

Teesside University MIDDLESBROUGH TS1 3BA

School of Computing – MSc Data Science

SYSTEMATIC LITERATURE REVIEW AND REPORT

Research Methods, CIS 4011-N

Big Data Analytics for Medical Imaging in Healthcare Sector

Author: ATHIRA REGHUNATH

Student ID: W9561613

Module Leader: Dr. Zia Ush Shamszaman

Submission Date: 11/05/2022

Word Count: 2733

Table of Contents

INTRODUCTION	
RESEARCH QUESTION	3
KEY CONCEPTS	
DATABASE RESEARCH	3
Database 1: ACM Library	3
Database 2: IEEE Xplore	
Database 3: Discovery	
LITERATURE EVALUATION	6
INCLUSION AND EXCLUSION PROCESS	
LIMITATIONS	10
CONCLUSION	10
REFERENCES	

INTRODUCTION

In recent years, one of the hottest topics of scientific discussion has been big data. An industry analyst sought to define big data in the early 2000s using the three Vs: volume, velocity, and variability. With new technologies like Hadoop, it is now possible to store and use massive amounts of data that arrive at an unparalleled rate. This data is highly varied because it might be in a various format, including text documents, voice or video recordings, and financial transactions. Science, sports, advertising, health care, genomic sequence data, and medical imaging have all proven to benefit from big data analytics. This research provides a high-level overview of big data analytics in medical imaging. Medical imaging data from X-rays, CT scans, and MRIs, among other sources, can be handled effectively and efficiently using Big Data analytics platforms to acquire superior diagnostic insights.

RESEARCH QUESTION

How Big Data Analytics works for Medical Imaging in Healthcare Sector?

KEY CONCEPTS

Big Data Analytics Big Data Analytics in Healthcare Big Data Analytics in Medical Imaging Medical Image Processing

DATABASE RESEARCH

For my research, I used these three databases to find results and information. To get more accurate results and progress in learning, I found some more databases to add to the literature review.

- 1) ACM Library
- 2) IEEE Xplore
- 3) Discovery

Database 1: ACM Library

Question	Keywords	Filters	Results	Notes
What Published on Big Data	"Big Data	None	559,912	There are too many
Analytics?	Analytics"			search results and it will
				take time to find the
				best results. As a result,
				I need to narrow my
				search to the last 7
				years.

What Published on Big Data Analytics in the previous 7 years?	"Big Data Analytics"	Year: 2015- 2022	211,564	It lowers the number of results, but still it's too much to read and find the answer. So, need to apply the keyword 'Medical Imaging'
What Published on 'Big Data Analytics' and 'Medical Imaging'?	"Big Data Analytics" and "Medical Imaging"	Year: 2015- 2022	86,488	There are too many results, including all the relative data. So, I need to add the keyword "Big Data Analytics" and title: "Medical Imaging"
What Published on 'Big Data Analytics' and the Document Title as 'Medical Imaging'?	"Big Data Analytics" and Title: as "Medical Imaging"	Year: 2015- 2022	5,412	Add keywords Title: as "Big Data Analytics" AND Title: as "Medical Imaging"
What Published on the Document Title as 'Big Data Analytics' and the Document Title as 'Medical Imaging'?	Title: as "Big Data Analytics" and Title: as "Medical Imaging"	Year: 2015- 2022	533	Still need to lower the results by adding the filter Year: 2020-2022
What Published on the Document Title as 'Big Data Analytics' and the Document Title as 'Medical Imaging' in the past 2 years?	Title: as "Big Data Analytics" and Title: as "Medical Imaging"	Year: 2020- 2022	232	Need to minimise the number of results again, so we filter them again using the publication type of journal only.
What are the Journals Published on the Document Title as 'Big Data Analytics' and the Document Title as 'Medical Imaging' in the past 2 years?	Title: as "Big Data Analytics" and Title: as "Medical Imaging"	Year: 2020- 2022 Publication: Journals	38	Got 38 journals but the required results are published by the IEEE Computer Society Press. So, we have to filter it also.
How Big Data Analytics works for Medical Imaging in Healthcare sector?	Title as "Big Data Analytics" and Title as "Medical Imaging"	Year: 2020- 2022 Publication: Journals Publisher: IEEE Computer Society Press	9	Now got the 9 journals for my research from this particular database.

