Computer Engineering and Applications Vol. 9, No. 1, February 2020

A Deep Learning Approach to Integrate Medical Big Data for Improving Health Services in Indonesia

Bambang Tutuko, Siti Nurmaini*, Muhammad Naufal Rachmatullah, Annisa Darmawahyuni, Firdaus Firdaus

Intelligent System Research Group, Faculty of Computer Science,
Universitas Sriwijaya, Palembang, Indonesia
**Correspondence: siti_nurmaini@unsri.ac.id*

ABSTRACT

Medical Informatics to support health services in Indonesia is proposed in this paper. The focuses of paper to the analysis of Big Data for health care purposes with the aim of improving and developing clinical decision support systems (CDSS) or assessing medical data both for quality assurance and accessibility of health services. Electronic health records (EHR) are very rich in medical data sourced from patient. All the data can be aggregated to produce information, which includes medical history details such as, diagnostic tests, medicines and treatment plans, immunization records, allergies, radiological images, multivariate sensors device, laboratories, and test results. All the information will provide a valuable understanding of disease management system. In Indonesia country, with many rural areas with limited doctor it is an important case to investigate. Data mining about large-scale individuals and populations through EHRs can be combined with mobile networks and social media to inform about health and public policy. To support this research, many researchers have been applied the Deep Learning (DL) approach in data-mining problems related to health informatics. However, in practice, the use of DL is still questionable due to achieve optimal performance, relatively large data and resources are needed, given there are other learning algorithms that are relatively fast but produce close performance with fewer resources and parameterization, and have a better interpretability. In this paper, the advantage of Deep Learning to design medical informatics is described, due to such an approach is needed to make a good CDSS of health services.

Keywords: Deep Learning, Medical Big Data, Decision Support System

1. INTRODUCTION

The use of electronic technologies for patient health medical information has produced a large amount of data from various populations, cell types, and disorders. Currently, it called Big data. Electronic health records (EHRs) one approach to support the technology [1][2]. In addition, the use of EHR can generate a lot of information about one's health digitally, due to the increasing popularity use of wearable devices/technologies and web applications in collecting patient health information. In the next few years, this type of data is widely accessible and even publicly available, allowing researchers everywhere to search for markers of certain biological processes or provide an appropriate therapy of disease for patients.

However, due to the increasing amount of data along with time and energy limitation of clinical experts in analyzing large amounts of data, a system is needed to help this process. Clinical decision support system (CDSS) is one approach that can be used to produce models or concepts from integration and analysis among data so that the information provided is more accurate and can help clinical experts in making decisions.[2]. CDSS based on open data initiatives which aim to serve as a hub for existing and future research data like cancer [3], biomolecular analysis [4], cardiovascular disease [2], type 2 diabetes and others will dramatically expand the public store of data available for both mining and integrated meta-analyses [5]. In addition, developments like such data have propelled medical informatics research into the era of big data.

One of the advantages of data mining techniques is hypothesis-free nature, so that the big data approach can be used to gain a global perspective that complements mechanistic studies in experimental biology and allows detection of a high-level feature that is unable to perceive. Such an approach can help clinical experts clarify the process of pathogenesis and classification of complex diseases appropriately, which is caused by several factors that are mutually involved [6]. Deep learning has also been recognized as an initial step toward precision medicine [7]. However, the process of integrating molecular science into clinical medicine requires great progress before big data can make a maximum contribution to health care. [8]. In order to address this challenge, the application of big data is used to develop statistical techniques as well as computational infrastructure that is able to integrate and analyze large and heterogeneous data sets, so that ultimately the information presented has clinically relevant insights that are able to handle unseen diagnostic and therapeutic needs in all broad medical disciplines [7].

The application of big data analytics in health care has and produces positive things for the world of health. Big data refers to large amounts of information consolidated analyzed by certain technologies and applied to the health sector, it will use certain health data from a population and potentially help prevent epidemics and cure disease [8]. In addition, big data can also play a role in obtaining patient's personal data and will reduce hospital errors in providing medication [7]. This usually occurs because the hospital lacks data from these patients, with big data this will not happen [8]. With big data, medical researchers can use a large amount of data on the initial recovery rate plan to find a high success rate in the real world and to increase the chances of recovery of patients with cancer. For example, with the big data, researchers can examine tumor samples in biobank, which are linked to records of patient care obtained through big data. In addition, researchers can see things like how certain mutations and cancer proteins will find trends that will lead to better patient outcomes [5][6].

