Journal of Arabic and Human Sciences
Qassim University, Vol. 14.No.1. PP 1-15. (August2020/ Muharram1442)

Remote Sensing Big Data: Challenges, Opportunities, Management and Application

Md. Enamul Huq
State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing
Wuhan University
Wuhan, China
enamul_huq@whu.edu.cn

Prof. Ahmed Abdullah Al Dughairi
Geography Department
Qassim University
Buraidah, Saudi Arabia
ahmadam320@gmail.com

Md. Mahabubur Rahman
Department of Computer Science and Engineering
Bangladesh University of Business and Technology
Dhaka, Bangladesh
lizububtcse@gmail.com

Abstract: Each single day many earth monitoring spaceborne and aerial sensors from various countries are collecting a huge quantity of remotely sensed data. This large amount of data is regarded as remote sensing (RS) big data or big RS data. Due to these massive amounts of data the world has entered in the age of big (RS) data. Big RS data have different physical, spatial and temporal features such as multisource, multiscale, high dimensional, dynamic state, isomer, and nonlinear properties. Moreover, RS big data are becoming as an economic resource and a modern key asset in several applications. The RS big data are using for various applications in resources as well as environment (both natural and built) for local and global observation, for example in agriculture, natural disaster monitoring, worldwide climate change, and urban planning. However, this paper aims to find the big RS data as well as data-intensive problems, counting RS big data challenges, analysis, methods, techniques, processing, managing, and effective utilization of big RS data. Furthermore, it seeks the methods and applications of RS big data. In the current study, the challenges and opportunities of RS big data applications were analyzed specifically. In order to explain the big RS data aspects, a case study discussing the application of RS big data was demonstrated. In this test case, RS big data were used to identify flood zones in Wadi Al Rimah in Kingdom of Saudi Arabia applying a big archive of RS data. In this study RS big data were applied to exhibit the substantial challenges along opportunities carried by the application of RS big data.

Keywords: *big data, remote sensing, challenges, methods, application*

Introduction :

As a moving data producers, human beings produce data each day. Remote sensing (RS) devices are using extensively to monitor our earth to make sure our lives more simple. The planet as well as moving data producers are the key actors to generate RS big data. According to IBM, approximately 2.5 quintillion of bytes data are generating every day over the world. A large quantity of RS data is freely accessible in NASA archives. It includes 7.5 petabytes (PB) data along 7000 individual data sets with 1.5 million users [1, 2]. The earth observatory data, generated from spacecraft nearly 2(1.7) GB, which is collected from a single satellite and raised a number of terabytes data in a day. The entire records of laboratory data of whole earth might exceed more than a exabyte [3]. However, till now there are some common problems remain to gain acumen into the RS big data. But the RS big data resource might lead a big opportunity. Big data is mostly characterized by the three attributes (i.e., volume, variety, as well as velocity). It is defined as the three “V” features. Big RS data can describe with its own features (denoted as the 3Vs). In terms of variety, it is obvious that big RS data comprise multi-source (e.g., optical, laser, radar, etc.), multi-temporal (hours, day) as well as multi-resolution (several spatial resolutions). These possibilities have been employed to manage various real-world problems with the RS big data, for example archaeology, crop estimation and yield predicting, food security, public health, land utilization planning, sustainable city planning, as well as management, forest monitoring, war, conflict studies [1].

A. Understanding RS Big Data

Usually remote sensing refers the technology to measure the features of an object or surface form a distance. The modern advances of RS with computer techniques raise an explosive development of RS data [4]. A Huge amount of data is generating by the satellite sensors. Therefore, storing and processing of RS data is a big challenge owing to its diverse variety and large volume. The big challenges also emerge to collect and manage RS big data [5]. The RS data is collected from satellite data center and intensely it increases into several terabytes in a day. RS data is documented as the big data in some particular senses. The RS big data not simply denotes to the large volume and diverse velocity of data which is challenging to storing and computing, but also represents wide variety as well as complexity [4]. To well understanding of big data, different perspectives need to be combined, such as data acquisition, data methodologies, as well as data applications [1].

