

Cloud Based Medical Image Data Analytics in Healthcare Management

R.A. Karthika^{1*}, K. Dhinakaran², D. Poorvaja³, A.V. Shanbaga Priya⁴

¹Associate Professor, Department of Computer Science and Engineering, Vels Institute of Science, Technology & Advanced Studies.

²Assistant Professor, Department of Computer Science and Engineering, Rajalakshmi Institute of Technology.

³UG Scholar, Department of Computer Science and Engineering, Rajalakshmi Institute of Technology.

⁴UG Scholar, Department of Computer Science and Engineering, Rajalakshmi Institute of Technology.

Abstract

In today's world, the images form a huge amount of unstructured data from the public and the corporate sector. As a result of the growth of these types of data, modern analytical systems need to interpret and assimilate images. This brings in the need of image processing which involves the transformation from images to analytically organized and structured data. It performs required operation on the given input image and returns the related outputs based on the query. Digital image processing has pushed the envelope for the appraisal in various domains such as healthcare, defense and security, remote sensing, robotic visions, pattern recognitions and satellites. The complication involved in the healthcare domain makes it suitable to explore and induce the concept of image processing and increases the potential for prescriptive analytics. The cloud combined with the Image Processing provides the best environment for analyzing these images using the proposed technique. The proposed work involves processing the input images using the cloud data analytics that provides a user-friendly environment and retrieves the relevant images as well as text for the given user query which produces the outputs that are more efficient in terms of parameters like time, size, security and speed comparatively with the existing data mining process.

Keywords: Cloud based medical image, data analytics, healthcare.

1. Introduction

Data analysis refers to the breaking of a whole data into its separate individual components for separate examinations. Analysis of data involves the process of inspecting the data, cleansing, transforming and modeling the data with the goal of finding out some useful information and conclusions. It obtains raw data and converts it into some useful information for decision-making by users. Data mining is the current data analysis technique that focuses on knowledge discovery and modeling for predictive purposes. [1] The precursor to data analysis is the data integration, whereas data analysis is more likely linked to data dissemination and data visualization. The collection of images is increased to petabytes in last few years similar to the text data. The need for analyzing data and proactively describe the actions is pervasive in nearly every vibrant growth industry, government and institutional sector. Image analytics is the logical analysis and automatic algorithmic extraction of information found in image data using digital image processing techniques. [2] To a computer, images are either a raster image or a vector image. To perform analytics on images or videos, the geometric encoding has to be transformed into constructs that depicts the physical features, objects and movement represented by the image or video that can be logically analyzed by a computer. The process of transforming big data (including image data) into higher-level constructs which could be analyzed by the system is organized in progressive steps that add value to the original information in a value chain. The image analytics science has scaled with advances in machine

vision, multi-lingual speech recognition and rules-based decision engines [3].

In this technology dependent world, extraction of data from images contain rich source of information. The data present in the image is efficient for automatic explanation, classification and structuring of images. The extraction process involves detection, localization, tracking, extraction, enhancement, and recognition of the text from a given input image. However the text varies in size, style, orientation, and alignment, as well as low image contrast and complex background which make the problem of automatic text extraction extremely challenging. Image data analysis has been proposed for various applications such as page segmentation, address block location, license plate location, and content-based image/video indexing. [9] The input images can be in gray scale or colour, compressed or un-compressed, and the text in the images may or may not move. The text characters that are embedded in images represent a huge source of information for content-based indexing and retrieval applications. However, these text characters vary in size, grayscale values and complex backgrounds which is difficult to be detected and recognized. Empirical image processing methods, statistical machine learning and modelling methods builds an efficient application system for text detection and recognition of the grayscale values embedded in images. These approaches are studied in two sub-problems: text detection and text recognition. Applying machine learning methods for text detection finds difficulties due to character size, grayscale variations and heavy computation cost. To overcome these problems, two-step localization/verification approach is proposed. The first step aims at quickly assigning candidate text lines, enabling the normalization of characters into a unique size.

