A Survey on Big Data Application for Modality and Physiological Signal Analysis

Madeleine Liu Qinghui
[1]School of Design, India China Institute, NY 10011.
[1]liuqin2@hotmail.com

ArticleInfo
R. Arulmurugan et al. (eds.), *First International Conference on Machines, Computing and Management Technologies*, Advances in Intelligent Systems and Technologies
Doi: https://doi.org/10.53759/aist/978-9914-9946-0-5_1

Abstract – An explosion of healthcare data has occurred in recent years due to the widespread availability of sophisticated physiological signal monitoring devices and the development of telemetry and cognitive communication systems. Additionally, the accessibility of medical data for the establishment of applications in big data has rapidly increased due to affordable and efficient storage and power techniques. With the current state of technology, healthcare professionals are unable to effectively handle and understand large, rapidly changing, and complex data; this is where big data applications come in. Making medical services more cost-effective and sustainable is a driving force behind the creation of such systems. In this article, we present a discussion of the present condition of big data applications that make use of physiological signals or derived metrics to aid in medical decision making in the home and in the hospital. Specifically, we examine critical care systems designed for continuous healthcare management and address the obstacles that must be surmounted before such systems may be used in real-world practice. Big data technologies might revolutionize future hospital administration if these problems are solved.

Keywords – Big Data, Big Data Analytics, Electronic Health Records, Data Modality, Physiological Signal Analysis

I. INTRODUCTION

Large amounts of physiological data collected both in hospitals and in patients' homes, have opened the door to a new age of data-driven methods for enhancing patient care in the areas of diagnosis, treatment, and results. The interpretation of available data is a significant obstacle on the path to better and quicker (often real-time) decision making, higher quality patient care, and improved health outcomes. With the use of machine learning techniques, large amounts of physiological data may help doctors make more accurate diagnoses and allow for real-time online monitoring in the clinic or at home. A clinical expert's confirmation bias may be avoided, and rapid, accurate decisions can be made at a lower cost with the use of clinical decision support systems, which make use of Deep Learning and Big Data techniques [1]. Nonetheless, healthcare datasets are often complex, noisy, and continuously changing. These datasets may include structured (such as lab results), semi-structured (such as sensor information), and unstructured (such as patient records in handwritten form), data. Due of the intricacy, accurate interpretation is challenging, and rigorous validation is necessary. In addition, new data sources, training, and escalation procedures will be needed to successfully implement these technologies into clinical practice.

In [2], Electronic Health Records (EHR) refers to the result of big data in hospitals, collecting details on patients' demographics, medical histories, diagnostics, and therapies. Despite the excellent quality of the data stored inside, its application in medical decision making is limited since the database is too massive and complex to be efficiently analyzed by humans. Clinical personnel need data analytics solutions, often powered by machine learning, to assist with data organization, pattern detection, result interpretation, and action threshold setting. Semantic context analysis of these records may also benefit greatly from the use of natural language processing, a big data-powered technology. In hospitals, for instance, big data analytics have already been put to use for purposes such as expanding existing knowledge, bettering clinical service, and streamlining public health monitoring. Smart systems based on the P4 clinical philosophy of preventive, predictive, participatory and individualized care may be developed by utilizing big data and establishing machine learning models, which provide continual learning strategies with real-time skill creation. One such innovative analytics platform is the Hospital Surveillance, Monitoring and Alert (HSMA) systems [3] that have been utilized at Sao Paolo Hospital from 2012. This system takes use of the massive amounts of data held inside Electronic Health Records (EHRs) in real-time and utilizes lightning-fast and sophisticated analytics to aid in clinical decision making.

More than 600 Key Performance Indicators (KPIs) are computed and generated by the system, many of which are directly connected to clinical care but also provide valuable information to hospital administration. Because having access to a database of thousands of patient records. The HSMA system delivers desktop dashboards with real-time information to hospital personnel, and it goes further by continuously tracking different physiological metrics from every patient during their stay and using clever prediction models for the early detection of patients at risk, indicated as red and yellow and amounting to treatments. When there is an abundance of data on the "normal" state but a dearth of "abnormal" data, novelty detection may be a useful machine learning strategy for delivering such alarms. Clinical applications of novelty detection approaches, which are often characterized as a one-class grouping methodology, have been made in the ED, the ICU, and for post-operative individuals. Even in the intensive care unit, Gaussian processes have been widely utilized to handle noisy and inaccurate physiological data. Biomedical decision support advancements have been recommended by Özen Kavas, Recep Bozkurt, Kocayiğit, and Bilgin [4] for various uses, including earlier impairment warnings in continual monitoring, assessment and improvement of protocol adherence, prescription reminders, improvement of screening, and prediction of hospital readmission.

The explosion of healthcare data in recent years may be traced back to the development of telemetry and smart communication models, as well as the widespread use of sensors capable of continuously monitoring a patient's physiological signals [5]. In the past, it was not common practice to keep the continuous physiological data provided by wearable sensors for extended periods of time. However, with the advent of efficient and cheaper storage and power systems, this has changed dramatically. The proliferation of big data healthcare applications is a direct result of the increasing complexity, variety, and contextual richness of medical data. The study of healthcare data sets that are too large, too quick, or too complicated to be handled by traditional data analysis methods is the focus of "big data" application in the medical industry. In light of the fact that the global population is aging at an alarming rate, there is a pressing need to build such systems in order to ensure the long-term viability of healthcare provision. To meet these societal and economic concerns, researchers are focusing on creating smarter healthcare systems with the intention of revolutionizing healthcare administration by lowering costs without compromising quality of treatment.