Database 2: IEEE Xplore

Question	Keyword	Filter	Results	Notes
What Published on Big Data Analytics?	"Big Data Analytics"	None	8,278	Too many irrelevant results are showing. Need to filter it with past 7 years for better results.
What Published on Big Data Analytics in the previous 7 years?	"Big Data Analytics"	Year: 2015- 2022	7535	Results are still more than expected and change the keyword as 'Big Data Analytics for Medical Imaging'
What Published on Big Data Analytics for Medical Imaging in the previous 7 years?	"Big Data Analytics for Medical Imaging"	Year: 2015- 2022	227	Results are comparatively less but need to apply the filter for accurate results Year: 2020-2022
What Published on Big Data Analytics for Medical Imaging in the previous 2 years?	"Big data Analytics for Medical Imaging"	Year: 2020- 2022	67	Select only conferences for better filtering
What all are the conference Paper published on Big data Analytics for Medical Imaging?	"Big data Analytics for Medical Imaging"	Year: 2020- 2022 Only Conferences	50	Add Keyword Big Data Analytics as Document Title and Medical Imaging as All Metadata for perfect results
What all are the conference Paper (with "Document Title": big data analytics AND "All Metadata": medical imaging) published on Big data Analytics for Medical Imaging?	"Big data Analytics for Medical Imaging" ("Document Title": big data analytics) AND ("All Metadata": medical imaging)	Year: 2020- 2022 Only Conferences	38	Then the result is showing the data from 2014-2021 we have to filter again it from 2018-2021
How Big Data Analytics works for Medical Imaging in Healthcare sector?	("Document Title": big data analytics) AND ("All Metadata": medical imaging)	Year: 2018- 2021 Only Conferences	11	Got 11 Conference Papers for the research

Database 3: Discovery

Question	Keyword	Filter	Results	Notes
What Published on Big Data Analytics?	"Big Data Analytics"	None	41,298	So many articles from different industries related to Big Data Analytics. Need to filter it with past 7 years.
What Published on Big Data Analytics in the past 7 years?	"Big Data Analytics"	Year: 2015-2022	32,331	Need to add a proper keyword for filter as 'Big Data Analytics for Medical Imaging'.
What Published on Big Data Analytics for Medical Imaging in the past 7 years?	"Big Data Analytics for Medical Imaging"	Year: 2015-2022	527	Need only past 5 years result. Add a filter for year
What Published on Big Data Analytics for Medical Imaging in the past 5 years?	"Big Data Analytics for Medical Imaging"	Past 5 Years	393	For better research only need to take Journal Articles and Conference Proceedings.
What are the Journal Articles & Conference Proceedings Published on Big Data Analytics for Medical Imaging in the past 5 years?	"Big Data Analytics for Medical Imaging"	Past 5 Years Content Type: Journal Articles & Conference Proceedings	372	Add a keyword 'How Big Data Analytics works for Medical Imaging in Healthcare sector'
What Published on How Big Data Analytics works for Medical Imaging in Healthcare sector ?	Analytics works for Medical Imaging in	Past 5 Years Content Type: Journal Articles & Conference Proceedings	10	Got 10 articles for my research

LITERATURE EVALUATION

When evaluating the whole research, so many filtering and keyword applications are done for better results. When we search without any filter, we get a huge number of results that contain irrelevant data from different sectors. For my research, I selected four databases primarily, but the fourth one was not showing the full research paper, so I removed that one and did research with ACM, IEEE Explorer, and Discovery. Among three of these, I got more relevant articles from IEEE. And it feels like conference papers are more than just articles. It is a little bit difficult to sort out the focused topics and make research among them. But the process is systematic and, through this, our research becomes much easier to complete.

From this research, we can evaluate that disease surveillance can benefit from the use of big data analytics. Medical imaging data sets can be evaluated quickly to produce more precise disease diagnoses and the necessary prevention and treatment strategies, leading in better crucial decision-making. Medical photos come in a variety of sizes, resolutions, and modalities. To handle this massive volume of image data, a specialized analytics platform is necessary.