Big Data, Big Challenges

The big data challenges in RS includes large volume of data, data gaining, storing, managing, and analyzing [6]. RS big data challenges are mentioned in the following according to trinity to understand of big data in RS.

B. Identifying Data

Identifying data is a significant characteristic of big data. Based on it we can identify the applications, utilities or designing of appropriate data methods for solving the actual problems (e.g., RS problem). The big challenges appear at each and every step when the big RS data is obtained and organized [7]. For example, RS data is received from the satellites, drone, airplanes, or any other sensing instruments at the same time, the other types of data also stored from the cyberspace. However, the main distinction between the conventional data and big RS data is to find the exact data sets as well as to integrate them to resolve the challenging and unique problem. Thus, data access, collaboration of domains and data organization should conduct in an effective way. It is one of the critical challenges in RS big data applications.

C. RS Big Data Methodologies

A RS big data methods should design scientifically to address the big data difficulties from various RS domains. This type of methodology could be applied to design a new data method for big RS data preparation, arrangement, mining, modeling, fusion, presentation, visualization and interpretation [8]. Moreover, data processing as well as analysis refer the multistep pipeline with data-driven methods that might be remarkably separate from the standpoint of particular applications as well as domains. In order to heterogeneity along with diverse dimensionality of RS big data, stakeholders often face significant computational and statistical hurdles associated with processing scalability, noise accumulations, synthetic correlations, secondary endogeneity, and measurement inaccuracy. To solve these challenges, it is required novel computational as well as statistical methods to analyze and process RS big data.

D. RS Big Data Applications

The applications of RS big data aim to find out the exact data to resolve the problem, because it is difficult to address or generally impossible to manipulate with traditional RS data. Then, the another challenge is collecting, organizing, and utilizing the RS big data for solving the real RS problems [9]. For identifying the exact RS data, it is needed to link closely with the original facet RS big data. After obtaining the exact data, for instance RS data, advanced data methodologies could be established for discovering, realizing, and demonstrating the significant of RS big data applications.

E. RS Big Data Computing

It is a great challenge to design a high-performance system to compute RS big data. Though cloud computing system has been developed for realizing the high level aggregated performance for RS applications, but still there are some challenges remain concerning the progressive integrated cloud computing concept in RS research [2, 6]. However, the energy consumption is an another issues for RS big data. Solving these issues are the important step to incorporate computing techniques in RS big data applications. Previous studies [3, 4, 8] on big data in RS mostly emphases on the volume issues to compute big data as well as evaluates it as the data rigorous computing problem. Although the data processing systems can do so, nevertheless it is a challenge to gather the dispersed data as well as to send those data in the exact computing node.

F. Proper RS Data Identification

Big RS data generally encompass in-domain as well as out-domain data. Formerly, those different types of data had combined rarely to accomplish RS applications/tasks. Hence, the data had no value without fulfil a particular task. However, the crucial difference in between the conventional RS data and big RS data is that, how to determine and integrate the various types of data to solve the real-world problems. The other important challenge of RS big data is to classify and harness the suitable data to resolve the imminent problem. In RS, there are different sorts of data, those include optical (multispectral, hyperspectral), radar (synthetic aperture radar), or laser (light detection and ranging) which are collected by airplane, satellite and ground sensors [2, 7, 10]. In addition, some image data are taken from different social networks which could assist in RS big data interpretation.

G. RS Big Data Organization

One more challenging issue of RS big data is to organize the data to utilize substantially. In the stage of RS big data utilization, RS big data organization incorporates data formulation, data management techniques, and data methods [3]. Therefore, data obtaining, data storing, data computing environment, as well as models building to get understanding of RS big data must be wisely fourmulated in data organization phase.

H. RS Big Data Visualization and Interpretation

Visualization allows the big data users to achieve well insights into RS big data and it is imperative to comprehend and analyze the RS big data to reveal data details and relevance to gain the recent aims/objectives [5]. Thus, the visualization should consider initially, along with different prior tasks, for example, data gaining and preprocessing. It needs an innovative visualization technique along with previous interdisciplinary domain information through close collaboration with the domain specialists who have pretend the task for addressing the real problems.

