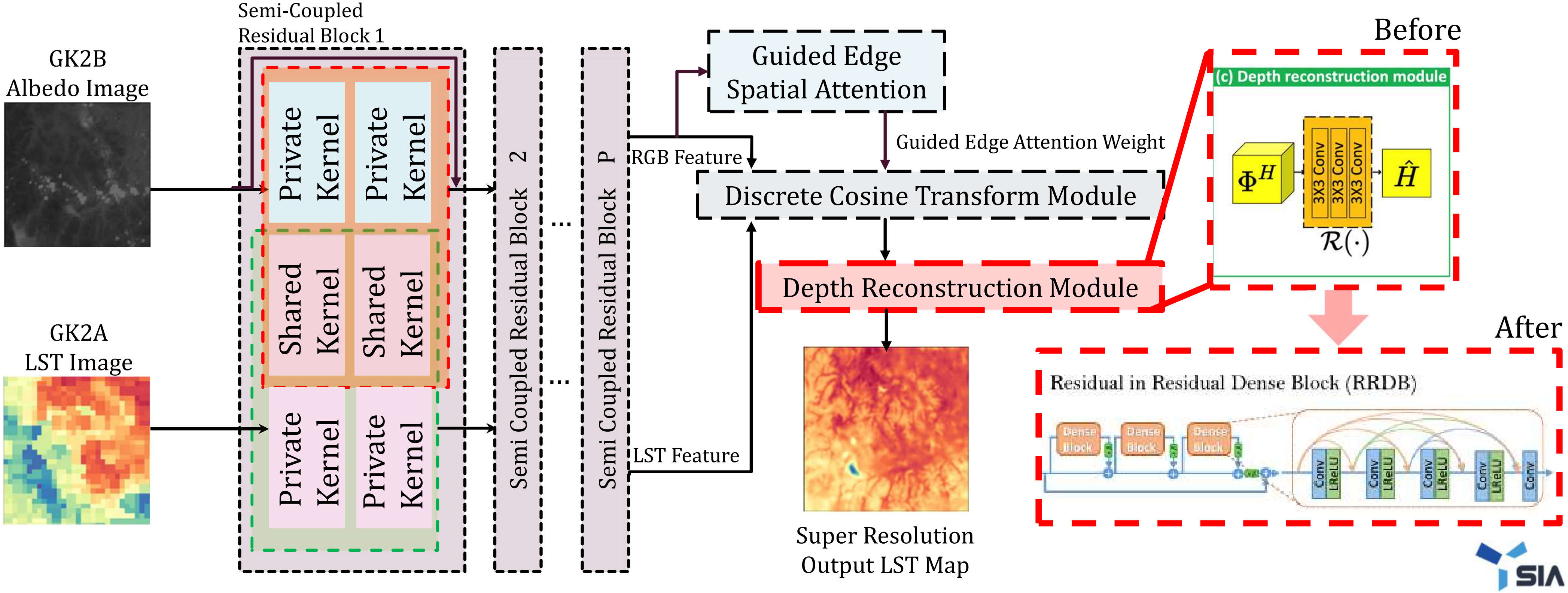
Comparison of Reference Image-based models (DCTNet) and Single Image-based models (ESRNet) to verify the effectiveness of the delivery of Reference Images.
(DataNormalization and Loss Function of the experiment are all the same)



Reference based Super-resolution

- Conv to RRDB (Convolution to Residual in Residual Dense Block)

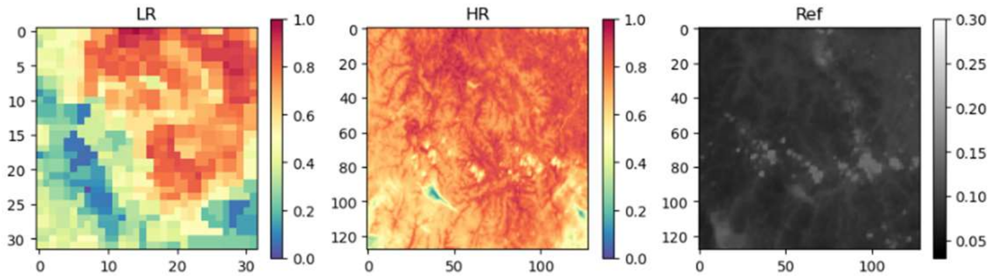
Blur results can be generated a little more detail through RRDB Block Architecture learning methods while using a network-based model with Discriminator as a generator structure.

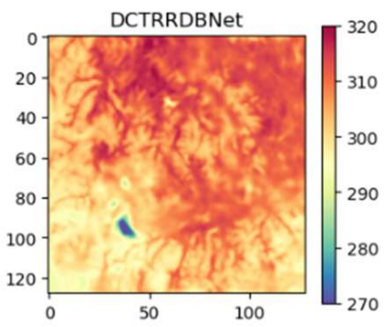


Reference based Super-resolution

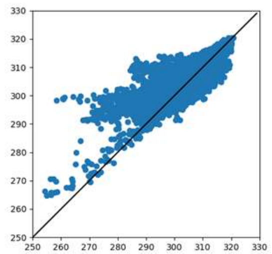
- Conv to RRDB (Convolution to Residual in Residual Dense Block) Result

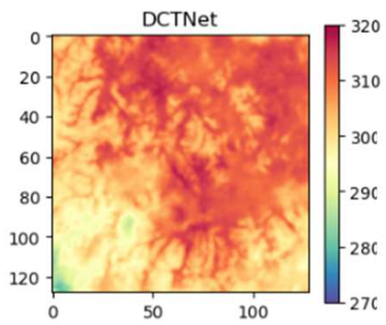
Blur results can be generated a little more detail through RRDB Block Architecture learning methods while using a network-based model with Discriminator as a generator structure.