However, to make this kind of insight more available, a database of patients from various patients in hospitals, universities, non-profit institutions needs to be connected, therefore it can be available for a broader health-database. For example, in Indonesia, cancer patients are expected to increase by about seven times by 2030 due to a lack of hospital services and a lack of personal data for each patient, with this the Indonesian hospital is in dire need of big data analytic that is useful for processing large amount. If talking about large amounts of data. Integration of data in health services is a very important thing, especially to provide the best and integrated health services for the community. This article discusses what are

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

important data sources in the health sector, and what are the important use cases to implement.

2. INDUSTRY KNOWLEDGE

The knowledge management system for healthcare is essential for improving the services and providing the best possible treatment [7][8]. It implies the combination of internal knowledge and externally generated information. The industry is growing rapidly, new technologies are created that require the collection of effective storage and distribution of various facts. Big data, with deep learning (DL) computation as a tool to guarantee the integration of various knowledge so that it can help health care providers to achieve optimal results. Such solutions reshape the medicine industry, uncover new insights, and turn brave ideas into reality [4]. There are four parameters to be defined explicitly including volume, variety, velocity and veracity when the technology is utilized to support all process and produce a good decision [4]. As presented in Figure 1 CDSS framework for improving health services is proposed to describe all aspect in the process. In the framework, continually *learning healthcare* infrastructure based on big data technology in the context of EHRs. The massive number of patient encounters results in high amounts of stored data. Transforming clinical data into knowledge to improve patient care. The process of knowledge management from data to knowledge and from knowledge to act as a cycle on a big data enabled architecture.