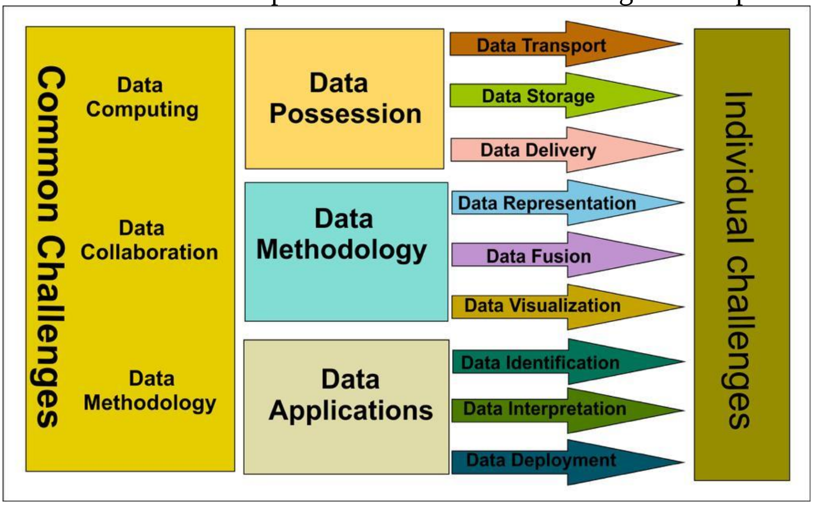

Fig. 1. Challenges of RS big data [1].

II. Big Opportunities

In spite of the above-mentioned big challenges, the possible values of RS big data is remarkable. In fact, the RS techniques have been effectively applied for various applications, for instance in agriculture (crop monitoring, grassland monitoring), hazard monitoring (flood monitoring), oceanography (ship identification), urban development, urban sprawl monitoring, human settlements (urban as well as rural), food safety monitoring, water quality observation, energy monitoring, ecosystem surveillance, global warming inspection, global change, global forest resources assessment, ancient site detection (archaeology), large-scale debris movement investigation disaster management (flood, drought, landslide, earthquake, etc.), and the observation of deep sea oil spills and so on [1, 10-12]. Combination of human activities and data from RS techniques have become a powerful tool for improving the production as well as operation for the welfare of human being. Thus, the big RS data offers the capacity for accomplishing the goals, aims and objectives which would hard or implausible to achieve with traditional techniques. Specifically, food safety is the main influential factor for the smart agricultural system and only the remote sensing observation from the earth monitoring satellites (Landsat, LiDAR, SAR, MODIS) can deliver constant, reliable, repeated, as well as good quality data to characterize and mapping features to estimate the global cropland as well as food security investigation. Similarly, for the other purposes (e.g., water quality observation, energy monitoring, ecosystem surveillance, global warming inspection, global change, disaster management) RS big data can be employed for smooth socioeconomic and environmental management.

III. RS Big Data Management

To solve the RS big data challenges, pipsCloud high-performance cloud computing [5] might be used. It can process RS big data on the cloud platform. It offers much efficient and user-friendly method for high-performance RS big data processing, and optimize quality of service (QoS) to conduct data-intensive issues. As shown in Fig. 2, pips Cloud holds multilevel system architecture. It includes, physical properties, cloud framework, virtual cluster environment for remote sensing (VE-RS), data management, virtual processing system for remote sensing (VSRS), and a cloud portal from bottom to top, where the cloud framework handle the physical assets to deliver ‘Infrastructure as a Service’ by the OpenStack. On the basis, of cloud framework, the VE-RS provides the virtual high-performance cloud (HPC) cluster environment as a service and the VS-RS offers a cloud empowered virtual RS big data processing system for online and large-scale RS data processing. The management, indexing, and sharing of RS big data are also handled as ‘Data as a Service’ Fig. 2.