These images reveal the inner workings of the body, which is useful for detecting malignancies, diabetic retinopathy, and artery stenosis, among other things. The use of Big Data Analytics platforms to store, retrieve, and analyse medical picture data has resulted in considerably faster diagnosis and treatment planning. Parallel programming and cloud computing have also aided in overcoming the difficulties of processing such large amounts of data. Images are retrieved and patterns are detected using the extracted data during medical image processing. Hadoop, MapReduce, YARN, Spark, Hive, and other tools and frameworks are utilised to solve the problem. Machine learning and deep learning techniques are commonly used to accomplish the required analytics. Several strategies are employed, including genetic algorithms and association rules.

INCLUSION AND EXCLUSION PROCESS

Sl No	Research Paper Title	Included/ Excluded	Reasons
1	"Data-Enabled Digestive Medicine: A New Big Data Analytics Platform (Lu Yan, Weihong Huang, Liming Wang, Song Feng, Yonghong Peng, and Jie Peng, May/June 2021)"	Excluded	Related to Big Data Analytics but not to Medical Imaging.
2	"A Latent Gaussian Copula Model for Mixed Data Analysis in Brain Imaging Genetics (Aiying Zhang, Jian Fang, Wenxing Hu, Vince D. Calhoun, Yu-Ping Wang, July-Aug. 2021)"	Included	This paper gives more information about Big Data Analytics in Brain Imaging
3	"Group Sparse Joint Non-Negative Matrix Factorization on Orthogonal Subspace for Multi- Modal Imaging Genetics Data Analysis (Peng Peng, Yipu Zhang, Yongfeng Ju, Kaiming Wang, Gang Li, Vince D. Calhoun, and Yu-Ping Wang, JanFeb. 2022)"	Excluded	Not fully related to my topic
4	"A Data-Driven Approach to Predict and Classify Epileptic Seizures from Brain-Wide Calcium Imaging Video Data (Jingyi Zheng,Fushing Hsieh,Linqiang Ge, NovDec. 2020)"	Included	Related to the selected topic
5	"Integration of Imaging (epi)Genomics Data for the Study of Schizophrenia Using Group Sparse Joint Nonnegative Matrix Factorization (Min Wang,Ting-Zhu Huang,Jian Fang,Vince D. Calhoun,Yu-Ping Wang, SeptOct. 2020)"	Excluded	Not related to Big Data Analytics but related to Healthcare sector
6	"Identify Consistent Cross- Modality Imaging Genetic Patterns via Discriminant Sparse Canonical Correlation Analysis (Meiling Wang, Wei Shao, Xiaoke Hao, Li Shen, Daoqiang Zhang, July-Aug. 2021)"	Excluded	Not gives any information related to Big Data Analytics

7	"Multi-Task Sparse Canonical Correlation Analysis with Application to Multi-Modal Brain Imaging Genetics (Lei Du, Kefei Liu, Xiaohui Yao, Shannon L. Risacher, Junwei Han, Andrew J. Saykin, Lei Guo, Li Shen, JanFeb. 2021)" "A Novel Encoding and Decoding Calibration	Included Excluded	Some areas give some information about medical imaging Paper is not related to big data
	Guiding Pathway for Pathological Image Analysis (Hansheng Li,JianPing Li,Yuxin Kang,Chunbao Wang,Feihong Liu,Wenli Hui,Qirong Bo,Lei Cui,Jun Feng,Lin Yang, JanFeb. 2022)"		analytics or medical imaging
9	"Celiac Disease Detection From Video capsule Endoscopy Images Using Strip Principal Component Analysis (Bing Nan Li,Xinle Wang,Rong Wang,Teng Zhou,Rongke Gao,Edward J. Ciaccio,Peter H. Green, July-Aug. 2021)"	Excluded	Not gives any information related to Big Data Analytics
10	"Using Big Data Analytics to Detect Fraud in Healthcare Provision (Spiros V. Georgakopoulos; Parisis Gallos; Vassilis P. Plagianakos, 2020)"	Excluded	Content of this paper is related to healthcare sector but not about the Medical Imaging
11	"Segmentation and Classification of Brain MR Images Using Big Data Analytics (Mahboob Alam;Mohd Amjad, 2018)"	Included	This paper gives information about the Brain MRI using Big Data Analytics
12	"Big Data Analytics of Inpatients Flow with Diabetes Mellitus type 1: Revealing new awareness with Advanced Visualization of Medical Information System Data (Olga Kolesnichenko; Elena Marochkina;Roman Komarov;Lev Mazelis; 2019)"	Excluded	It's only talks about diabetic mellitus
13	"Health Analytics on Big COVID-19 Data (Nguyen Duy Thong Tran; Carson K. Leung; Daryl L.X. Fung; Thanh Huy Daniel Mai, 2021)"	Excluded	Not related to the subject. It's only talking about the covid 19 data
14	"An Intelligent Visual Big Data Analytics Framework for Supporting Interactive Exploration and Visualization of Big OLAP Cubes (Carlos Ordonez;Zhibo Chen;Alfredo Cuzzocrea;Javier Garcia-Garcia, 2020)"	Excluded	This paper is not related to healthcare industry.
15	"The Use of Big Data Analytics in Medical Images: a Survey (Mohammed Shatnawi;Muneer Bani Yassein;Ibtihal Jalabneh, 2020)"	Included	It's very much related to the topic Big Data Analytics for Medical Imaging
16	"Big Data Analytics in Healthcare (M. Ambigavathi; D. Sridharan, 2018)"	Included	Medical Imaging is a small part of this paper
17	"Big Data Analytics in Healthcare: Design and Implementation for a Hearing Aid Case Study (Jeppe Høy Christensen; Michael Kai Petersen; Niels Henrik Pontoppidan Eriksholm Research	Excluded	Not related to medical imaging

