# Parallel Image Segmentation Using Map-Reduce Framework

Mohammad Nishat Akhtar, Junita Mohamad Saleh*, Elmi Abu Bakar and Ayub Ahmed Janvekar

***Abstract*—** As a result of the expansive information set size of high-resolution image data, most desktop workstations do not have sufficient configurable scheduling to perform image processing assignments in a convenient manner due to which the image processing tasks are meant to be divided into straight forward assignments. The processing power of any regular computing machine in this way becomes a severe bottleneck with respect to high execution time and low throughput. Many image processing tasks exhibit a high level of information region and parallelism and map quite readily to a parallel computing system. This paper shows an alternative to sequential image processing by introducing Map-Reduce technique to segment multiple images with the help of Hadoop framework. The evaluation of the proposed scheduling algorithm is done by implementing parallel image segmentation algorithm to detect lung tumor for up to 1 GB size of CT image dataset. The results have shown improved performance with parallel image segmentation when compared to sequential image segmentation method particularly when data capacity reaches a particular threshold. This is because the process of parallel image processing has been able to exploit the multi-cores thread level parallelism which ultimately gave the CPU usage with octacores up to 96%, hence reducing the task execution time up to approximately 1.6 times compared with the sequential style of image segmentation using Map-Reduce algorithm implemented with FIFO scheduler. The proposed parallel image segmentation design has shown to be useful for researchers at performing bulk image segmentation in parallel, which can save tremendous execution time.

***Keywords*—** Hadoop, Execution Time, Task parallelism, Parallel Image Segmentation, Map, Reduce.

## I. Introduction

High end computing machines have not been savvied enough (as far as the necessary equipment and programming speculation) to increase across the broad usage. Maybe, it appears that in near future, parallel computing will be dominated by medium-grain distributed memory machines in which every processing node will have the capabilities of a desktop workstation [1]. In reality, as network innovations keep on maturing, bunches of workstations are themselves being progressively seen as a parallel computing asset. The upsides of medium-grain standard computing are low cost and high quality. The disadvantage comprises irregular load designs on the processing nodes [2]. The proposed research depicts the outline and implementation of parallel image segmentation using Hadoop framework. It is also worth to be noted that Hadoop framework is not based on the model of Message Passing Interface (MPI) standard and is specifically designed to support parallel execution on heterogeneous workstation nodes [3, 4]. Many image processing algorithms exhibit natural parallelism in a sense that the input image data required to compute a given portion of the output is spatially localized and is compatible to be implemented on a cloud framework [5, 6]. In the simplest case, the output image could be computed simply by independently processing single pixels of the input image.

Image Processing with parallel computing is a viable approach to take care of image processing issues that require extensive processing time [6, 7]. It is evident that restorative imaging requires heaps of memory space and time to process, so by parallelizing, it would be helpful to discover productive and quick outcome. In parallel processing, a program can make numerous assignments that cooperate to take care of the issue of multi-tasking [8]. Parallel image processing cannot be connected to all issues, or in other words it can be stated that not every one of the issues can be coded in a parallel shape. A parallel program ought to must have a few elements for a right and proficient operation; else, it is conceivable that run-time may not have the normal execution. These components incorporate the processing parameters such as granularity, coarse grained and fine-grained parallelism [9]. The remaining parts of this manuscript are arranged as follows. Section II highlights the background for parallel image segmentation. Section III describes the model for multiple image segmentation simultaneously. Section IV describes and demonstrates the proposed parallel image segmentation algorithm along with illustration of mapper and reducer for parallel image segmentation. Section V shows the results and discussion followed by a conclusion in Section VI.

Mohammad Nishat Akhtar is with School of Aerospace Engineering, Universiti Sains Malaysia, 14300, Nibong Tebal, Penang, Malaysia (email: nishat.akhtar2000@gmail.com)
*Junita Mohamad Saleh is affiliated with School of Electrical and Electronics Engineering, Universiti Sains Malaysia, 14300, Nibong Tebal, Penang, Malaysia (e-mail: jms@usm.my; Tel: +60194732732)
Elmi Abu Bakar is affiliated with School of Aerospace Engineering, Universiti Sains Malaysia, 14300, Nibong Tebal, Penang, Malaysia (email: meelmi@usm.my)
Ayub Ahmed Janvekar is affiliated with School of Mechanical Engineering, Universiti Sains Malaysia, 14300, Nibong Tebal, Penang, Malaysia (email: ayubjanvekar@gmail.com)