In the verification step, a trained support vector machine or multi-layer neural network is applied on background independent features to remove the false alarms. Localizing text from the given input images finds its application in navigation, assistive reading, geocoding, and content-based image retrieval. Tracking involves recognition of the images which we can follow it to extract extra information.

The proposed idea makes use of the image analytics process for the retrieval of the related information from the raw input image uploaded which helps in transformation of the unorganised data into a structured format. The uploaded image relates to the query submitted by the user. The whole process takes place in a cloud environment which also makes use of image analytics technique to sort the images relevantly. Different data sets are maintained at different regions of the globe, thus allowing access to any user located at any part of the world. The related search to the uploaded picture is returned to the user. This process become more helpful for analysing the images related to health care where the images related to their search can be retrieved efficiently.

2. Related Work

Hadoop Image Processing Interface

HIPI is a tool used for image processing and it is programmed to be combined with the Apache Hadoop MapReduce technique with parallel programming framework which facilitates the efficient and high-throughput image processing executed on a cluster. It generates a solution of how to store a huge collection of images on the Hadoop Distributed File System (HDFS) and make them available for a better distributed processing. The base input to a HIPI program is a Hipi Image Bundle (HIB). [4] It is an enormous collection of images portrayed as a single file on the Hadoop File System. The initial processing step of a HIPI program is an extracting step that filters the images in the HIB based user-defined conditions which are achieved using the culler class. The culled Images that are not fully decoded which saves the processing time. The images are then assigned to individual map to maximize data locality which is achieved through the HibInput Format class Individual images are presented as objects to the Mapper which is derived from the HipiImage abstract base class along with an associated HipiImageHeader object. At last the task is reduced by the request of the user are executed in parallel and the appropriate output is aggregated and written to the HDFS.

Distributed Image Processing Using HIPI

The collection of images is increasing in today's world which occupies an enormous memory. Our personal computers cannot analyze such huge accumulation of images. Hence to process these kinds of images, the distributed computing is needed. The framework called HadoopMapReduce Image Processing (Hair) framework provides an ability to implement the image processing technique in the distributed computing. The Apache Hadoop which is an open source is used to implement the HMIPr. The HMIPr is based on the map reducing technique. Various kinds of representation of images are provided by the HMIPr in the internal format of Hadoop input and output tools, for the integration of processing images in to Hadoop data work flow. Common image processing libraries are used for the image format in Hadoop MapReduce Image Processing framework. [1] This provides a high level image processing API for the developers who are not much familiar with the Hadoop. The Sequential functions can be generated using APIs in order to process one or group of related images. The HMIPr framework applies this kind of function to the large quantity of images in parallel. HMIPr also includes MapReduce implementation for popular image processing algorithms that is used for distributed image processing without

the software development process. The HMIPr is found to be more efficient than the existing models.

Advances in Medical Image Processing Image Processing

Medical image processing is an interesting and active field of research. It deals with the generation and reconstruction of the images, pre-processing and enhancement, analysis and quantification, visualization and management of the images. The most repeatedly used steps in the image processing framework are image registration and image segmentation. Image registration aims at sorting the image data from different subjects, modalities, or points of time. Image segmentation aims at delineation and localization of related objects in two-dimensional (2D) and three dimensional (3D) projections of visualization. The present topic of discussion in medical image processing involves the visualization and documentation of the medical image data. Medical image processing is one of the most mature fields of the present time. Real time image processing and composition has become a major focus of research, whereas fields such as quantitative measurement and segmentation are still major concerns of the current ongoing research.