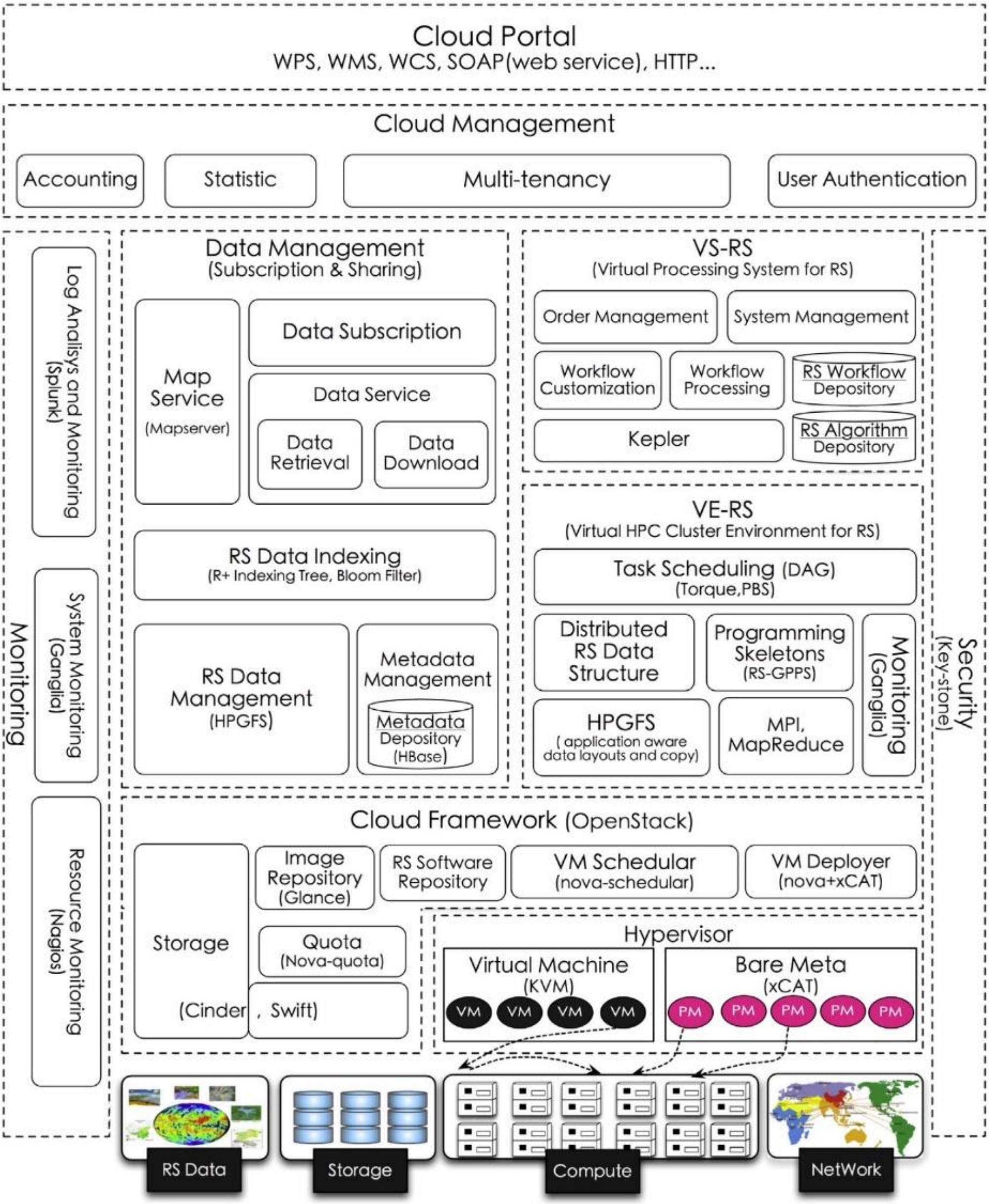

Fig. 2. System architecture for pipsCloud [5].

