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

Author: Dr. Wei Wan
China Center for Resources Satellite Date and Application (CRESDA), China, wanwei829@163.com

Mr. Yechao Sun
China Aerospace Science and Technology Corporation (CASC), China, wanwei829@163.com
Dr. Xuetao Hao
China Aerospace Science and Technology Corporation (CASC), China, wanwei829@sina.com
Prof. Xiangjun Min
CASC, China, wanwei829@sina.com

DEEP LEARNING FROM HIGH RESOLUTION EARTH OBSERVATION DATA TO EXTRACT SEMANTIC INFORMATION

Abstract

Deep learning for computer vision becomes attractive topic in recent years, especially because of Convolutional Neural Networks (CNN) that are able to learn powerful and expressive descriptors from images for a wide range of tasks. Moreover, new network architectures derived from CNN have appeared, such as Fully Convolutional Networks (FCN), able to output pixel-wise labelling. These networks have quickly become state-of-the-art for achieving fine-grained image classification. In the Earth Observation(EO) filed, the deep network methods are now especially appealing, as dense prediction allows us performing semantic mapping without requiring pre-processing work other than labelling training data. In this work, we want to show that deep learning is suitable for accurate semantic segmentation of EO data. We study deep network architectures that leverage different methods of model optimizing, data augmentation, parameter tuning, and filed expertise. Especially, we look into FCN-based network and coarse-to-fine segmentation, and investigate the use of geospatial data as an enhanced input layer to be included into multi-spectral band data. We present an end-to-end solution for learning several varieties of structures in high resolution images. We design a network with a final stage that integrates activations from multiple preceding stages for pixel-wise prediction, and introduce the geospatial data and index function of structure as the input representations, which has an enhanced representation power. The use case is automatic building and road extraction from two high spatial resolution datasets: aerial and satellite imagery, which is challenging due to large variations of building appearances. The trained network achieves good performance on test data, demonstrating that the proposed method provides a promising solution for automating such labor-intensive task. The results also show that knowledge of remote sensing expertise and geospatial data can be efficiently integrated into the vision-based deep learning model, and that it can improve the accuracy and the convergence speed of the networks. Therefore, by using deep learning means we can shift from traditional region-based classification to fully supervised semantic segmentation, for land cover classification, object detection and scene understanding in EO images to achieve better results.