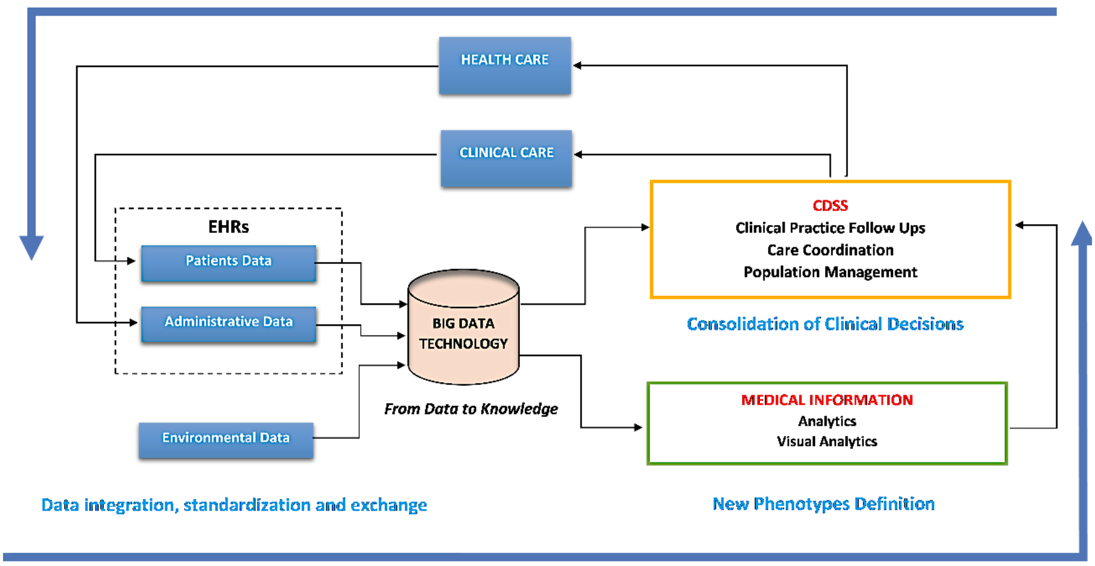


FIGURE 1. CDSS framework

Big data analysis using DL, has a great advantage for assimilating and evaluating complex and plentiful health-care data. In addition, other advantages include optimal flexibility and scalability when compared to traditional bio-statistical methods, which make it usable for many tasks, such as risk stratification, diagnosis and classification, and survival prediction [9]. Another advantage of the deep learning method is its ability to analyze various of data types (such as population data, laboratory findings, medical images or doctor's notes) and combine them to risk factor prediction of a disease, diagnosis, prognosis and appropriate health-care treatment for patients [10]. Apart from this advantage, the application of deep learning in the delivery of health services also provides several challenges such as data processing, design and training of models and system improvements related to clinical problems. In addition, to maximize the use of digital devices in health-care there are some limitations that must be faced and resolved, such as the application of appropriate clinical systems and ethics in the provision of health services. Furthermore, other important ethical considerations include medico-legal applications, doctors' understanding of DL devices, and the privacy and security of patient data.

3. DEEP LEARNING APPROACH

Deep learning (DL) mimics the operation of the human brain using multiple layers of conventional neural network which can learn a specific pattern directly from data and generate an automated prediction from input. DL is a basic assumption of Artificial Intelligence (AI) that increasingly interested to support better patient care while improving efficiencies and reducing costs in healthcare organizations [10]. DL with AI has developed rapidly for healthcare. It is potential to analyze data with a high precision. Over a relatively short period of time, some stakeholders, providers, payers with a dizzying array of technologies and strategies have able to show AI in outstanding performance. DL is a subset of new machine learning technique.

Technically, DL methods can be constructed by composing simple but nonlinear layers which transform the raw input representation into a higher and slightly more abstract level [11]. Large complex functions can be learned with the composition of enough such transformation. As it is known, DL is representation-learning based with multiple levels of representation [11][12][13][14]. Higher layers of representation amplify the input aspect which is important to remove irrelevant information and to increase feature discriminant. The DL processing pipeline to support CDSS is presented in Figure 2. The process can be divided into six phases such as capturing, storing, sharing, analyzing, searching and decision support system. The place of DL process in phase analyzing, because all raw data will be extracted become knowledge. Before the action is produced some tasks like classification, clustering and regression must be done.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

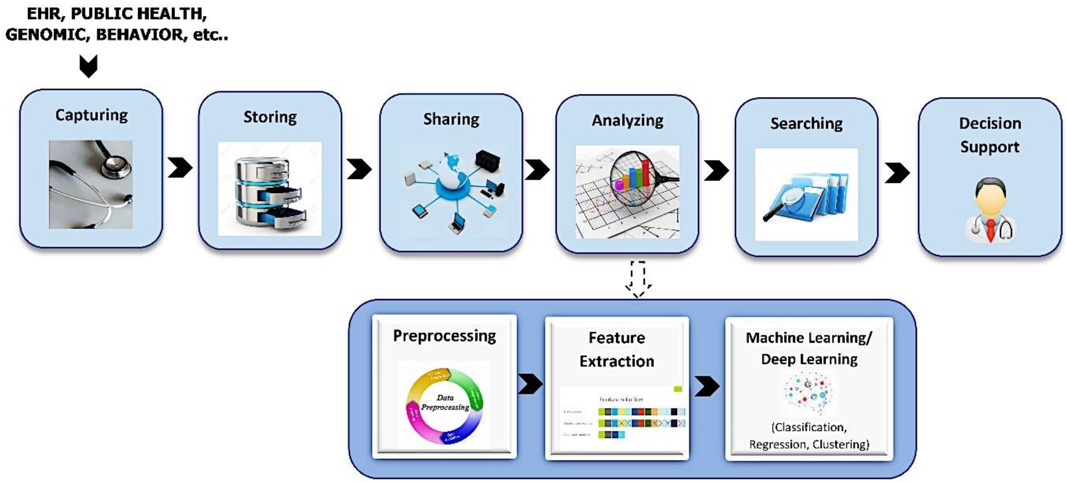


FIGURE 2. Deep learning processing pipeline for DCSS

Some of DL approach including deep neural networks (DNNs), convolutional neural network (CNN) and recurrent neural networks (RNNs) have been proposed in CDSS. In the process of analysis and diagnosis of biomedical imagery, CNN has particularly well suited to detect and to classify MRI images and x-rays [15]. The DL approach for natural language processing and speech processing for health care industry has also been widely used such as dictating documentation and translating speech-to-text [10]. In addition, the discovery of new medicines and precision prescribing have also been the focus of research by DL developers. In Table 1. some research by using DL approach in medical have been developed with good results. However, there are some limitations of DL such as [11]; (i) getting high diagnostic accuracy on new cases from available data; (ii) dealing with missing and noisy data; (iii) reducing the number of training and testing required for diagnosis, and (iv) minimizing time complexity of the whole process from acquisition to decision making.

