Distributed Computing Framework for Ship Detection from Panchromatic VNREDSat-1 Satellite Images

T. Q. Luu[1], **H. V. Luu**[2], **T. T. N. Nguyen**[3], **and H. Q. Bui**[4]

Abstract: In the past decades, satellite images for Earth Observation (EO) have been largely using in many research studies such as Land Cover Land User Change, Air Pollution Forest Fire, Urban Classification, or Ship detection, etc. Together with technology development, amount of satellite data captured for EO, in terms of diversity and volume, is increasing dramatically. Therefore, innovative tools, processing methodologies and development frameworks are requested in order to collect, store and process satellite data efficiently. In this paper, we propose a distributed storage and processing architecture for satellite images based on Hadoop and Spark, namely Hadoop-Spark cluster. A test-bed of Hadoop-Spark cluster is conducted for ship/boat detection from panchromatic VNREDSat-1 images recorded by the first Vietnamese satellite launched in 2014. Each panchromatic VNREDSat-1 image has about one hundred million pixels at Very High Resolution (VHR) (i.e. 2.5 meters). The distributed ship detection algorithm, consists of three steps: anomaly detection, segmentation and classification, is designed to run on a Hadoop-Spark cluster. To evaluate the performance of the proposed platform, we compare the running time of ship detection algorithm on a single node and a cluster, in terms of input sizes, number of satellite images, number of processors, and Hadoop-Spark combination modes. Experiments have shown promising results were running time is significantly improved on ship detection from VNREDSat-1 images by the proposed Hadoop-Spark cluster.

[1,2,3,4] Center of Multidisciplinary Integrated Technology for Field Monitoring
VNU University of Engineering and Technology
144 Xuan Thuy, Cau Giay, Hanoi, VIETNAM
{*thanglq, hunglv, thanhntn, hungbq*}*@fimo.edu.vn*