As of late, 2D medical imaging analysis by Masuda, Shimizu, Nakazawa, and Edamoto [6] has shown the great promise of Deep Learning (DL), an effective methodology for generative and discriminative tasks. Nonetheless, physiological data one-dimensional signals have not been exploited from this new methodology to achieve desired clinical tasks. As a result, in this study, we provide a comprehensive review of the most recent scientific literature on the application of deep learning to electroencephalogram (EEG) [7], electromyogram (EMG) [8], electrooculogram (EOG) [9] and electrocardiogram (ECG) [10]. We compiled a total of 147 articles from different journals and publishers that were all published between January 2018 and October 2019. The purpose of this work is to undertake a comprehensive analysis of the deep-learning algorithms utilized in physiological signal evaluation for distinct clinical applications in order to better understand, classify, and compare these methods' most important aspects.

We examine the data input type, dataset sources, training infrastructures, deep-learning framework, and deep-learning issues as the major parameters of the deep-learning technique. These are the most influential settings for each given system. We classify the literature on the employment of deep-learning algorithm to physiological signal analysis according to the following criteria: (1) Physiological Signal Analysis and Modality, (2) Data Modality and Medical Application in continuous patient monitoring, and (3) Real-World Clinical Problems (evaluated in separate section). The criteria are presented in the following article organization: Section II focuses on Big Data and physiological signals, where the aspect of physical signal analysis and modality, and data modality and medical application in continuous patient monitoring, has been discussed. Section III focuses on the challenges of real-world clinical problems. Section IV presents concluding remarks to the article.

II. BIG DATA AND PHYSIOLOGICAL SIGNALS

Physiological Signal Analysis and Modality

Human health status may be estimated via the study of physiological signals. Reports, readings, and behaviors are the three forms of measurements used to capture physiological signals. The "report" is the respondents' ratings of their own physiological conditions in response to a questionnaire. In this context, "reading" refers to the recorded data obtained from a device that measures some aspect of human body e.g., brain functionality, heart rate, and muscular strength, etc. In the "behavior" metric, activities such as eye blinking and other subtle gestures are recorded. Because the "report" answer is more subjective, has less precision, and has greater variation in its question scale, we chose not to analyze it in this article. Here, we zero down on a specific method of measuring "reading" and "behavior" in which the data are presented in the form of electrical muscle activity, electrocardiogram activity, electrooculogram activity, or a combination of these.

Physiological signal modality employed in the development of medical applications is outlined in **Table 1**. According to Zhou, Yu, Chen, Wang, and Arshad [11], EMG signal's "muscle tension pattern" decodes the motion of the hand and the muscles' activities. Heart condition, sleep phase, emotional state, age, and gender may all be identified from a person's unique heart rhythm. Brain illness, mood, sleep-stage, motion, gender, speech, and age may all be classified using the EEG signal's wide variety of responses. Classifying the stages of sleep based on the EOG signal's variations in ocular corneo-retinal potential is possible.

In this section, we classify the different deep-learning models according to the modalities of the physiological signal data they need. We show how our work significantly improves medical system performance and utility. We classify works in deep learning for EMG, ECG, EEG, EOG, and combined data. We evaluate the deep-learning model quantitatively and qualitatively. Statistically, this diagram shows how many deep-learning models have been put to use in the medical field thus far. Since the performance criteria is not consistently supplied, we use an accuracy number as a starting point for a more holistic evaluation of the results.

Table 1. Clinical applications in the analysis of physiological signal

Signal Modality	Clinical Application
Combination of signals	Sleep-stage classification
EOG	Sleep-stage classification
EEG	Age classification, Words classification, Gender classification, Motion classification, Emotion classification, Brain disease classification, Brain functionality classification
ECG	Gender and age prediction, Emotion classification, Heart disease classification, Heartbeat signal classification, Sleep-stage classification
EMG	Muscle activity identification, Hand motion identification

Deep Learning with Electromyogram (EMG)

Non-invasive EMG electrodes, such as the popular MYO Armband on the market, capture electromyogram (EMG) signals, which are data on changes of skeletal muscles. The ability to recognize a motion pattern, e.g., closed or open hand, is due to the fact that each muscle data is featured by various activities. The several deep-learning models (**Fig. 1**) used to identify hand motion and (**Fig. 2**) utilized to identify muscle activation are shown below. CNN+RNN or CNN algorithms were widely utilized for hand motion identification. CNN framework is widely utilized for identification of muscle activities.

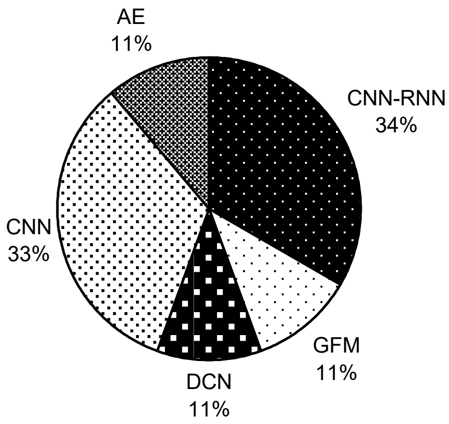

Fig 1. No. of DL algorithms utilized in EMG data for (a) detection of the hand motion.

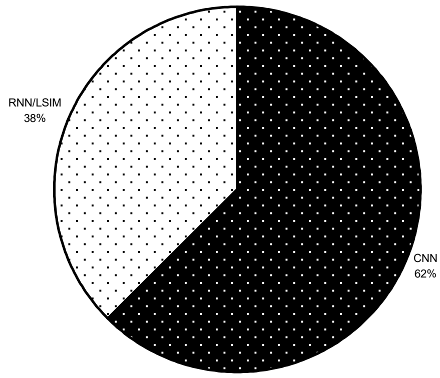

Fig 2. No. of DL algorithms utilized in EMG data for (b) identification of muscle activity

Deep Learning with Electrocardiogram (ECG)

An electrocardiogram (ECG) records information about variations in heart rate and rhythm. In 47 studies, researchers have used deep learning to the analysis of ECG data [12]. Their major contributions may be broken down into five distinct groups: gender and age prediction; emotion recognition; sleep stage categorization; heart disease categorization; heartbeat signal categorization.