IV. Case Study

In the present study, Wadi Al Rimah in Saudi Arabia was selected as the case study area. Wadi systems extend in different parts of Saudi Arabia; for example, Wadi Al Rimah is one of the longest Wadi systems in Saudi Arabia (Fig. 3) and extends for 970 km from Harrat Khaybar in the west to the east of Saudi Arabia toward the Arabian Gulf [13]. The uppermost Wadi channel is a narrow gorge that cuts into Precambrian metamorphic and igneous rocks, namely schists, phyllites and granites [14], as well as large basalt flows from the Tertiary and Quaternary ages [15, 16]. The mid-drainage area is flat, with a low gradient (0.24° per km on average) and gentle slopes [17]. Approximately 4 km to the west of Burydah City, the wadi runs into the surface of a cuesta, which consists mainly of Cambrian–Devonian sandstones. According to Hotzl et al. (1969) [14], Wadi Al Rimah is a large, old river system crossing the Burydah area from west to east, with a channel of varying width at the eastern margin of the study area which cuts across a sabkhah until it becomes blocked by the sand dunes of Nafud Al Thuwayrat. Beyond Nafud Al Thuwayrat, Wadi Al Rimah takes on another name, ‘Wadi Al Ajradi’; it then crosses the Al Taysiyah plateau, which is composed of limestone, dolomite and clays. When the Wadi crosses the western edge of the Sulb Plateau it takes a turn towards Wadi Al Batin and continues on towards the eastern Saudi border with Iraq.

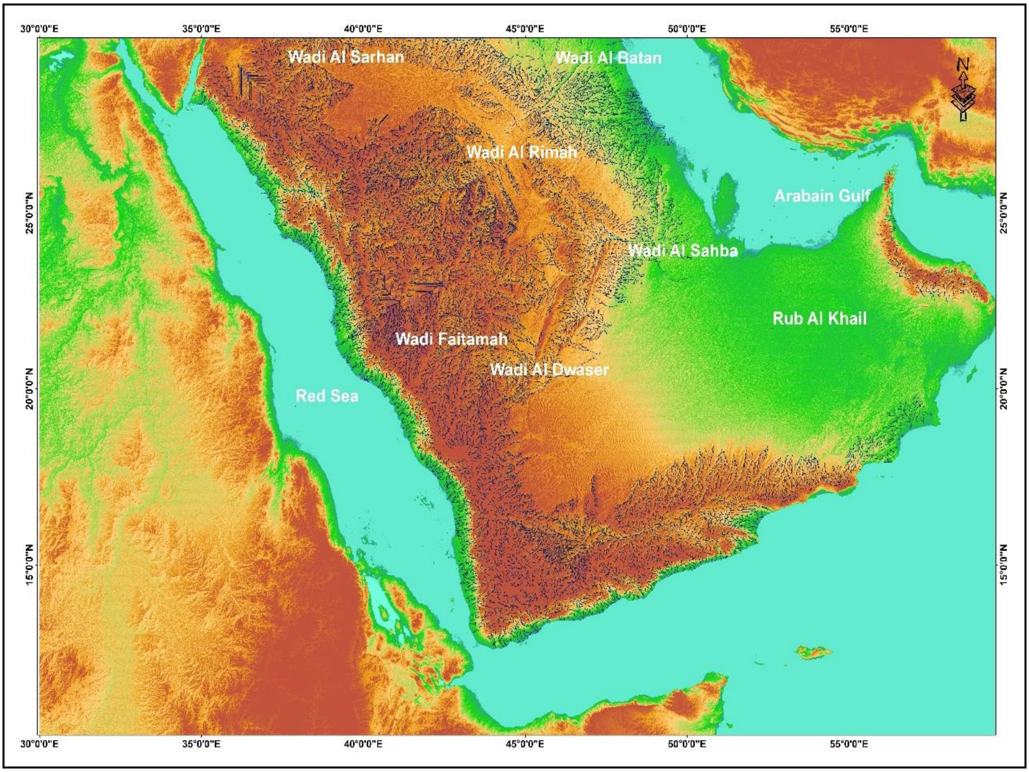

Fig. 1. A general map showing the main Wadis in Saudi Arabia, box represents the study area. ***Source: Saudi Atlas 19***

The case study has been included based on reconstructed channel and flood of Wadi Al Riham that indicating the efficiency of RS big data application.

