

Special focus on deep learning in remote sensing image processing*

As a newly emerging technology, deep learning is a very promising field in big data applications. Remote sensing applications often involve huge volume data obtained daily by numerous in-orbit satellites. This makes it a perfect area for data-driven applications. Over the past years, there has been an exponentially increasing interest in deep learning for remote sensing image processing, including not only optical imagery but also synthetic aperture radar (SAR) imagery. In addition to the rapidly growing size and spectral, spatial and temporal resolution of remote sensing data, there are other challenges that are unique in this area, e.g. the intrinsic complexity and particularity of each specific sensor and their multi-modality, the fundamental physical properties embedded in the data, the underlying principles for information retrieval, etc. To promote the research in this area, we have organized a special focus feature on deep learning in remote sensing image processing in the *SCIENCE CHINA Information Sciences*.

Spatio-temporal fusion (STF) aims at fusing (temporally dense) coarse resolution images and (temporally sparse) fine resolution images to generate image series with adequate temporal and spatial resolution. In the contribution entitled “Spatio-temporal fusion for remote sensing data: an overview and new benchmark”, Li et al. provide a comprehensive overview on the development of this field. Most importantly, a robust benchmark STF dataset that includes three important characteristics: 1) diversity of regions, 2) long timespan, and 3) challenging scenarios, is introduced. In the contribution “A new sensor bias-driven spatio-temporal fusion model based on convolutional neural networks”, Li et al. propose a STF model to mitigate the differences between the spectral and spatial distortions in traditional methods.

One of the major challenges for the application of deep learning to SAR remote sensing lies in the limited availability of training samples. In the contribution “FUSAR-Ship: building a high-resolution SAR-AIS matchup dataset of Gaofen-3 for ship detection and recognition”, Hou et al. build the FUSAR-Ship SAR-AIS dataset for Gaofen-3 (GF-3) radar satellite ship detection and recognition. It uses 126 scenes of GF-3 SAR images matched-up with ship information from the automatic identification system (AIS). It has over 6000 annotated chips of ship SAR images which could be used as benchmark dataset for ship detection and recognition.

Radar remote sensing also promotes ecological research, such in predicting large-scale airborne animal migration. In the contribution “Deep-learning-based extraction of the animal migration patterns from weather radar images”, Cui et al. use convolutional neural networks to classify and segment weather radar images to extract biological information. Through the analysis of massive weather radar data, airborne animal migration patterns such as intensity and speed are partly revealed.

In the contribution “Hybrid first and second order attention Unet for building segmentation in remote sensing images”, He et al. deal with the problem of building segmentation in aerial images. The proposed method is featured with the hybrid first and second order attention Unet, which explores the inner-product among different channels. Its performance is demonstrated on three widely-used datasets.

The lack of samples in hyperspectral images limits the performance of classification, while the existing methods for hyperspectral image sample generation often disregard the integrity of spatial-spectral

*Citation Xu F, Hu C, Li J, et al. Special focus on deep learning in remote sensing image processing. *Sci China Inf Sci*, 2020, 63(4): 140300, <https://doi.org/10.1007/s11432-020-2810-x>

information. In the contribution “Cascade conditional generative adversarial nets for spatial-spectral hyperspectral sample generation”, Liu et al. propose a cascaded GAN to generate samples containing detailed spatial-spectral information from hyperspectral images to supplement training samples.

Satellite video processing is also a new and emerging topic in remote sensing. In the contribution “Deep feature extraction and motion representation for satellite video scene classification”, Gu et al. propose a two-stage satellite video scene classification method in which both image-domain spatial features and target motion’s temporal features are extracted. These two kinds of features are fused and used in a scene classification problem involving video data acquired by Jilin-1 and UrtheCast satellites.

Guest Editors:

Feng XU

Fudan University, China

Cheng HU

Beijing Institute of Technology, China

Jun LI

Sun Yat-sen University, China

Antonio PLAZA

University of Extremadura, Spain

Mihai DATCU

German Aerospace Center, Germany