3. Proposed System

The architecture of the proposed work is depicted in the Fig.1. The process involves users from various regions around the world accessing the forum that is created as an App from their own personal devices. The authentication for the users to login into the forum requires certain mandatory fields which have to be filled by the respective users to access their own account in the forum. The user's login is checked and the permission to access is granted. The major purpose of the forum is to perform analytics on the images present in the database and give out the suitable output. The user consists of two options in the forum to access: 1. Search 2. Upload. The search field allows the user to access the database to find in the relevant image along with its details by uploading a raw image for which the user needs information about. The image is uploaded into the search field of the forum where the image is processed using the image processing technique implemented by the cloud image analytics. The raw image that is fed by the user is segmented into parts and the specific areas of detailed and important features are extracted by making use of the feature extraction process. These features are compared with the images present in the database by finding out the relationship of the image with time variables and features. The time stamp of the image is also retrieved. These analytically prepared features are compared with the image existing in the database of various regions. The image matching the input image is extracted from the datacenter and given as the output to the user. The upload field present in the forum helps the users to upload their own files and images to be stored in the data center present at various regions. This proposed work is very helpful for the specialists of doctors around the world to find out the details about some unknown new diseases from all over the world. The forum can be created for the specialists of doctors in the various fields from each country. They can access the services of the portal using their unique details. This work helps in to identify the patient's disease just by uploading the raw image and get the detailed information about the image along with similar images and the treatment given by some specialist given at some part of the world. It helps in gaining knowledge about the various diseases, problems and treatments. This method is efficient for its time, speed and accuracy of the result.

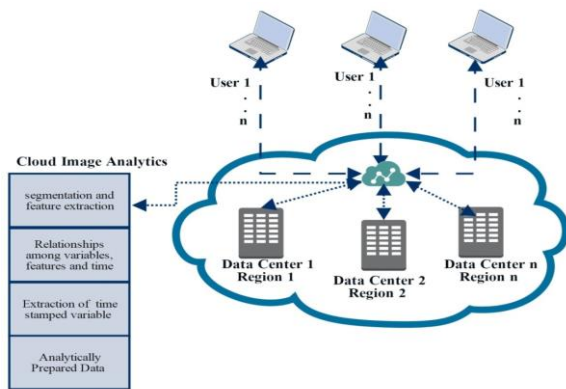


Fig. 1: Architecture of proposed work

Simple Algorithm

Input - Image- I and Patient Report

Output - Relevant data

Step 1: Authorized user login (i1, i2, i3.....in)

Step 2: If Searching: Upload the image into the cloud application.
Else: Go to Step 5

Step 3: Match the input image with existing image using cloud image analytics.

Step 4: Retrieve the relevant report.

Step 5: Upload: Submit the image and patient record into the cloud

The algorithm explains about the cloud based medical image data analytics in healthcare management system. The input consists of the image-I and the Patient Report. The initial step involves the authentication of the user login. The forum can grant access to n authorized users from various regions. The forum provides two features either to search relevant image or to upload data. If the users desire to search for similar image, they are required to upload a raw image about their query. [14][15]The uploaded image is processed using the cloud image analytics. The output of the process consists of the relevant images about the query submitted by the user. If the users desire is to share their case studies with the other users, they can upload their collection of information as images into the forum using upload option that is available which gets stored in the cloud database.[14]

4. Implementation and Result

The experimental setup is implemented using the system having a capacity of 16GB RAM and working on an i5 core processor. The forum is created with Amazon S3 as the cloud platform offered by the Amazon Web Services that is used for the storage of the data and the data analytical process for processing the images is combined with it. This provides a solution for the extraction of the result based on the images submitted by the user that the user has as a query. Initially, a sample of about 200 images is uploaded into the forum for the searching purpose. A raw input image is uploaded for search and the relevant report is obtained that matches the user's input raw image.

The sample output of the image is shown Fig. 2. and this describes about the process of raw image converted as a machine understandable image for the healthcare process.

