Research Proposal

Jiang He
Wuhan University

Earth observation data plays an essential role in understanding and exploring our planet. However, various degradation phenomena can occur, affecting the quality and accuracy of the captured data due to the atmosphere, topography and sensors. Improving the quality of big earth data is crucial for accurate analysis, insightful decision-making, and the advancement of various scientific disciplines.

My core research focuses on satellite image quality improvement and remote sensing information extraction. In particular, my work centers on three areas: ***1) Spatial-spectral fusion***, integrating multi-source data into a spatial-spectral cube, ***2) Spectral super-resolution***, learning spectral transformation from spectral library and increasing the spectral resolution of available spectra, and ***3) Earth observation semantic extraction***, extracting semantic information from satellite data.

The main algorithms in my research involve variational optimization, general deep learning, embedding physical mechanisms into deep learning, multi-task learning, and self-learning.

In the following, first I give a brief overview of my research. The second part of the document contains a detailed description of my research as well as the related future research directions.

Summary of Research

Spectral super-resolution: Spectral super-resolution is a crucial technique to obtain hyperspectral images from multispectral images. My early work considered using deep learning to address spectral super-resolution. However, as deep learning-based algorithms are always blamed for the lack of physical interpretability. Thus, I developed a new approach that employing physical model to guide the network construction of deep learning, which addressed this problem well and achieved good performance.

Spatial-spectral fusion: Spatial-spectral fusion is a very productive technique to integrate fine spatial details and rich spectral information together. Embedding physical mechanisms into deep learning, I proposed a novel fusion model, and the fused data can be used to extract finer vegetation index. Furthermore, with the experience of studying on spectral super-resolution, I designed an efficient ground truth-free fusion framework through self-learning and multi-task collaboration.

Multi-task collaboration: There are various image quality improvement tasks in remote sensing. In general, people address different tasks separately while ignore the inter-task relevance. Considering the model similarity between spectral degradation and spatial-spectral degradation model, I developed a general spectral super-resolution framework that can handle spectral super-resolution with spatial degaradation. Meanwhile, I also use spectral super-resolution to enhance the spatial details in spatial-spectral fusion.

Glacier semantic extraction: Glaciers serve as sensitive indicators of climate change, making accurate glacier boundary delineation crucial for understanding their response to environmental and local factors. We propose a Transformer-based deep learning approach to effectively extract glacier by using multi-source data.

Previous Projects

1 Spectral response function–guided deep optimization-driven network for spectral super-resolution (He et al., 2022a)

Early works about spectral super-resolution all concern on dictionary learning, which can be regarded as the model-based algorithms, because they are derived from spectral degradation models. Deep learning-based methods are always blamed for their low physical interpretability.

We started from the spectral degradation model and gived a model-based solutions with the help of *Alternating Direction Method Of Multipliers* (ADMM). Considering the similarity between the mathematical equation of the prior sub-problem and denoising tasks, we utilized a data-driven *Convolutional Neural Network* (CNN) to learn the hyperspectral prior implicitly and use 1×1 convolutional layers to simulate spectral degradation. Thus, we can build a end-to-end CNN following the dataflow of the model-based solutions, which is now known as deep unfolding. At the time, it was the first deep-unfolding network proposed in spectral super-resolution.

Furthermore, we found that the spectral correlation between adjacent bands is strong, while the correlation between distant bands is poor. Directly recovering all bands with bands through the whole wavelength may lead to the mutual interference of spectral information. According to spectral response functions, we can easily group the bands with spectral similarity, which improves the performance greatly and reduces the runtime.

2 A knowledge optimization-driven network with normalizer-free prior for spatial-spectral fusion (He et al., 2022c)

Similar to spectral super-resolution, normal deep learning-based spatial-spectral fusion algorithms follow an end-to-end operation and run in a black-box manner. We proposed a knowledge optimization-driven network to address spatial-spectral fusion, where the degradation models are used as the knowledge guide. Besides, we also designed a *Normalizer-Free* (NF) ResNet prior with out *Batch Normalization* (BN). On the one hand, BN incurs computation memory overhead and breaks the independence of distributions between training examples within a batch, while, on the other hand, without BN, the training of ResNet would be unstable.

