Chapter 3
Evolution of Big Data in Medical Imaging Modalities to Extract Features Using Region Growing Segmentation, GLCM, and Discrete Wavelet Transform

Yogesh Kumar Gupta
https://orcid.org/0000-0002-4572-178X
Banasthali Vidyapith, India

ABSTRACT

Big data refers to the massive amount of data from sundry sources (gregarious media, healthcare, different sensor, etc.) with very high velocity. Due to expeditious growth, the multimedia or image data has rapidly incremented due to the expansion of convivial networking, surveillance cameras, satellite images, and medical images. Healthcare is the most promising area where big data can be applied to make a vicissitude in human life. The process for analyzing the intricate data is mundanely concerned with the disclosing of hidden patterns. In healthcare fields capturing the visual context of any medical images, extraction is a well introduced word in digital image processing. The motive of this research is to present a detailed overview of big data in healthcare and processing of non-invasive medical images with the avail of feature extraction techniques such as region growing segmentation, GLCM, and discrete wavelet transform.

DOI: 10.4018/978-1-7998-2795-5.ch003

INTRODUCTION

In this digital, the astronomically immense data is a very critical quandary because of the tremendous amount of data engendered in routine life from sundry sources such as online transaction, e-mails, research journals and articles, convivial media sites (Facebook, Twitter, WhatsApp etc.) and web forums, different sensor's data composed from sundry sources such as healthcare science or medical data, environmental organizations, meteorological department, business strategically data, trading market, company data being engendered daily life in different format such as structured, semi-structured and unstructured with a great velocity is customarily referred to as sizably voluminous Data (Singh, 2016). Data can be engendered on web in sundry forms like texts, audios, videos, images, texture or gregarious media posts data etc. This tremendous data is no more time stable in environment; rather it is updated according to time at rapid celerity. So that put an immensely colossal number of critical challenges on sizably voluminous data processing and storage. As an outcome, the conventional database computation implements and algorithms as well as data storage and management techniques has not able to deal with these data (Altera, 2016). Thus the astronomically immense data need to describe sundry innovative storage and processing implements (such as Hadoop, MapReduce, NoSQL database, HPCC and Apache Hive etc.) to acquire, store, distribute, handle and analyze. The process for analyzing the perplexed data mundanely concerned with the disclosing the hidden patterns. Big data involves sizably voluminous distributed file systems in commodity hardware for storing, which should be more flexible, fault tolerant, scalable and reliable (Singh, 2016).

This digitized era needs to process the images for the manipulation of image properties. As immensely colossal data comes with the flood of intricate and digitize data so as to manage this data, there is need to have a discussion in an optimistic manner which describes its valuable and valid positive aspects. Healthcare industry conventionally has engendering astronomically immense amounts of data. As authors ken that historically the data stored in hard copies but nowadays there is a rapid digitization of this massive amount of stored data. The immensely colossal data analysis holds a broad accumulation of medical and healthcare applications such as public health management, clinical assessment and disease scrutiny. Big data analytics in healthcare has the puissance to decrement the costs of treatment, amend care, preserve lives and additionally ameliorate the quality of life.

From the commencement of evolution to 2003, only 5 Exabyte's of information has engendered, currently by 2012 the authors engender that equipollent amount in just two days, if authors consider data of digital macrocosm that will cultivate to 2.72 Zeta bytes and by 2015 that will twice over every two years to reach 8 Zeta bytes (Tamilselvan, 2015).