Below, we can see how many deep-learning frameworks are employed in ECG signal analysis: (**Fig. 3**) shows how CNN is used for classifying heartbeat signals, (**Fig. 4**) shows how CNN is used for classifying heart diseases, (**Fig. 5**) shows how CNN is used for classifying sleep stages, (**Fig. 6**) shows how CNN+RNN and RNN/LSTM are employed for classifying emotions, and (**Fig. 7**) shows how CNN is used for classifying age and gender.

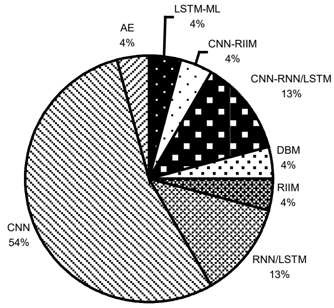

Fig 3. Numerous DL models are applied to ECG data for the following purposes: (a) classification of heartbeat signals;

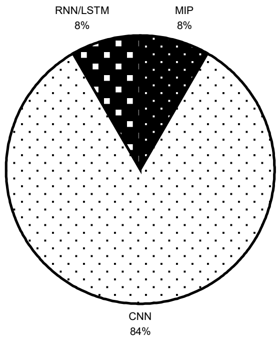

Fig 4. Classification of heart conditions.

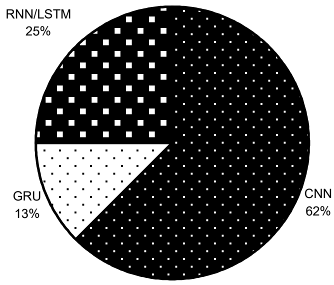

Fig 5. Detection of sleep stages

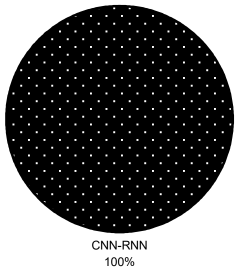

Fig 6. Detection of emotions

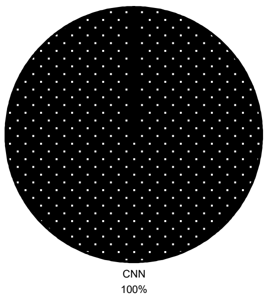

Fig 7. Classification of age and gender.

Deep Learning with Electroencephalogram

An electroencephalogram (EEG) [13] is capable of recording the brains' electrical activities and displays the results graphically on the user's head. There have been 79 studies conducted that have used deep-learning techniques to examine EEG data. Classification of brain functions, classification of brain diseases, classification of emotions, classification of sleep stages, classification of movements, classifications of words, classification of gender, and classification of age are all among their most significant contributions. The number of deep-learning systems employed for EEG data analysis is shown below. (**Fig. 8**) shows examples of classification activities where the CNN frameworks is most popular; (**Fig. 9**) indicates the examples of the classification activities where CNN is most popular when dealing with brain diseases; (**Fig. 10**) shows examples of classification activities where CNN is most popular when dealing with emotions; (**Fig. 11**) shows examples of classification activities where CNN is most popular when dealing with sleep-stage classification; (**Fig. 12**) shows examples of classification activities where CNN is most popular when dealing with motion; **Fig. 13**) shows gender categorization using only CNN, **Fig. 14**) shows word recognition using just AE, and **Fig. 15**) indicates the classification of age using CNN alone.