Let x_i^l present the i-th batch features after l-th residual block, and $x_i^{l+1} = x_i^l + f(x_i^l)$ always holds in residual blocks, where $f(\cdot)$ denotes the residual branch. Focusing on the variance in ResNet, we can find that the variance of the activations before and after the residual block satisfy:

$$\mathbf{Var}(x_i^{l+1}) = \mathbf{Var}(x_i^l) + \mathbf{Var}(f(x_i^l)) = \mathbf{Var}(x_i^l) + 1 \tag{1}$$

With good variance preserving of weight-scaled convolutions, we change the ResNet from $x_i^{l+1} = x_i^l + f(x_i^l)$ into $x_i^{l+1} = x_i^l + \alpha f(x_i^l/\beta_l)$, where α denotes the rate of variance growth, and β_l is fixed as $\sqrt{\mathbf{Var}\left(x_i^l\right)}$. In this way, the variance $\mathbf{Var}\left(x_i^{l+1}\right)$ will be changed into:

$$\mathbf{Var}(x_i^{l+1}) = \mathbf{Var}(x_i^l) + \alpha^2\mathbf{Var}(f(x_i^l/\beta_l)) = \mathbf{Var}(x_i^l) + \alpha^2 \tag{2}$$

Eq (2) has a similar variance growth with BN-ResNet which is just multiplied with α^2. With such a strategy, we could build a ResNet without BN but with similar variance growth as BN-ResB, which keeps the strong prior learning ability of ResNet but does not introduce more computational cost.

3 A optimization-based network addressing spectral super-resolution and spatial-spectral fusion simultaneously (He et al., 2022d)

Spectral super-resolution is a very important technique to obtain hyperspectral images from only multispectral images. However, in practice, multispectral channels or images captured by one satellite are often with different spatial resolutions, which brings a severe challenge to spectral super-resolution.

Building a degradation model considering both spectral and spatial degradation, we proposed a new concept called generalized spectral super-resolution, and solved it in a model-driven way. According to our previous works, we deep unfolded it into an end-to-end CNN. Besides, we proposed an efficient cross-dimensional channel attention with 1D and 2D convolutions.

4 Spectral super-resolution injects more spatial details for ground truth-free spatial-spectral fusion (He et al., 2023c)

The fine spatial detail in spatial-spectral fusion is the priority among priorities. Without ground truth as training samples, deep learning-based methods can hardly achieve a satisfactory result. The key difficulty is that spatial details between panchromatic and multispectral images differ both in spatial resolution and spectral domain. Based on our previous works about spectral super-resolution, we found that spectral super-resolution only improve the spectral resolution while maintain spatial information.

With the help of spectral super-resolution injection, our methods shows the transcendental performance of ground truth-free spatial-spectral fusion. The proposed spectral super-resolution injection can be easily deployed on existing fusion methods and further improve their spatial details. In this work, we recognized the universality and individuality of multiple remote sensing tasks, thereby building a new future directions in multi-task collaboration.

5 Glacier extraction using Transformer from multi-source data (Peng et al., 2023)

Extracting accurate glacier boundary delineation is crucial for understanding their response to environmental and local factors. However, traditional semi-automatic remote sensing methods for glacier extraction can hardly fuse local-global information and fail to leverage multi-source data.

We propose a Transformer-based deep learning approach to address these limitations. Our method employs a U-Net architecture with a Local-Global Transformer encoder and multiple Local-Global CNN Blocks in the decoder, which can integrate both global and local information. The multi-source data include Sentinel-1 SAR data, Sentinel-2 multispectral data, High Mountain Asia DEM, and Shuttle Radar Topography Mission DEM. In this work, our findings highlight the significant contribution of Sentinel-2 data to glacier extraction.

Current and Future Projects

1 Remote sensing computational imaging

The previous works about satellite image quality enhancement only involve the post-processing after image acquisition. However, using principles of rapid computational imaging combining the front-end optics and post-detection signal processing, we can design an instrument to make optical measurements from which images, and other scene information, can be derived with a capacity that far exceeds the physical limits of the optics.