	Centre, Snekkersten, Denmark; Marco Cremonini. 2018)"		
18	"Using Big Data Analytics to Create a Predictive Model for Joint Strike Fighter (Ryan Norman;Jason Bolin;Edward T. Powell;Sanket Amin;John Nacker, 2018)"	Excluded	Not the area related to the topic
19	"Big Data Analysis and Services: Visualization on Smart Data to Support Healthcare Analytics (Carson K. Leung; Yibin Zhang; Calvin S.H. Hoi; Joglas Souza; Bryan H. Wodi, 2019)"	Excluded	Not a paper is talking about medical imaging
20	"Big Data Analytics APIs Architecture for Formative Assessors (Wassim Mahfouz;Heinz-Dietrich Wuttke, 2021)"	Excluded	This paper is not about medical industry
21	"Editorial for Special Issue of Journal of Big Data Research on "Big Medical/Healthcare Data Analytics" (Sakr, Sherif. 2018)"	Included	Topic is the part of this paper
22	"Big Healthcare Data Analytics in Internet of Medical Things (Allen, Margaret, 2020)"	Excluded	Related to medical field but not to medical imaging
23	"Big data analytics in medical engineering and healthcare: methods, advances and challenges (Wang, Lidong. 2020)"	Included	In this article topic is included and it's useful
24	"Cognitive Internet of Medical Things, Big Healthcare Data Analytics, and Artificial intelligence-based Diagnostic Algorithms during the COVID-19 Pandemic (Morrison, Michael; Lăzăroiu, George. (2021)"	Excluded	Not a relevant paper for my research area
25	"Smart Biomedical Sensors, Big Healthcare Data Analytics, and Virtual Care Technologies in Monitoring, Detection, and Prevention of COVID-19 (Morris, Kevin. 2021)"	Excluded	Not particularly speaks medical imaging
26	"BIG DATA ANALYTICS CAPABILITIES, THE BUSINESS VALUE OF INFORMATION TECHNOLOGY, AND HEALTHCARE ORGANIZATIONS: THE NEED FOR CONSENSUS IN EVIDENCE-BASED MEDICAL PRACTICES (Coatney, Karen. 2018)"	Excluded	Not related to medical field
27	"SNOMED CT-Based Standardized e-Clinical Pathways for Enabling Big Data Analytics in Healthcare (Ayman D. Alahmar, Rachid Benlamri. 2020)"	Included	This paper is useful in terms of medical imaging
28	"Deep Learning and Big Data in Healthcare: A Double Review for Critical Beginners (Luis Bote- Curiel, Sergio Muñoz-Romero, Alicia Gerrero- Curieses, José Luis Rojo-Álvarez. 2019)"	Included	Some areas speak about the topic.