In this case study, a novel data processing method and a powerful computing

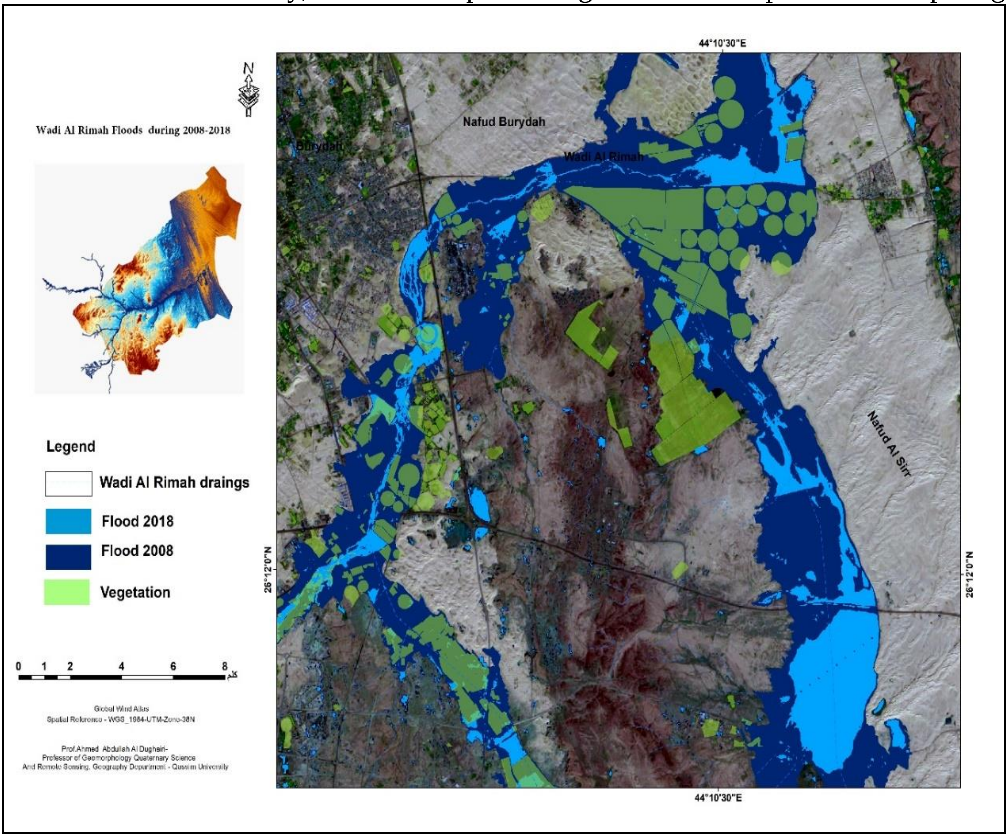

Fig. 4. Flood zone detection on multitemporal and multisource spaceborne remote sensing images using big data in remote sensing for Wadi Al Riham of KSA.

architecture was employed to address the automatic flood zone identification. For this study multispectral as well as hyperspectral RS data and other domains were used. However, in the traditional remote sensing techniques, categorized samples are taken according to the field survey and image photointerpretation. *In situ* field survey provides very high accuracy of classification but it is expensive as well as time consuming. In contrast, image photointerpretation technique is quick and low cost but cannot assure the accuracy classification quality. Till now it is a complex process to mark flood zone with the fusion resolution of RS data, providing by air/spaceborne devices owing to flood water float and circulation. Hence, to identify the flood zone is a big challenge. In this exploration, initially suitable data format of big RS data were identified and to tackle the classification challenge an unique data methodology was applied. Particularly, in this study the large flood events happened in 2008 and 2018 were selected. The optical RS data containing multitemporal as well as multisource images were used. These various types of big data are employed to enhance flood zone recognition accurately. Classification is a lengthy process as well as the exterior data unable to incorporate all the pixels in RS images. Consequently, it is imperative to choose a concise number of useful samples logically to classify that makes sure the accuracy for classification task. Afterwards eliminating data that is severely tainted by clouds. Therefore, some popular classifiers (support vector machine (SVM), backpropagation neural networks, and k-nearest neighbor classifier) were applied. The SVM classified map of flood zones of 2008 and 2018 in Wadi Al Riham is presented in Fig. 4.