Dual super-resolution helps spatial-spectral fusion: In our future work, we plan to build a two-stream model with spectral super-resolution branch and spatial super-resolution branch. Dual super-resolution improve the spatial details and spectral fidelity respectively. The further concern is how to build spatial-spectral constraints. There are two assumptions. The first is following the spectral constraint in (**He** et al., 2023c), where a simulated degradation module and self-supervised learning is employed to degrade the fused results back to the original low-resolution domain. The second is directly using the degradation modules learned in the two-stream model. The spatial degradation module in spatial super-resolution branch can be used to build spectral constraint in spectral super-resolution branch, while the spectral degradation module in spectral super-resolution branch can be used to build spatial constraint in spatial super-resolution branch.

Spectral super-resolution & Cloud Removal: High-resolution hyperspectral images is demanded for precision agriculture, ecological factor estimation, and environmental monitoring, while various degradation appear during imaging processing, including spatial degradation and clouds. Traditional algorithms improve the spatial details and remove clouds respectively. Considering the feature similarity in spectral super-resolution and cloud removal, we plan to utilize a universal framework to address spectral super-resolution and cloud removal simultaneously. Fed with multi-temporal cloudy multispectral images, we hope that the proposed algorithms can recover cloud-free hyperspectral images.

2 Geoscience low-high-level task collaboration

Traditional geoscience tasks can be divided into two groups similar to computer vision field, including low-level geoscience tasks and high-level geoscience tasks. Low-level geoscience tasks only aims to improve the data quality and achieve fine Earth observation. And high-level geoscience tasks involves the physical meaning and semantic information, which provide more information for us to explore and understand our planet. Low-high-level task collaboration aims to achieve mutual promotion between low-level and high-level geoscience tasks.

Spectral super-resolution & Fine classification: Rich spectral information plays an essential role both in spectral super-resolution and fine classification. Thus, there must be the great potential to build a model sharing spectral features. Besides, the category semantic information in classification helps spectral super-resolution recover more accurate spectra. On the contrary, spectral super-resolution provides more texture and continuous spectral information for fine classification (**He** et al., 2022b).

Time series enhancement of vegetation products: With a successful track record of recovering continuous spectral information from low-spectral channels, our next step is to enhance the temporal information of vegetation products. The idea is motivated by the similarity between the spectral curve and plant growth curve. In multiple life cycles, plants grows and go to death while the growth curves change periodically. Thus, the temporal information in vegetation products can be learned from complete vegetation database.

3 Quality improvement jointed with quantitative retrieval

Quantitative retrieval is a crucial technique in Earth observation, which provides more large-scale and long-time-series phenological information for people to make decisions. Unfortunately, as it requires satellite images and products for data source, the retrieval precision is always limited by the quality of satellite images and products. Addressing them separately fails to exploit their relevance. Thus, I draw my last future research route on quality improvement jointed with quantitative retrieval.

Spectral super-resolution & Vegetation index retrieval: Vegetation index retrieval from optical remote sensing data requires rich spectral information, while hyperspectral data is not always readily available. With low-cost multispectral images, sharing the same spectral features mutually benefits spectral super-resolution and vegetation index retrieval. The early work may concern on leaf area index, and the further works would involve net primary productivity, gross primary productivity, and crop yield prediction.

Glacier area calculation & Glacier volume prediction: After extract glacier area, glacier volume is also a crucial ecological factor for detecting global climate change and water cycle. Traditional methods assumed that glacier is a motionless rigid body and glacier volume only depends on the glacier area and mass balance (melting and accumulation of glaciers). However, in fact, the ice flowing would also influence the final glacier volume. Thus, we would like to introduce a glacier dynamics model with machine learning to extract glacier area and predict glacier volume.

Unity of AI-driven Earth data quality improvement: Earth data involves remote sensing images and products. Improving their quality provides more accurate and continuous data source for exploring and understanding our planet. Artificial intelligence (AI)-driven remote sensing image enhancement and remote sensing product quality improvement are always regarded as two independent fields. In our views, they can be involved into AI-driven Earth data quality improvement. Our future work is to explore the opposition and unification between remote sensing image enhancement and product quality improvement.