TABLE 1.
Deep learning research for medical data services

Method	Research Focus	Data	References
CNN	Volumetric medical image segmentation using V-Net Architecture	Clinical Imaging	[17]
	Arrythmias automated detection for different interval of tachycardia segment using CNN	Electronic Health Records	[18]
	Deep convolutional neural network for detecting atrial fibrillation in ECG.	Electronic Health Records	[19]
	Water–fat separation and parameter mapping for cardiac MRI using	Clinical	[15]

	convolutional neural network	Imaging	
RNN	Medical event detection in electronic health using bidirectional RNN	Electronic Health Records	[20]
	Risk predictions of dynamic mortality in pediatric critical care using RNN	Electronic Health Records	[21]
	Combination of RNN and CNN for multiclass myocardial infraction classification in portable ECG devices	Electronic Health Records	[22]
AE	The application of improved DAE for ECG signal enhancement	Electronic Health Records	[23]
	Variational autoencoder for DNA exploration of Lung cancer sample.	Genomic	[24]
	Combination of biologic feature extraction and unsupervised variational autoencoder for breast tumor genome-wide DNA methylation data	Genomic	[25]
	Denoising autoencoder for association analysis of Deep Genomic Feature	Genomic	[26]
DBN	Classification of ECG signal quality with DBN	Electronic Health Records	[27]
	Name Entity Recognition (NER) of Chinese electronic medical record using improved DBN Model	Electronic Health Records	[28]
	Fault detection of roll bearing with local linear embedding and DBN	Genomic	[29]
RBM	Deep Learning Approach for manifold learning of brain MRIs	Clinical Imaging	[30]
	PPG-based identification using adaptive deep learning approach	Mobile	[31]
	Health-care decision making for Electronic medical record (ERM) using deep learning.	Electronic Health Records	[10]

Healthcare and medicine stand to benefit immensely from DL. Its ability to generate the sheer volume of data, learn arbitrarily complex relationships, and incorporate existing knowledge that being generated as well as the increasing proliferation of medical devices and digital record systems. However, the flexibility of a deep learning algorithm creates another challenge, that is creating a high level of trust for users, especially doctors and regulators, which requires strong reasons

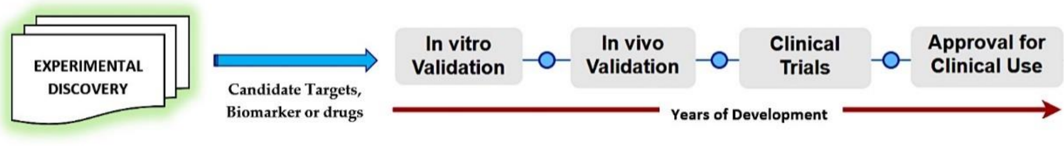
ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

and considerations from various aspects for making decisions. We argue that the challenge can be overcome by training a deep model so that the results can be trustworthy to make predictions.

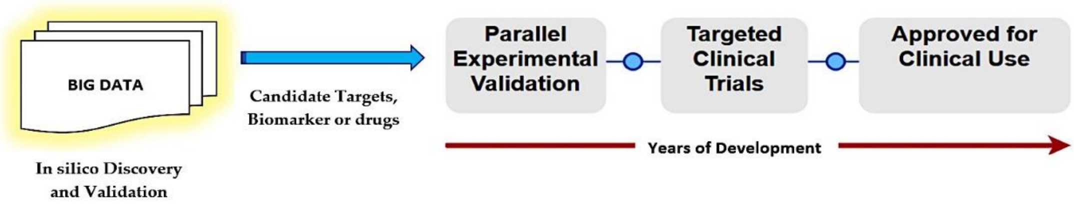
4. BIG DATA INTEGRATION SYSTEM FOR E-HEALTH