V. Conclusions

This study proposed that RS big data have a great potential to serve disaggregated as well as timely a real method to monitor economic, and environmental changes globally. The growing accessibility of high quality global observing of the glob from space and RS big data put forward new possibilities to quantify and identify the economic, environmental problems simultaneously solutions, where the conventional data sources are habitually scarce. The flood related RS big data provide different information such as water flow, level to produce and propose real time flood map. Agricultural studies face the challenge of uncertainties from the variations of weather conditions and management strategies. RS big data are the valuable resource for precision agriculture to potentially make robust distributions of agricultural variables, such as yield and other biotic and abiotic indicators of crops, to tackle the uncertainties from experiments and analysis from different sites and farms. It is also one of the backbone technologies for precision agriculture, within-field variability for local regional and farm scale instead of uniform management as in traditional agriculture. It is also an important information source for monitoring land use and land cover, water mapping for urban environmental management. The high-resolution night lights remote sensing big data can be used to estimate electricity access in remote areas to guide the planning and monitoring of grid extension programs and locate existing electricity infrastructure.

References

[1] M. Chi, A. Plaza, J.A. Benediktsson, Z. Sun, J. Shen, Y. Zhu, Big data for remote sensing: Challenges and opportunities, Proceedings of the IEEE, 104 (2016) 2207-2219.

[2] Z. Shao, L. Wang, Z. Wang, J. Deng, Remote Sensing Image Super-Resolution Using Sparse Representation and Coupled Sparse Autoencoder, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12 (2019) 2663-2674.

[3] D. Rajyalakshmi, K.K. Raju, G.S. Varma, Taxonomy of Satellite Image and Validation Using Statistical Inference, 2016 IEEE 6th International Conference on Advanced Computing (IACC), IEEE, 2016, pp. 352-361.

[4] Y. Ma, H. Wu, L. Wang, B. Huang, R. Ranjan, A. Zomaya, W. Jie, Remote sensing big data computing: Challenges and opportunities, Future Generation Computer Systems, 51 (2015) 47-60.

[5] L. Wang, Y. Ma, J. Yan, V. Chang, A.Y. Zomaya, pipsCloud: High performance cloud computing for remote sensing big data management and processing, Future Generation Computer Systems, 78 (2018) 353-368.

[6] Z. Shao, J. Cai, P. Fu, L. Hu, T. Liu, Deep learning-based fusion of Landsat-8 and Sentinel-2 images for a harmonized surface reflectance product, Remote Sensing of Environment, 235 (2019) 111425.

[7] Z. Shao, Y. Pan, C. Diao, J. Cai, Cloud Detection in Remote Sensing Images Based on Multiscale Features-Convolutional Neural Network, IEEE Transactions on Geoscience and Remote Sensing, 57 (2019) 4062-4076.

[8] N. Levin, S. Ali, D. Crandall, S. Kark, World Heritage in danger: Big data and remote sensing can help protect sites in conflict zones, Global Environmental Change, 55 (2019) 97-104.

[9] Z. Shao, J. Cai, Remote sensing image fusion with deep convolutional neural network, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11 (2018) 1656-1669.

[10] Y. Song, B. Huang, Q. He, B. Chen, J. Wei, R. Mahmood, Dynamic assessment of PM2.5 exposure and health risk using remote sensing and geo-spatial big data, Environmental Pollution, 253 (2019) 288-296.

[11] Y. Huang, Z.-x. Chen, T. Yu, X.-z. Huang, X.-f. Gu, Agricultural remote sensing big data: Management and applications, Journal of Integrative Agriculture, 17 (2018) 1915-1931.

[12] Z. Shao, J. Deng, L. Wang, Y. Fan, N. Sumari, Q. Cheng, Fuzzy autoencode based cloud detection for remote sensing imagery, Remote Sensing, 9 (2017) 311.