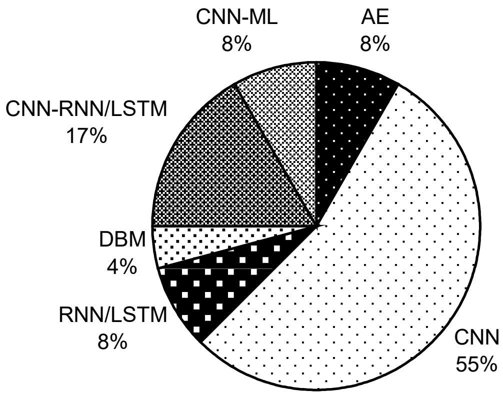

Fig 8. Classifying brain functions

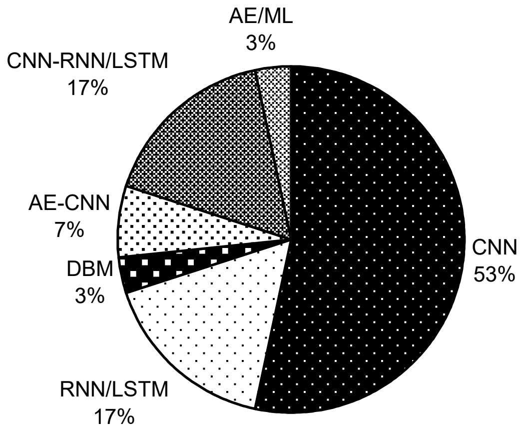

Fig 9. Classifying brain diseases

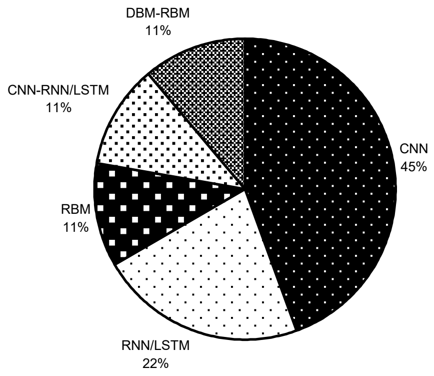

Fig 10. Classifying emotions

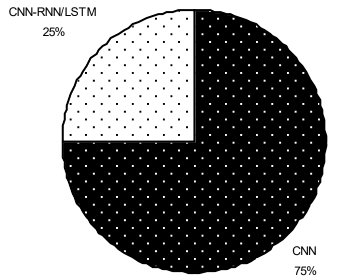

Fig 11. Classification motion

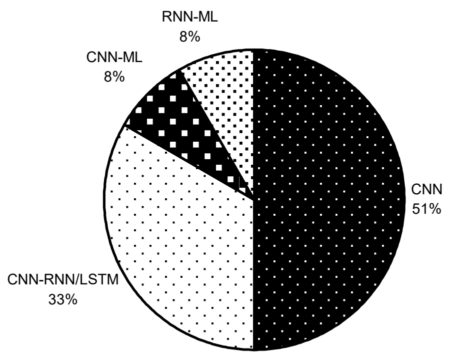

Fig 12. Classifying sleep stages

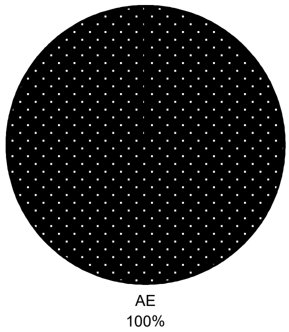

Fig 13. Classifying gender

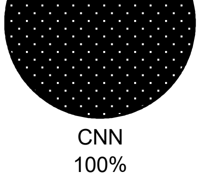

Fig 14. Classifying language

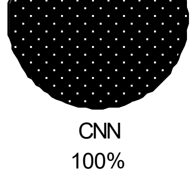

Fig 15. Classifying age

Data Modality and Medical Application in Continuous Patient Monitoring

Using either continual physiological spectral information or inferred physiological characteristics obtained from various sources, big data analytics systems in distant patient monitoring aim to offer advanced warning of impairments. Prior until recently, most patients surveillance systems depended on a single data source (such as a single pulse oximeter), which might cause "alarm fatigue," a situation in which healthcare workers get desensitized to and eventually dismiss signals from the surveillance technology. By combining data from numerous sources, including multiple concurrently obtained vital signs and other patient data (of diverse phenotypes), big data applications have opened up the possibility of developing enhanced and more complete early warning systems. This has opened up the possibility of examining the interconnections and interactions among heterogeneous medical time series information or waveforms for obtaining clinical knowledge, which has the added benefit of enhancing alerting accuracy.

An Intensive Care Unit (ICU) is a hospital setting where a large volume of physiological data is generated rapidly due to the use of multiparameter monitors to constantly capture waveforms and provide periodic metrics of vital signs e.g., Temperature (T), Respiratory Rate (RR) [14], Arterial Blood Pressure (ABP) [15], peripheral arterial oxygen saturation (SpO_2) [16], and heart rate (HR) [17]. Data from electronic health records, pharmaceutical infusion pumps, and ventilators may all be used to construct big data application for clinical decision support and patient monitoring in real-time. Numerous papers have surfaced in the literature suggesting big data techniques, indicating widespread recognition of the possibility for establishing these systems in an ICU setting. Many studies have concluded that conventional methods (such as single pharmacological treatments) are unable to deal with the intricacies of critical disease (frequently with comorbidities), and that big data-driven data fusion methods may be a better fit. An example of a big data application in critical care is a patient care managerial model, which utilizes constant and static data supervised from critically ill patients that have been admitted to perform data mining and alert clinical staff for critical medical situations in real-time (a program designed for a neonatal ICU that used streaming).

Logistic regression was used to create predictive models to create an estimation of ICU transfers, mortality rates and cardiac arrest in ward patients in one of the first researches about ICU patients. Improved predictive accuracy for early identification of deterioration was shown in a follow-up analysis by Janous, Kosan, Talla, and Peroutka [18], who swapped out logistic regression for advanced ML algorithms including gradient enhancement and random forest techniques. Utilizing the Physionet database of the Multi-parameter Intelligent Monitoring in Intensive Care (MIMIC) study, Saeed et al. [19] were able to properly estimate mortality rate. Using clustering algorithms to categorize patients according to their organ dysfunction, the scientists were able to increase the AUC to 0.91. When Bergquist et al. [20] assessed the predictive performance of approximately twelve mortality prediction frameworks (both parametric and non-parametric) using the same database, they found that ensemble approaches using weighted output of various algorithms (Superlearner) has a greater projected performance, with AUB of approximately 0.88.

Researchers now have the chance focus on a research on the ideology of big data applications for information exchange thanks to datasets like MIMIC III (Multiparameter Intelligent Monitoring in the Intensive Care), which includes physiologic sensory information and vital signs series data procured from clinical facilities and extensive medical data gathered from medical data systems [21]. Consequently, the MIMIC dataset has been put to use for a variety of purposes, such as comparing patients within carefully chosen cohorts and creating algorithms that combine different waveform information to forecast cardiovascular instabilities in patients. In addition, the MIMIC II database's collection of numerous waveforms was used to create a method for detecting hemodynamic instabilities in patients at an early stage. Gericke [22] developed a system to estimate cardiac output using pulses contour analysis, Saha [23] developed a system to identify hypovolemia using photoplethysmography data, and Elgendi, Galli, Ahmadizadeh, and Menon [24] developed a system to predict hyperlactemia using integrated physiological data.

Wearable sensors of consumer quality are becoming more common, opening the door for the establishment of big data applications that utilize physiological sensor dataset outside of a clinical setting. Because of these sensors, formerly bulky and expensive physiological monitoring equipment may now be carried about easily. A number of methods have been developed for early diagnosis of cardiac arrest in hospitals by combining telemetric ECG readings with patient demographics such medical history, pericardial effusion, laboratory results, and medicines. However, as in clinical applications, it might be difficult to combine data from several mobile devices at once. Lee [25] conduct a systematic review of the many different types of sensors, both stationary and mobile, that may be used for data analysis in the healthcare industry and how this information can be used to advance the state of the art in patient care technology.