Big data analytics has opened a new era for improving services and completing health problems. Many health care institutions and various countries have test the use of big data analytics, and successfully solve basic problems within administering health insurance [1][4][8]. In the UK, big data utilization has been developed by NHS to become more complex in a machine learning that involves the participation of patients, practitioners, researchers, health facilities, and policymakers [32]. In South Korea, NHIS has collected at least 3.4 trillion data including names, addresses, income, assets, medical history, and other information obtained from tax service offices, ministries, service providers for retirees, and institutions that oversee the welfare of workers [33]. Meanwhile, in Ghana, NHIS Ghana utilizes big data to control tuberculosis. The use of big data includes measuring the effectiveness of interventions clinically, drug procurement management, disease, and clinical audit [34]. However, there are a number of factors that need to be carefully considered when using big data analytics for health services. It must be systematically prepared to produce good quality. The results must also be reviewed with varied expertise to create a comprehensive unity of understanding. Big data utilization has been developed to become more complex in a machine learning that involves the participation of patients, practitioners, researchers, health facilities, and policymakers.

Especially in Indonesia, the utilization of big data is important to measure clinically the effectiveness of interventions, drug procurement management, disease burden, clinical audit, and others. This case study illustrates the revolutionary benefits of telemedicine as the latest innovation in health services. With telemedicine technology, the limitations of medical personnel due to the uneven distribution of health workers and Doctor throughout Indonesia including the limited number of specialists in a number of cities can be overcome. Such telemedicine technology allows patients to get consults from doctors who practice in other cities with more health providers. This telemedicine technology allows health practitioners to provide diagnosis and treatment of patients remotely, without reducing the quality of professional services available to health professionals (see Figure 3). The use of secure connectivity and data servers allows patients' medical records to be accessed and analyzed before doctors meet directly with patients. The benefits of this telemedicine can also be directly felt by health care providers and patients as recipients of these services.



(a) Conventional approach



(b) Big data approach

FIGURE 3. Comparison of big data-driven approach

Currently, the access of Indonesian people to health services has become increasingly widespread. Based on a report from the Indonesian Minister of Health in 2014, Jakarta had the highest number of general practitioners. On average 16,000 doctors are placed in major cities in Indonesia and serve more than 10 million residents. The region with the least number of doctors in West Sulawesi, with more than 100 doctors serving around 1.3 million residents. There are so many ways to use telemedicine. The health industry must be able to store and manage large amounts of data for their patients, especially with the rapid growth in information supply. This includes the ability to process large amounts of data from various sources, such as wearable devices. Such as real-time monitoring of the patient's health conditions through wearable devices it allows doctors to provide a much better diagnosis. With this innovation, fatal medical conditions can be prevented and it is no longer impossible to deal with. In fact, this technology has been used to detect some diseases through a combination of connectivity, cloud data center, and Big Data analysis. Connectivity allows medical researchers to collect more data from patients

How do we do it? First of all, the health industry in Indonesia needs to digitize patients' medical records. This practice has actually been carried out by nurses or health workers throughout the archipelago. But keep in mind that existing data must be stored in a standard protocol to ensure interoperability of the data. Data security also needs to be guaranteed by providing strong data and backup storage using online backup or cloud storage. In addition, safe computing practices need to be maintained by regularly archiving data systems whether it is daytime or nighttime. Currently, the big data revolution in the health sector continues. Imagine how much data is generated and used by someone for their medical records? How can we manage all the data and information very much and take all the benefits offered? We can solve critical issues in this health service by utilizing IoT to collect data digitally and use connectivity to store data in the cloud data center while doing big data analytics. Now is the time to develop the health sector with appropriate data management facilities in Indonesia.

