SCALABLE EVOLUTIONARY COMPUTATION FOR EFFICIENT INFORMATION EXTRACTION FROM REMOTE SENSED IMAGERY

L. M. Almutairi[a], S. Shetty [a], H. Momm [b]

[a] Dept. of Electrical and Computer Engineering, Tennessee State University, Nashville, TN 37209 USA-
(lalmutai ,sshetty)@tnstate.edu
[b] Department of Geosciences, Middle Tennessee State University, Murfreesboro, TN 37132 USA Henrique.Momm@mtsu.edu

Commission VI, WG VI/4

KEY WORDS: evolutionary computation, distributed computing, Hadoop, MapReduce, remote sensing, information extraction

ABSTRACT:

With the ever-increasing scale of remote sensing imagery, it is paramount to develop scalable methods to extract specific information from imagery to generate geospatial intelligence in a timely, consistent, and cost-effective way. To address this task, novel biologically inspired framework was developed to evolve solutions composed of multiple image-derived cues combined in a non-linear way and designed to increase the spectral separation between features of interest and the remaining image background. Despite the framework's documented success, its adoption by practitioners is limited due to the large computational cost involved during the supervised solution development mode. In this mode, the algorithm has to calculate a score value for all candidate solutions for each iteration of the process. With problems involving tens of thousands of candidate solutions and thousands of iterations, this process could take between hours to days. This research enhanced the existing framework through the utilization of distributed computing technology to expedite the supervised solution development stage. The evolutionary computation framework to extract materials/features of interest from the original remotely sensed imagery were implemented on a MapReduce/Hadoop based distributed computing platform. Preliminary results indicate that distributing and parallelizing the task of evaluating all candidate solutions can significantly reduce the computational cost involved during the solution development mode. Preliminary results demonstrated that increasing the number of distributed map tasks for the first generation of candidate solutions decreases elapsed real time by more than 91%. The expected outcome of this project will be a robust, non-linear system for analysis of multi-band imagery and will constitute an attractive alternative for practitioners interested in information extraction from remotely sensed imagery.