III. BIG DATA ANALYTICS AND CHALLENGES IN REAL-LIFE CLINICAL PROBLEMS

A. Big Data Theoretical Framework

The rapidly expanding availability of high-quality data is poised to dramatically alter the healthcare industry in the not-too-distant future. According to Resnyansky [26], the concept of Big Data Analytics (BDA) is a crucial tool for implementing management principles in the healthcare industry. Therefore, the concept has attracted much attention of the general public and the digital world. In order for statistics to be significant in the management of medical organization, it is fundamental to first assess their quality, which necessitates establishing a premise: statistics alone explain nothing. In other words, the dataset needs to be processed in the appropriate way using the appropriate tools before any conclusions can be reached.

The use of technology for healthcare BDA is growing rapidly and will have an increasing impact on managerial decision-making in the future. For instance, the project of IBM Watson is a "super-computer", which utilizes AI advancements (such as Machine Learning) to analyze millions of scientific publications accumulated over a twenty year period to link medical symptoms and forecast likely diagnosis scenarios. By analyzing this scenario, we can learn more about what and how BDA could aid medical managers enhance the quality of their decision-making process and the effectiveness of their medical organizations.

According to Patrizio [27], the absolute quantity of data is now an issue. Networking giant Cisco estimates that by 2025, the digital universe will fill 44 zettabytes, with daily data creation possibly reaching 463 exabytes. Today marks the beginning of a new era in which robots, alongside humans, will play a vital role in the establishment and management of human knowledge. The healthcare and pharmaceutical industries will both benefit greatly from the implementation of IoT and AI technologies in the near future. In modern medicine, IoT-based applications allow for remote monitoring of clinical data, with information being generated by specific pieces of equipment (e.g., wearable devices) and then made available to doctors rather than caregivers in the form of reports.

The extent to which healthcare providers are focusing on novel management strategies on the employment of big data could be determined by investigating the extent of the general market. The health care big data marketplace is projected to approach $80 billion by 2030, an increase of 568 percent over the preceding decade. In addition to posing a challenging problem, the introduction of such a technology presents an exciting new opportunity for everyone involved in the healthcare supply chain that plays a role in making important decisions. Also, this technology will improve the efficiency and effectiveness of healthcare delivery, and it may even play a role in determining the creation of creative management techniques inside healthcare organizations. Indeed, healthcare administrators utilize big data technology to gather information like a staff roster, a list of pharmaceuticals and their expiry dates, etc., to improve overall management, streamline decision-making, and elevate service quality while decreasing resource waste.

Which goals does Big Data Analytics (BDA) serve, and how may its potential in healthcare administration be maximized? Many healthcare administrators and politicians have given serious thought to this question. Multiple studies and articles published in the last several years highlight the benefits of BDA adoption medical organization, especially in the fields of process management and resource efficiency. In 2013, Dutta [28] established a framework using the BDA Capabilities (BDAC) and Resource-Based Theory to define the connectiions between BDA, value generation and the benefits of medical providers. According to Bui et al. [29], a BDAC is "a facility with tools, processes, and procedures that allows an organization to process, organize, present, and analyze data, therefore providing insights that enable data-driven operational strategy, decision-making, and execution." BDAC in the medical sector is helpful in facilitating the capacity to process, analyze, store and collect massive amounts of medical data from distinct mobile sources to enhance the process of decision-making driven by data.

Indeed, Rahman and Reza [30] work has shown how separate "path-to-value" may be established; this has been corroborated empirically by 109 cases of BDA frameworks employed in 63 medical institutions. Different levels of applicability of the suggested approaches have shown that there are various benefits for healthcare organizations alongside the challenges of installing particular BDA instruments. The ability to assess large datasets with the help of the ideology of Information Lifecycle Management (ILM) [31] has been described in the preliminary study. According to this theory, the primary function of the BDA in medical organizations is to evaluate health data collected from a variety of sources and then report their findings to the healthcare management in a relevant way. In order to boost their company's performance, managers may use BDA to identify timely indicators and develop business strategies, and then put in place long-term goals, effective tactics, and streamlined procedures.

Authors generally point to the incorporation of new features and tools as the driving force for BDA's efficacy. First and foremost are data processing tools (OLAP, machine learning, and natural language processing, for example), then data aggregation tools (data storage tools), and the capabilities and tools for data presentation (reporting interfaces/systems, and visual systems/dashboards). Among the many benefits that using BDA has given to the healthcare sector, the ability to analyze data more effectively has been the most significant. This possibility involves using descriptive analytic techniques to examine time-sensitive, heterogeneous, and massive amounts of clinical data in various formats (from text to graph). As such, it is fundamental to consider that BDA-oriented management models are the only ones capable to evaluate unstructured and semi-structured data. This is crucial in identifying patterns of association that are obscured by more traditional forms of management. Moreover, a healthcare organization that implements such systems has the ability to precisely regulate care service and process outputs to sustainably boost the efficiency of its operations. In conclusion, the following are characteristics of management systems based on BDA that are used in a hospital setting: (i) the ability to look at data, find vital trends, patters, corrections, and project their likelihood for the future; (ii) the capacity to connect processes, and data to enhance the process of sharing, collaboration, and management across various medical care systems.

Improved healthcare decision quality and accuracy, quicker issue resolution, and the capacity to provide therapies prior to a worsening in patients' situations are some of the most important improvements from BDA implementation, according to the study of Noby, Rady, and Gaber Abd Eljalil [32]. IT infrastructure benefits including standardization, reduced expenses for redundant equipment, and rapid data transmission across IT systems came in a close second. Their impact on the area is significant; they developed a useful business model to help hospital administrators decide which levers to pull in tandem with the introduction of BDA-oriented management frameworks. In addition to re-iteration of the apparent benefits, Bibri and Krogstie [33] discuss particular instances of how certain BDA technologies may facilitate faster and more accurate decision-making on the part of hospital managers.