29	"Visualization for Quality Healthcare: Patient	Excluded	Not a relevant document for
	Flow Exploration (Veronika Domova; Shiva		my studies
	Sander-Tavallaey. 2019)"		
30	"Influential Usage of Big Data and Artificial	Excluded	Not about the specific topic.
	Intelligence in Healthcare (Yan Cheng Yang, Saad		Only related to Big Data
	Ul Islam, Asra Noor, Sadia Khan, Waseem Afsar,		
	Shah Nazir. 2021)"		

Total Included = 10**Total Excluded** = 20

LIMITATIONS

The major limitation of this topic is the lack of relevant papers related to this particular topic. We will get so many results for the healthcare industry and data science field. But it is so difficult to find a paper about medical imaging with data analytics. Another limitation is the restrictions on choosing only research papers and articles. So much relevant information is in other types of publications. Another significant limitation was that the information technology industry is expanding and technologies are changing all the time. need to filter the results within 5 years. A major limitation is the shortage of papers related to the specific topic.

CONCLUSION

So many industries are widely using big data analytics. However, the clinical, medical imaging, and ECG sectors are still in the early stages of adopting big data, and research in these fields is still limited. Using proper analytics, we can extract more valuable information from the healthcare data generated by organizations. Hadoop can be used to generate meaningful real-time analytics for predictions on emergencies or assist in planning, strategizing, and making decisions. This research mainly spotlights the potential of big data analytics in the healthcare industry. specifically, for medical imaging. The purpose of this paper was to discuss the importance of big data analytics in medicine, medical imaging, and ECG, highlighting how they can improve the ECG field. From this research, we can conclude that big data analysis can play a major role in the healthcare industry particularly in medical image processing.

REFERENCES

- "Data-Enabled Digestive Medicine: A New Big Data Analytics Platform (Lu Yan, Weihong Huang, Liming Wang, Song Feng, Yonghong Peng, and Jie Peng, May/June 2021)"
- "A Latent Gaussian Copula Model for Mixed Data Analysis in Brain Imaging Genetics (Aiying Zhang, Jian Fang, Wenxing Hu, Vince D. Calhoun, Yu-Ping Wang, July-Aug. 2021)"
- "Group Sparse Joint Non-Negative Matrix Factorization on Orthogonal Subspace for Multi-Modal Imaging Genetics Data Analysis (Peng Peng, Yipu Zhang, Yongfeng Ju, Kaiming Wang, Gang Li, Vince D. Calhoun, and Yu-Ping Wang, Jan.-Feb. 2022)"
- "A Data-Driven Approach to Predict and Classify Epileptic Seizures from Brain-Wide Calcium Imaging Video Data (Jingyi Zheng,Fushing Hsieh,Linqiang Ge, Nov.-Dec. 2020)"
- "Integration of Imaging (epi)Genomics Data for the Study of Schizophrenia Using Group Sparse Joint Nonnegative Matrix Factorization (Min Wang,Ting-Zhu Huang,Jian Fang,Vince D.