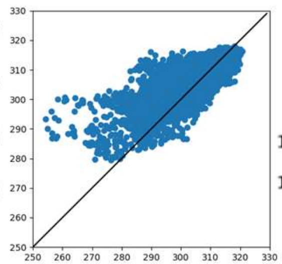


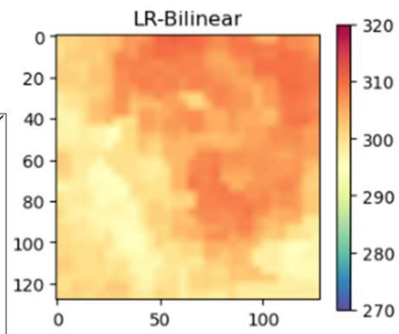
PSNR : 36.36
SSIM : 0.92
MSE : 15.02



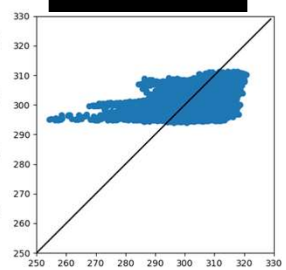


PSNR : 35.27
SSIM : 0.90
MSE : 19.32





PSNR : 31.89
SSIM : 0.84
MSE : 42.07



Reference based Super-resolution

- Total Result

It is a table that can evaluate the performance of our models through experiments so far.

MSE, PSNR and SSIM were used as corresponding indicators,
and can be compared and analyzed for each model.

The red bold indicates the highest score, and the blue underlined indicates the second highest score.
As a test set, we built on 10 randomly sampled images.

	PSNR	SSIM	MSE
Bilinear x4	35.253	0.931	20.403
ESRNet x4	35.470	0.916	20.163
DCTNet x4	<u>35.792</u>	**0.938**	<u>19.409</u>
DCTRRDBNet x4(Ours)	**36.885**	<u>0.937</u>	**14.939**

Challenging point

- Data needs to be used as a pair for GK2B, GK2A, and LANDSAT, making it difficult to build a dataset
- Super Resolution Task is challenging because of limited features that can be extracted from small patches when doing 8x, unlike 4x
- Normalization and Denormalization are difficult due to differences in data domains, and images from models do not fully follow the value range because they do not know the domain of GT during Denormalization

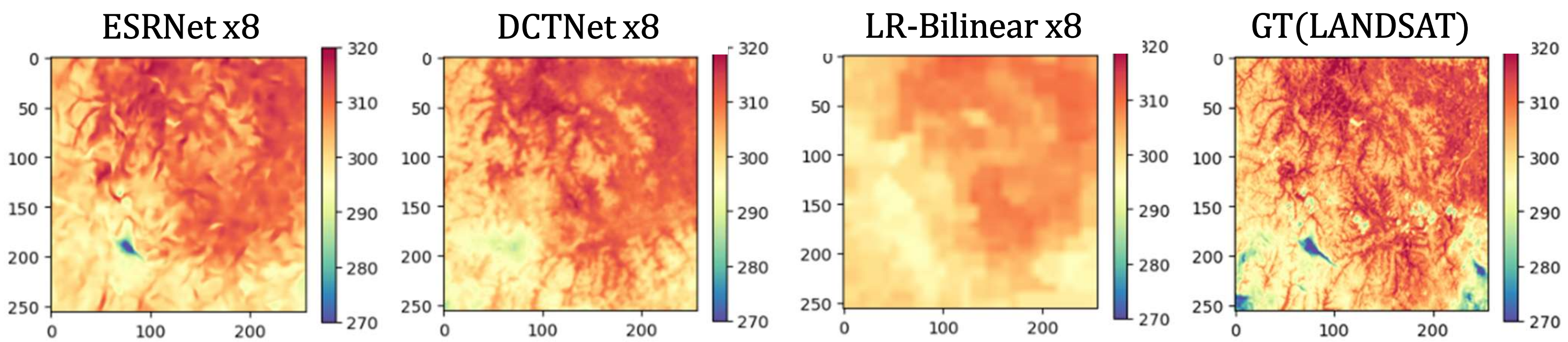
Future work

- X8 Super Resolution Model

The advantage of GAN, a generative model, is that it separates the high frequency area from the low frequency area while clearly changing the blur.

However, in the case of 8 scale, the patch size of the Low Resolution image that can extract features was 32x32, which was too small to proceed with proper learning.

Therefore, I think that better results can be achieved if we can create more HR, LR, and Reference Images, and build a learning dataset with larger images.



Thank you

Thank you for attention

2023. 11.08

www.si-analytics.ai

70 Yuseong-daero 1689 beon-gil, Yuseong-gu, Daejeon, 34047, Republic of Korea