Enikeeva [34] examined 26 case scenarios linked to the applications in the medical organizations and identified five different "capacities" of BDA: (i) the capacity to analyze care patterns; (ii) the capacity to analyze unstructured data; (iii) the capacity to provide decision support; (iv) the capacity to forecast outcomes; and (v) the capacity to leave a trail of information. As exciting as the study's mapping of precise benefits is, the study's suggestions for how healthcare organizations could effectively employ BDA are even more so. These strategies are vital for reaching one's goals by making the most of BDA's potential.

By outlining objectives, procedures, and KPIs, big data-based governance may be established as the first efficient method. Integrating information systems and standardizing data standards, which might originate from contradictory sources within healthcare organizations, are crucial to the success of executing such a plan. With regards to the second, it has to do with encouraging a culture of open data sharing. The third aspect is training for healthcare administrators who need to understand BDA in order to supervise the usage of data mining and BI technologies. As a fourth option, businesses may ditch expensive Network-Attached Storage (NAS) devices in favor of more modern and flexible alternatives like cloud computing to store huge amounts of data in several forms. The broad use of prognostic BDA models is the last critical component. In order to effectively use BDA-enabled process management methodologies, healthcare managers and organizations must have a firm grasp of KPIs, interactive visualizations, and data gathering tools such as reports and dashboards.

Recent works by Juma and Kilani [35] have examined the supply chain management practices of the healthcare sector. Through in-depth interviews with China's top hospital administrators, Zhang [36] show how BDAC has a positive impact on all the three aspects of medical Supply Chain Integration (SCI) (hospital-supplier integration [37], hospital-patient integration [38], and inter-functional integration [39]). The ability of a healthcare institution to adapt to new circumstances by making necessary adjustments to routine procedures without affecting the quality of care provided to patients is what is meant by the phrase "operational flexibility". By showing how BDAC, SCI, and operational flexibility are all intertwined, the researchers have made a substantial contribution by offering practical management advice to hospital executives and supply chain executives. By processing and reviewing managerial and clinical data using cutting-edge analytic approaches, the Chinese health system was able to track humans' whereabouts during the Coronavirus period of lockdown, ascertain prevailing medical trends, and government pharmaceutical distributions. This theoretical model might help you comprehend the advantages of BDA-inspired best practices in healthcare delivery.

Using Big Data for Challenges in Healthcare

Intel and Assistance Publique-Hôpitaux de Paris (AP-HP), Europe's biggest hospital institution, have developed a cloud-based technology that utilizes a variety of data sources to improve the prediction of ER visits and hospital admissions. By using this predictive analytics tool, healthcare management at AP-HP health care facilities will be able to decrease patient wait times, maximize available Human Resources (HR) in light of anticipated demands, plan patient loads with pinpoint accuracy (even at the pathology level), and enhance the quality and effectiveness of care delivered. Patients with chronic illnesses impose a load on healthcare budgets and risk health decline without regular drug adherence. The European Commission's initiative to launch manufacture of the medicine Enerzair Breezhaler is another example of a project that effectively integrated BDA technologies into health administration. It was the first time asthma medications were combined and prescribed using the Propeller digital platform. The program reminds the user to take their medicine and logs their dosages so that the user may share this information with their doctor. A study found that patients who used the Propeller platform had a 63% improvement in asthma control, a 58% improvement in adherence to treatment, and a 57% decrease in asthma-related admissions and emergency room visits.

Even while the suggested practical structure is supported by some real-world experience, it is insufficient to fully demonstrate the power of BDA. As BDA-based management systems are rolled out throughout an organization's healthcare infrastructure, it is expected that a positive feedback loop would be triggered, resulting in the collection of ever-better quality patient data. BDA will permit predictive modeling utilizing cutting-edge AI methods, which will help doctors create more efficient diagnostic paths and managers put those findings to use. Its use in healthcare decision-making will benefit both healthcare professionals and patients, as well as save money and improve service quality for healthcare providers.

There has been a lot of excitement about using big data in healthcare [40] recently, but many obstacles still need to be cleared away before big data technologies can be put to use solving actual clinical issues. While data quality and analysis are not healthcare-specific difficulties to be addressed, the collecting of data from real patients presents extra difficulties not seen in many other professions. Datasets containing missing or damaged data are common when working with data acquired from people; these issues, if overlooked, might distort or corrupt the analysis, leading to incorrect decision making. Such problems may arise, for example, when data is acquired from wearable sensors due to a misplaced or improperly mounted sensor or due to motion. The ergonomics and attachment tactics of sensors need to be enhanced to limit the amount of erroneous data they produce. There may be a way to fix this issue using non-contact vital sign monitoring.

If contaminated data has already been obtained and cannot be erased, data imputation and quality review may be valuable. Linking different data sources is a challenge when employing physiological set of data in big data clinical application. It can be difficult to incorporate different sources of data in applications of big data, despite the fact that data fusion (processing data from various sources) has been widely used to provide a more comprehensive understanding of the problem and to integrate measurements where quality issues frequently occur. However, more work is required so that a higher value may be procured from simultaneously retrieved physiological data. Data fusion methodologies, for example, Kalman Filters (Durrant-Whyte and Henderson), feature selection methods, and Many Kernel Learning (MKL), are techniques that allow the incorporation of multiple sources of data concurrently.