- Calhoun, Yu-Ping Wang, Sept.-Oct. 2020)"
- "Identify Consistent Cross-Modality Imaging Genetic Patterns via Discriminant Sparse Canonical Correlation Analysis (Meiling Wang, Wei Shao, Xiaoke Hao, Li Shen, Daoqiang Zhang, July-Aug. 2021)"
- "Multi-Task Sparse Canonical Correlation Analysis with Application to Multi-Modal Brain Imaging Genetics (Lei Du, Kefei Liu, Xiaohui Yao, Shannon L. Risacher, Junwei Han, Andrew J. Saykin, Lei Guo, Li Shen, Jan.-Feb. 2021)"
- "A Novel Encoding and Decoding Calibration Guiding Pathway for Pathological Image Analysis (Hansheng Li, Jian Ping Li, Yuxin Kang, Chunbao Wang, Feihong Liu, Wenli Hui, Qirong Bo, Lei Cui, Jun Feng, Lin Yang, Jan.-Feb. 2022)"
- "Celiac Disease Detection From Video capsule Endoscopy Images Using Strip Principal Component Analysis (Bing Nan Li, Xinle Wang, Rong Wang, Teng Zhou, Rongke Gao, Edward J. Ciaccio, Peter H. Green, July-Aug. 2021)"
- "Using Big Data Analytics to Detect Fraud in Healthcare Provision (Spiros V. Georgakopoulos; Parisis Gallos; Vassilis P. Plagianakos, 2020)"
- "Segmentation and Classification of Brain MR Images Using Big Data Analytics (Mahboob Alam; Mohd Amjad, 2018)"
- "Big Data Analytics of Inpatients Flow with Diabetes Mellitus type 1: Revealing new awareness with Advanced Visualization of Medical Information System Data (Olga Kolesnichenko; Elena Marochkina; Roman Komarov; Lev Mazelis; 2019)"
- "Health Analytics on Big COVID-19 Data (Nguyen Duy Thong Tran; Carson K. Leung; Daryl L.X. Fung; Thanh Huy Daniel Mai, 2021)"
- "An Intelligent Visual Big Data Analytics Framework for Supporting Interactive Exploration and Visualization of Big OLAP Cubes (Carlos Ordonez; Zhibo Chen; Alfredo Cuzzocrea; Javier Garcia-Garcia, 2020)"
- "The Use of Big Data Analytics in Medical Images: a Survey (Mohammed Shatnawi; Muneer Bani Yassein; Ibtihal Jalabneh, 2020)"
- "Big Data Analytics in Healthcare (M. Ambigavathi; D. Sridharan, 2018)"
- "Using Big Data Analytics to Create a Predictive Model for Joint Strike Fighter (Ryan Norman; Jason Bolin; Edward T. Powell; Sanket Amin; John Nacker, 2018)"
- "Big Data Analysis and Services: Visualization on Smart Data to Support Healthcare Analytics (Carson K. Leung; Yibin Zhang; Calvin S.H. Hoi; Joglas Souza; Bryan H. Wodi, 2019)"
- "Big Data Analytics APIs Architecture for Formative Assessors (Wassim Mahfouz; Heinz-Dietrich Wuttke, 2021)"
- "Editorial for Special Issue of Journal of Big Data Research on "Big Medical/Healthcare Data Analytics" (Sakr, Sherif. 2018)"
- "Big Healthcare Data Analytics in Internet of Medical Things (Allen, Margaret, 2020)"
- "Big data analytics in medical engineering and healthcare: methods, advances and challenges (Wang, Lidong. 2020)"
- "Cognitive Internet of Medical Things, Big Healthcare Data Analytics, and Artificial intelligence-based Diagnostic Algorithms during the COVID-19 Pandemic (Morrison, Michael; Lăzăroiu, George. (2021)"
- "Smart Biomedical Sensors, Big Healthcare Data Analytics, and Virtual Care Technologies in Monitoring, Detection, and Prevention of COVID-19 (Morris, Kevin. 2021)"
- "BIG DATA ANALYTICS CAPABILITIES, THE BUSINESS VALUE OF INFORMATION TECHNOLOGY, AND HEALTHCARE ORGANIZATIONS: THE NEED FOR CONSENSUS IN EVIDENCE-BASED MEDICAL PRACTICES (Coatney, Karen. 2018)"

- "SNOMED CT-Based Standardized e-Clinical Pathways for Enabling Big Data Analytics in Healthcare (Ayman D. Alahmar, Rachid Benlamri. 2020)"
- "Deep Learning and Big Data in Healthcare: A Double Review for Critical Beginners (Luis Bote-Curiel, Sergio Muñoz-Romero, Alicia Gerrero-Curieses, José Luis Rojo-Álvarez. 2019)"
- "Visualization for Quality Healthcare: Patient Flow Exploration (Veronika Domova; Shiva Sander-Tavallaey. 2019)"
- "Influential Usage of Big Data and Artificial Intelligence in Healthcare (Yan Cheng Yang, Saad Ul Islam, Asra Noor, Sadia Khan, Waseem Afsar, Shah Nazir. 2021)"
- "5: Big Data analytics in medical imaging Applications of Big Data in Healthcare (Siddhant Bagga Sarthak Gupta Deepak Kumar Sharma, 2021)"
- "Radiogenomics for Precision Medicine With a Big Data Analytics Perspective (Andreas S. Panayides; MariosS.Pattichis;Stephanos Leandrou;Costas Pitris;Anastasia Constantinidou;Constantinos S. Pattichis. 2019)"