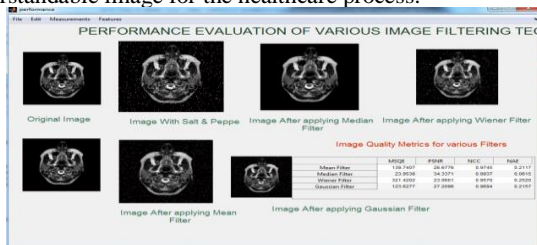


Fig. 2: Brain image with expected output

5. Conclusion

The framework designed in this paper includes the cloud based services combined with the image analytics process which is very advantageous in the field of health care. It provides a great solution for finding the significant result in a valuable time period along with the previous patient records. The use of the forum provides a convenient solution to the unknown or rare diseases. The use of cloud service reduce the memory consumption problem as it is totally based on internet-computing and offers a wide range of memory space and security to the data. The use of image analytics along with the cloud platform provides a consistent result with speed and standard image format. The result produced by the search is very accurate such that if the user looks for a specific report, it is clearly obtained by the processing of the data present in the database of the forum. The proposed work thus produces efficiency in the parameters such as time, speed and accuracy.

References

- [1] Dhinakaran K, Prasanthi K & Kumar P, "Distributed Image Processing Using HIPI Distributed Image Processing Using HIPI", *IJCTA*, Vol.9, No.12, (2016), pp.5583-5589.
- [2] Dhinakaran K, Silviya Nancy J & Saranya R, "Classification and Prediction of Earthquake and Tsunami using Big Data Analytics", *Australian Journal of Basic and Applied Sciences*, Vol.9, No.11, (2015), pp.611-616.
- [3] Dhinakaran K, Silviya Nancy J & Duraimurugan N, "Video Analytics Using HDVFS In Cloud Environment", *ARNP Journal of Engineering and Applied Sciences*, Vol.10, No.13, (2015).
- [4] Nikhita Reddy G & Ugander Reddy GJ, "Study of Cloud Computing in HealthCare Industry", *International Journal of Scientific & Engineering Research*, Vol.4, No.9, (2013).
- [5] Saichandana, B., Rachana Sri, G., Surendar, A., & Suniltej, B. (2017). Controlling of wall lamp using arduino. *International Journal of Pure and Applied Mathematics*, 116(24 Special Issue).
- [6] Bahrami M & Singhal M, "The role of cloud computing architecture in big data", *Information granularity, big data, and computational intelligence*, (2015), pp.275-295.
- [7] Vimalkumar, M. N., Helenprabha, K., & Surendar, A. (2017). Classification of mammographic image abnormalities based on em and LS-SVM techniques. *Research Journal of Biotechnology*, 12(1), 35-40
- [8] Lin CH, Huang LC, Chou SCT, Liu CH, Cheng HF & Chiang, JJ, "Temporal event tracing on big healthcare data analytics", *IEEE International Congress on Big Data (BigData Congress)*, (2014), pp.281-287.
- [9] Kumar M, "Digital Image Processing", *Photogrammetry and Remote Sensing Division Indian Institute of Remote Sensing*.
- [10] Lee JS, "Digital Image Enhancement and Noise Filtering by Use of Local Statistics", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (1980).
- [11] Jung K, Kim KI & Jain AK, "Text information extraction in images and video: a survey", *Pattern recognition*, Vol.37, No.5, (2004), pp.977-997.
- [12] Syal N & Garg NK, "A Study of Text Localization Algorithms for Complex Images", *International Journal of Innovative Research in Computer and Communication Engineering*, Vol.2, No.4, (2014).
- [13] Wadhawan K & Gajendran E, "Automated Recognition of Text in Images: A Survey", *International Journal of Computer Applications*, Vol.127, No.15, (2015).
- [14] G Ainabekova, Z Bayanbayeva, B Joldasbekova, A Zhaksylykov (2018). The author in esthetic activity and the functional text (on the basis of V. Mikhaylov's narrative ("The chronicle of the great jute"). *Opción*, Año 33. 63-80.
- [15] D, Ibrayeva, Z Salkhanova, B Joldasbekova, Zh Bayanbayeva (2018). The specifics of the art autobiography genre. *Opción*, Año 33. 126-151.