To manage unevenly-sampled, heterogeneous, sparse, incomplete, and noisy datasets, Multi-Task Gaussian Processes (MTGP) approaches have been developed. In addition, high-quality, low-cost storage and processing methods that provide fast data pulls and commits in response to analytics needs must be established. In addition, there are a number of governance difficulties that must be surmounted, such as those interconnected with the enhancement of suitable data protocols and standards, as well as those concerning data privacy. Another challenge in creating comprehensive big data applications is the perception that data obtained from various sources may be stored in separate databases. To successfully implement this relationship, a number of legal and ethical considerations must be met, and this sometimes necessitates collaboration amongst otherwise rival system makers.

There are several operational hurdles to overcome when using big data approaches in real-world clinical applications, in addition to the aforementioned issues with analysis and quality. It calls for fresh approaches in therapeutic practice and thought. The use of data mining and machine learning in big data analysis necessitates being open to unexpected results from the search for patterns. It is important to use rigorous procedures for verifying results in order to assure their credibility and statistical significance, which is where this method varies dramatically from the traditional scientific strategy of beginning with a particular study topic.

IV. CONCLUSION

The development of high-tech tools for monitoring physiological signals, together with advances in cognitive communication and telemetry, has resulted to a deluge of data in the healthcare industry in recent years. More affordable and efficient power and storage methods have also considerably enhanced the presence of medical data for big data applications and development. The applications are necessary because doctors and nurses presently lack the tools to deal with and make sense of data sets that are very massive, dynamic, and complex. One motivation for developing such systems and applications is the need to make medical care more efficient and sustainable. In this article, we reviewed the current status of big data applications that leverage physiological fluctuations or derived metrics to improve in-home and in-clinic medical decision making. The main obstacle in the way of big data applications making it into clinical practice is training clinicians in new ways of thinking, inferring information, reacting to situations, and taking action based on what they've learned. Therefore, researchers in the healthcare sector should focus their attention not only on the technological application of such models, but also on the right processes for the integration of medical practices. Once this issue has been solved, big data applications could have a profound effect on not just clinical trials but also community health.

References

[1]. G.-G. Wang, X.-Z. Gao, and Y. Pei, “Call for special issue papers: Deep learning assisted big data analytics for biomedical applications and digital healthcare: Deadline for manuscript submission: August 20, 2022,” Big Data, vol. 10, no. 1, pp. 85–86, 2022.

[2]. H. Maarsingh et al., “Implementing electronic health records on a medical service trip improves the patient care process,” Front.Health Serv., vol. 2, 2022.

[3]. K. J. McKay, C. Li, and R. Z. Shaban, “Using video-based surveillance for monitoring hand hygiene compliance according to the World Health Organization (WHO) Five Moments framework: A pragmatic trial,” Infect. Control Hosp. Epidemiol., pp. 1–7, 2022.

[4]. P. Özen Kavas, M. Recep Bozkurt, İ. Kocayiğit, and C. Bilgin, “Machine learning-based medical decision support system for diagnosing HFpEF and HFrEF using PPG,” Biomed. Signal Process. Control, vol. 79, no. 104164, p. 104164, 2023.

[5]. J. Coffin, “Strategic placement. Addressing the data explosion in healthcare requires a sound strategy for turning data into information,” Health Manag. Technol., vol. 29, no. 2, pp. 46, 48, 2008.

[6]. K. Masuda, T. Shimizu, T. Nakazawa, and Y. Edamoto, “Registration between 2D and 3D ultrasound images to track liver blood vessel movement,” Curr. Med. Imaging Rev., 2022.

[7]. A. N. Tasyakuranti, H. Sumarti, H. H. Kusuma, I. Istikomah, and I. S. Prastyo, “Analysis of the effect of istighfar dhikr to adolescent anxiety at beta wave activity using Electroencephalogram (EEG) examination,” J. NEUTRINO, vol. 15, no. 1, pp. 31–37, 2022.

[8]. S. M. Debbal, “Pathological Electromyogram (EMG) Signal Analysis Parameters,” Clinical Cardiology and Cardiovascular Interventions, vol. 4, no. 13, pp. 01–14, 2021.

[9]. F. A. Azhar et al., “The classification of electrooculogram (EOG) through the application of linear discriminant analysis (LDA) of selected time-domain signals,” in Lecture Notes in Electrical Engineering, Singapore: Springer Singapore, 2022, pp. 583–591.

[10]. H. Ko, K. Rim, and J. Y. Hong, “Bio-metric authentication with electrocardiogram (ECG) by considering variable signals,” Math. Biosci. Eng., vol. 20, no. 2, pp. 1716–1729, 2023.

[11]. J. Zhou, K. Yu, F. Chen, Y. Wang, and S. Z. Arshad, “Multimodal behavioral and physiological signals as indicators of cognitive load,” in The Handbook of Multimodal-Multisensor Interfaces: Foundations, User Modeling, and Common Modality Combinations - Volume 2, Association for Computing Machinery, 2018, pp. 287–329.

[12]. E. A. Maharaj and A. M. Alonso, “Discriminant analysis of multivariate time series: Application to diagnosis based on ECG signals,” Comput. Stat. Data Anal., vol. 70, pp. 67–87, 2014.

[13]. L. D. Barnes, K. Lee, A. W. Kempa-Liehr, and L. E. Hallum, “Detection of sleep apnea from single-channel electroencephalogram (EEG) using an explainable convolutional neural network (CNN),” PLoS One, vol. 17, no. 9, p. e0272167, 2022.

[14]. L. Gagliardi, F. Rusconi, and the working party on respiratory rate, “Respiratory rate and body mass in the first three years of life,” Arch. Dis. Child., vol. 76, no. 2, pp. 151–154, 1997.

[15]. H. Mizuno, S. Hoshide, R. Nozue, D. Shimbo, and K. Kario, “Associations of office brachial blood pressure, office central blood pressure, and home brachial blood pressure with arterial stiffness,” Blood Press. Monit., vol. 27, no. 3, pp. 173–179, 2022.

