Data Overloading in Medical Imaging: Emerging Issues, Challenges and Opportunities in Efficient Data Management

K. H. Ng1,∗, Oliver Faust2, Vidya Sudarshan3, and Subhagata Chattopadhyay4

1Department of Biomedical Imaging University of Malaya, Kuala Lumpur Malaysia
2School of Science and Engineering, Habib University, Karachi, 75350, Pakistan
3Department of Electronics and Computer Engineering Ngee Ann Polytechnic, Singapore
4Department of Computer Science and Engineering National Institute of Science and Technology Palur Hills, Berhampur 761008, Orissa, India

Imaging techniques have gained tremendous popularity and pace in current trend of medical diagnostics. The key reason lies in the fact that such techniques extract clinical information with higher speed and accuracy compared to manual diagnosis by doctors. Moreover, the images serve as important records/evidences of diseases, hence the images are stored for the future use. However, among these advantages, there are growing issues with widespread applications of imaging techniques. To be specific these issues are data (i) overloading, (ii) security/privacy, (iii) storage, (iv) transfer (especially, across institutes, regions, countries). In this paper we review the individual sources of medical image data. Based on this review, we discuss enabling technologies and present Computer-Aided Diagnosis (CAD) as well as data mining as coping strategies for the overwhelming amount of digital image data encountered in modern clinical work flows.

Keywords: Data Overloading, Medical Imaging, Systems Thinking, Workflow, Infrastructure, Data Management.

1. INTRODUCTION
Analytical science observes nature and analyzes the data from these observations. The drive is to realize more relevant data to inject precision in decision making amidst an overwhelming data pool. Biomedical imaging techniques are specialized types of analytical science and therefore this area is no exception to the general drive for more precise data. It also strives towards extracting relevant information that could accurately diagnose a disease or predict its prognosis. For example, surgeons go through a large number of images as part of the surgery planning process. In particular before the neurosurgery, they have a Functional Magnetic Resonance Imaging (MRI) done to plan the surgery and future course of management. There are several such practices in other disciplines as well.

For biomedical imaging systems, the popular term progress∗ means to produce an ever increasing amount of image data. In technical terms, it denotes increasing both sensitivity and resolution while decreasing scanning time and cost. Hence, sensitivity, resolution, time, and cost are considered to be four performance parameters of imaging. Sensitivity can be related to the number of computer bits spent to encode an intensity level for each pixel and the resolution is a measure for the number of pixels in an image. Together it renders the level of visual elaboration of a given clinical condition and hence the diagnostic or prognostic accuracy. Reducing the scanning time means that, it is possible to scan more patients in a unit time and reducing the cost refers to the cost of establishment from the vendor’s and treatment from the patients’ perspective. Taken all these parameters into account, the net result is the intended maximization of its use by healthcare professionals and in turn production of high volume of image data. This phenomenon could be termed as data overload. In 2008 it has been estimated that all kinds of medical images occupied 1250 Petabytes. In comparison, the giant Internet search company Google used an estimated disk space of 20 to 200 Petabytes in August 2009. To calculate storage capacity in Petabytes is cutting edge, the first industrial vendor to produce such a Petabyte storage array was launched in Jan 2006 at a cost of around US$4 million.1

In order to understand the kind of data overloading that we are facing today, we need to analyze the data volume per procedure. For example, a simple Digital Radiography (DR) such as a chest X-ray takes up about 30 Mb (Megabyte). Mammogram scans