Published Articles

L.-J. Deng, G. Vivone, M. E. Paoletti, G. Scarpa, J. **He**, Y. Zhang, J. Chanussot, and A. Plaza. Machine learning in pansharpening: A benchmark, from shallow to deep networks. *IEEE Geosci. Remote Sens. Mag.*, 10(3):279–315, 2022. [PDF].

X. Jin, J. **He**, Y. Xiao, and Q. Yuan. Learning a local-global alignment network for satellite video super-resolution. *IEEE Geosci. Remote Sens. Lett.*, 20:1–5, 2023.

D. Liu, J. Li, Q. Yuan, L. Zheng, J. **He**, S. Zhao, and Y. Xiao. An efficient unfolding network with disentangled spatial-spectral representation for hyperspectral image super-resolution. *Inf. Fusion*, 94:92–111, 2023.

Y. Peng, J. **He**, Q. Yuan, S. Wang, X. Chu, and L. Zhang. Automated glacier extraction using a transformer based deep learning approach from multi-sensor remote sensing imagery. *ISPRS J. Photogramm. Remote Sens.*, 202:303–313, 2023. [PDF].

J. **He**, J. Li, Q. Yuan, H. Li, and H. Shen. Spatial–spectral fusion in different swath widths by a recurrent expanding residual convolutional neural network. *Remote Sens.*, 11(19):2203, 2019.

J. **He**, J. Li, Q. Yuan, H. Shen, and L. Zhang. Spectral response function-guided deep optimization-driven network for spectral super-resolution. *IEEE Trans. Neural Netw. Learn. Syst.*, 33(9):4213–4227, 2022a. [PDF].

J. **He**, Q. Yuan, J. Li, Y. Xiao, X. Liu, and Y. Zou. Dster: A dense spectral transformer for remote sensing spectral super-resolution. *Int. J. Appl. Earth Observ. Geoinf.*, 109:102773, 2022b. [PDF].

J. **He**, Q. Yuan, J. Li, and L. Zhang. A knowledge optimization-driven network with normalizer-free group resnet prior for remote sensing image pan-sharpening. *IEEE Trans. Geosci. Remote Sens.*, 60:1–16, 2022c. [PDF].

J. **He**, Q. Yuan, J. Li, and L. Zhang. Ponet: A universal physical optimization-based spectral super-resolution network for arbitrary multispectral images. *Inf. Fusion*, 80:205–225, 2022d. [PDF].

J. **He**, Q. Yuan, and J. Li. Generalized spectral super-resolution for multispectral satellite imagings. *Acta Photonica Sinica*, 52(2): 0210002–0210002, 2023a.

J. **He**, Q. Yuan, J. Li, Y. Xiao, D. Liu, H. Shen, and L. Zhang. Spectral super-resolution meets deep learning: achievements and challenges. *Inf. Fusion*, 97:101812, 2023b. [PDF].

J. **He**, Q. Yuan, J. Li, Y. Xiao, and L. Zhang. A self-supervised remote sensing image fusion framework with dual-stage self-learning and spectral super-resolution injection. *ISPRS J. Photogramm. Remote Sens.*, 204:131–144, 2023c. [PDF].

Y. Xiao, Y. Wang, Q. Yuan, J. **He**, and L. Zhang. Generating a long-term (2003-2020) hourly 0.25° global pm2. 5 dataset via spatiotemporal downscaling of cams with deep learning (deepcams). *Sci. Total Environ.*, 848:157747, 2022a.

Y. Xiao, Q. Yuan, J. **He**, Q. Zhang, J. Sun, X. Su, J. Wu, and L. Zhang. Space-time super-resolution for satellite video: A joint framework based on multi-scale spatial-temporal transformer. *Int. J. Appl. Earth Observ. Geoinf.*, 108:102731, 2022b.

Y. Xiao, Q. Yuan, K. Jiang, J. **He**, Y. Wang, and L. Zhang. From degrade to upgrade: Learning a self-supervised degradation guided adaptive network for blind remote sensing image super-resolution. *Inf. Fusion*, 96:297–311, 2023.

L. Zhang, J. **He**, Q. Yang, Y. Xiao, and Q. Yuan. Data-driven multi-source remote sensing data fusion: Progress and challenges. *Acta Geodaetica et Cartographica Sinica*, 51(7):1317–1337, 2022. [PDF].