[16]. P. Sirohiya et al., “A correlation analysis of peripheral oxygen saturation and arterial oxygen saturation among COVID-19 patients,” Cureus, vol. 14, no. 4, p. e24005, 2022.

[17]. G. S. Costa, L. S. Julião-Silva, V. S. Belo, H. C. F. de Oliveira, and V. E. Chaves, “A systematic review and meta-analyses on the effects of atorvastatin on blood pressure and heart rate,” Eur. Heart J. Cardiovasc. Pharmacother., 2022.

[18]. S. Janous, T. Kosan, J. Talla, and Z. Peroutka, “Improved accuracy of model predictive control of Induction motor drive using FPGA,” in 2019 IEEE International Symposium on Predictive Control of Electrical Drives and Power Electronics (PRECEDE), 2019.
[19]. M. Saeed et al., “Multiparameter Intelligent Monitoring in Intensive Care II: a public-access intensive care unit database,” Crit. Care Med., vol. 39, no. 5, pp. 952–960, 2011.
[20]. T. Bergquist et al., “Evaluation of crowdsourced mortality prediction models as a framework for assessing AI in medicine,” bioRxiv, 2021.
[21]. R. Bharathi and T. Abirami, “Energy aware clustering with medical data classification model in IoT environment,” Comput. Syst. Sci. Eng., vol. 44, no. 1, pp. 797–811, 2023.
[22]. O. R. Gericke, “Spectrum and contour analysis of ultrasonic pulses for the determination of microstructure in metals,” J. Acoust. Soc. Am., vol. 32, no. 11, pp. 1499–1499, 1960.
[23]. S. Saha, “Fuzzy logic analysis of physiological data for hypovolemia class level detection,” in 2014 International Conference on Informatics, Electronics & Vision (ICIEV), 2014.
[24]. M. Elgendi, V. Galli, C. Ahmadizadeh, and C. Menon, “Dataset of psychological scales and physiological signals collected for anxiety assessment using a portable device,” Data (Basel), vol. 7, no. 9, p. 132, 2022.
[25]. I. Lee, “An analysis of data breaches in the U.S. healthcare industry: diversity, trends, and risk profiling,” Inf. Secur. J. Glob. Perspect., vol. 31, no. 3, pp. 346–358, 2022.
[26]. L. Resnyansky, “Conceptual frameworks for social and cultural Big Data analytics: Answering the epistemological challenge,” Big Data Soc., vol. 6, no. 1, p. 205395171882381, 2019.
[27]. A. Patrizio, “IDC: Expect 175 zettabytes of data worldwide by 2025,” Network World, 03-Dec-2018. [Online]. Available: https://www.networkworld.com/article/3325397/idc-expect-175-zettabytes-of-data-worldwide-by-2025.html. [Accessed: 07-Dec-2022].
[28]. D. K. Dutta, “Path dependence, VRIN resource endowments, and managers: Towards an integration of resource-based theory and upper echelons theory,” J. Bus. Theory Pr., vol. 1, no. 1, p. 109, 2013.
[29]. T.-D. Bui, J.-W. Tseng, T. P. T. Tran, H. M. Ha, M.-L. Tseng, and M. K. Lim, “Circular business strategy challenges and opportunities for Industry 4.0: A social media data-driven analysis,” Bus. Strat. Environ., 2022.
[30]. M. S. Rahman and H. Reza, “A systematic review towards big data analytics in social media,” Big Data Min. Anal., vol. 5, no. 3, pp. 228–244, 2022.
[31]. M. Jeusfeld, I. Morshedzadeh, and A. H. C. Ng, “Managing manufacturing data and information in product lifecycle management systems considering changes and revisions,” Int. J. Prod. Lifecycle Manag., vol. 13, no. 3, p. 244, 2021.
[32]. O. Noby, A. Rady, and S. Gaber Abd Eljalil, “The influence of big data analytics on hotel performance efficiency in Egyptian hotels,” مجلة كلية السياحة والفنادق. جامعة المنصورة, vol. 11, no. 2, pp. 535–567, 2022.
[33]. S. E. Bibri and J. Krogstie, “The core enabling technologies of big data analytics and context-aware computing for smart sustainable cities: a review and synthesis,” J. Big Data, vol. 4, no. 1, 2017.
[34]. R. A. Enikeeva, “Using of the new technologies for producing medical oxygen and applications of it in military medical organizations,” Bulletin of the Russian Military Medical Academy, vol. 19, no. 3, pp. 81–83, 2017.
[35]. L. Juma and S. Kilani, “Adoption enablers of big data analytics in supply chain management practices: the moderating role of innovation culture,” Uncertain Supply Chain Manag., vol. 10, no. 3, pp. 711–720, 2022.
[36]. Zhang J., “Application of supply chain integration management of medical consumables,” Zhongguo Yi Liao Qi Xie Za Zhi, vol. 37, no. 4, pp. 304–307, 2013.
[37]. S. Alshahrani, S. Rahman, and C. Chan, “Hospital-supplier integration and hospital performance: evidence from Saudi Arabia,” Int. J. Logist. Manag., vol. 29, no. 1, pp. 22–45, 2018.
[38]. J. K. Smith and A. Ashcraft, “Moving the needle on patient cancellations through mobile integration: A hospital-based quality improvement project: A hospital-based quality improvement project,” Gastroenterol. Nurs., vol. 45, no. 6, pp. 419–427, 2022.
[39]. M. Fallatah, “Offshoring and organizational innovation: The moderating roles of absorptive capacity and inter-functional integration,” Int. Bus. Res., vol. 12, no. 10, p. 57, 2019.
[40]. K. Batko and A. Ślęzak, “The use of Big Data Analytics in healthcare,” J. Big Data, vol. 9, no. 1, p. 3, 2022.