A. Al Welaie, Geomorphology and Geomorphology of Saudi Arabia. Al]13[Obeikan Library. Riyadh, Saudi Arabia, (1996).

[14] A. Hotzl, W. David, W. Graham, B. P. Colin, H. Stephen, Monsoon Climate and Arabian Sea coastal Upwelling Recorded in Massive Corals from Southern Oman. Society for sedimentary Geology, 11 (1969) 347-361.

[15] G. Brown, D. Schmidt, A. Huffman, Geology of the Arabian Peninsula: Shield area of western Saudi Arabia. U. S. Geological Survey Professional Paper, (1989) 560-A.

[16] B. Bramkamp, D. Fleitmann, M. Mudelsee, U. Neff, A. Matter, A. Mangini, A780-year annually resolved record of Indian Ocean monsoon precipitation from a speleothem from south Oman. Journal of Geophysical Research, (2002) 107.

[17] A. Al Dughairi, Geomorphologic Al watah in Saudi Arabia. Masters Thesis, King Saud University, Riyadh Saudi Arabia (2003).

بيانات الاستشعار عن بعد الكبرى : التحديات والفرص والإدارة والتطبيق

د. انعام حق

مختبر الهندسة للبحث والاستشعار عن بعد جامعة أوهان

أ د أحمد عبدالله الدغيري

قسم الجغرافيا -جامعة القصيم كلية اللغة العربية والدراسات الاجتماعية

د. محمد محبوب الرحمن

قسم العلوم والهندسة الحاسبات جامعة بنغلاديش للأعمال والتكنولوجيا

ملخص البحث: تقوم العديد من أجهزة الاستشعار المحمولة على المنصات الجوية بمراقبة الأرض بشكل يومي ومن مسافات بعيدة في كافة اصقاع العالم، هذه البيانات الضخمة التي توفرها تقنيات الاستشعار عن بعد مكنت العالم من دخول ما يعرف بالاستشعار عن بعد الضخم ، تميزت هذه البينات الضخمة بخصائص فيزيائية ومكانية وزمنية مختلفة بالإضافة إلى تنوع وتعدد نطاقاتها ومصادرها ، وكذلك تعدد أبعادها وتوافرها بصور ديناميكية ، علاوة على ذلك أصبحت البيانات الكبيرة في منظومة الاستشعار عن بعد مورد اقتصادي وأصل رئيس حديث في العديد من التطبيقات ، و أضحت هذه البيانات الضخمة تستخدم في الكثير من تطبيقات الموارد المختلفة وكذلك البيئة (الطبيعية والمبنية على حد سواء) و كذلك للمراقبة المحلية والعالمية ، على سبيل المثال في الزراعة ومراقبة الكوارث الطبيعية وتغير المناخ العالمي والتخطيط الحضري. وعلى هذا ، فإن هذه الورقة تهدف إلى ثبت تحدي بيانات الاستشعار عن بعد الضخمة في مجال مشكلات كثافة البيانات ، تحليلات البيانات الكبيرة كالتحليل ، والأساليب ، والتقنيات ، والمعالجة ، والإدارة ، والاستخدام الفعال

لبيانات الكبيرة. علاوة على ذلك، فالبحث يحاول تتبع أساليب وتطبيقات البيانات الكبيرة للنقاط المرجعية.

في الدراسة الحالية، تم تحليل التحديات والفرص المتعلقة بتطبيقات البيانات الضخمة في قطاع النقل على وجه التحديد، من أجل تبيان جوانب بيانات الاستشعار الكبيرة في هذا المجال، وتم تقديم دراسة حالة تناقش تطبيق بيانات الاستشعار الضخمة في تحديد نطاقات فيضان وادي الرمة في المملكة العربية السعودية واسهام ارشف بيانات الاستشعار الضخمة في استعراض الفرص والتحديات التي تقدمها هذا البيانات التطبيقية الضخمة

الكلمات المفتاحية: البيانات الضخمة، الاستشعار عن بعد ، التحديات ، الأساليب ، التطبيقات.