Paper ID: 69647
student

IAF EARTH OBSERVATION SYMPOSIUM (B1)
Interactive Presentations - IAF EARTH OBSERVATION SYMPOSIUM (IP)

Author: Dr. Lingyun Gu
School of Aerospace, Tsinghua University, Beijing, China

Dr. Eugene Popov
Peter the Great St. Petersburg Polytechnic University, Russian Federation
Prof. Ge Dong
Tsinghua University, China

FAST FOURIER CONVOLUTION BASED REMOTE SENSOR IMAGE OBJECT DETECTION FOR EARTH OBSERVATION

Abstract

Remote sensor image object detection is an important technology for Earth observation, and is used in various tasks such as forest fire monitoring and ocean monitoring. Image object detection technology, despite the significant developments, is often struggling to handle large remote sensor images and small-scale objects, due to the limited pixels of small objects. In this paper, a global receptive field-based image object detection for Earth observation is proposed. By extending the receptive field of networks, features from surroundings of objects are introduced as auxiliary information to help detect small objects. For clarity, the main contributions of this work can be summarized as: 1) A detection network based on fast Fourier convolution (FFC) is proposed. FFC is able to extract global features by its unique Fourier unit, the details are: first transform the input image to the frequency domain by fast Fourier transform (FFT), then the convolution block is performed to extract the frequency features, obviously the convolution has a global receptive field in the spatial domain covering the whole image, and finally the features are reduced to the spatial domain using the inverse FFT. 2) A cross-layer connected module is proposed to achieve information enhancement of small objects by context information and thus alleviate the disturbance of low resolution. Context features from surrounding information of objects are extracted from the higher layers, and then connected to the object features as auxiliary information to assist in object detection. 3) A multi-spectral channel attention module is proposed for enhancing the features of the object itself. While traditional attention mechanisms utilize only the lowest frequency domain features, the multichannel attention module introduces rich information by adding features of the high frequency components, thus enhancing the features of small targets. A remote sensor image dataset is proposed to demonstrate the effectiveness of the approach, which contains 1560 images from several satellites such as Sentinel-1, TerraSAR-X, and 20 classes of objects such as ships, aircraft, and ports. On this dataset, the method achieved an average recall of 92.2% and a Mean Average Precision (MAP) of 85.6%, proving that it is suitable to the mission of remote sensor image object detection for Earth observation. Notably, the data processing of this method is mainly in the frequency domain, so it has a good potential to migrate to raw remote sensing data in future to avoid time-consuming imaging operations.