5. CONCLUSION

There have been significant demonstrations of the potential utility of AI approaches based on DL in medical diagnostics. The applications of deep learning

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

algorithms in clinical settings provide the potential of consistently delivering high-quality results. On the clinical side, digital image analysis on medical data will be the focus of research in the near future, due to the fact that deep learning has good initial steps in many high-value applications. However, the use of DL in clinical applications is only one small part of how DL changes the functionality of the health care system. This approach requires many integrated technologies such as chat-bot systems, m-Health applications, and virtual personal assistants such as Alexa, Siri or Google Assistant to support the system in helping consumers. The use of various types of technology can radically change the way patients interact with the health care system, offering home-based chronic disease management programs access to basic triage, and new ways to complete administrative tasks. However, the use of big data in the health sector with DL in Indonesia is still very minimal due to lack of supporting technology, even though it is estimated that in 2030 Indonesia will experience an increase in cancer patients by 70%. In conclusion, with the combination of several advanced methods and devices, improvement in processing power, and growing interest in developing innovative methods in preventing, predicting, and reducing health care costs, is likely to be a good initial step for deep learning in the health. With the increasing number of promising use cases, which are supported by strong investment from the industrial sector, and the growing amount of data supported by an up-to-date clinical analysis, DL will undoubtedly play a leading role in an attempt to provide high-quality care services to consumers in several coming decades.

REFERENCES

[1] D. Blumenthal and M. Tavenner, "The 'meaningful use' regulation for electronic health records," *N. Engl. J. Med.*, vol. 363, no. 6, pp. 501–504, 2010.

[2] J. L. Hall *et al.*, "Merging electronic health record data and genomics for cardiovascular research: a science advisory from the American Heart Association," *Circ. Cardiovasc. Genet.*, vol. 9, no. 2, pp. 193–202, 2016.

[3] E. Cerami *et al.*, "The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data." AACR, 2012.

[4] D. W. Bates, S. Saria, L. Ohno-Machado, A. Shah, and G. Escobar, "Big data in health care: using analytics to identify and manage high-risk and high-cost patients," *Health Aff.*, vol. 33, no. 7, pp. 1123–1131, 2014.

[5] K. Häyrinen, K. Saranto, and P. Nykänen, "Definition, structure, content, use and impacts of electronic health records: a review of the research literature," *Int. J. Med. Inform.*, vol. 77, no. 5, pp. 291–304, 2008.

[6] B. J. Wells, K. M. Chagin, A. S. Nowacki, and M. W. Kattan, "Strategies for handling missing data in electronic health record derived data," *Egems*, vol. 1, no. 3, 2013.

[7] S. Witz *et al.*, "Using informatics to improve medical device safety and systems thinking," *Biomed. Instrum. Technol.*, vol. 48, no. s2, pp. 38–43,

2014.
[8] M. Adibuzzaman, K. Musselman, A. Johnson, P. Brown, Z. Pitluk, and A. Grama, “Closing the data loop: An integrated open access analysis platform for the mimic database,” in *2016 Computing in Cardiology Conference (CinC)*, 2016, pp. 137–140.
[9] R. Fang, S. Pouyanfar, Y. Yang, S.-C. Chen, and S. S. Iyengar, “Computational health informatics in the big data age: a survey,” *ACM Comput. Surv.*, vol. 49, no. 1, p. 12, 2016.
[10] Z. Liang, G. Zhang, J. X. Huang, and Q. V. Hu, “Deep learning for healthcare decision making with EMRs,” in *2014 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2014, pp. 556–559.
[11] A. G. Siti Nurmaini, Radiyati Umi Partan, Muhammad Naufal Rachmatullah, “Cardiac Arrhythmias Classification Using Deep Neural Networks and Principal Component Analysis Algorithm,” *Int. J. Adv. Soft Comput. Its Appl.*, vol. 10, no. 2, pp. 14–32, 2018.
[12] A. Darmawahyuni, S. Nurmaini, W. Caesarendra, V. Bhayyu, M. N. Rachmatullah, and others, “Deep Learning with a Recurrent Network Structure in the Sequence Modeling of Imbalanced Data for ECG-Rhythm Classifier,” *Algorithms*, vol. 12, no. 6, p. 118, 2019.
[13] A. Darmawahyuni, “Coronary Heart Disease Interpretation Based on Deep Neural Network,” *Comput. Eng. Appl. J.*, vol. 8, no. 1, 2019.
[14] Siti Nurmaini, Radiyati Umi Partan, Muhammad Naufal Rachmatullah, “Deep Neural Networks Classifiers on The Electrocardiogram Signal for Intelligent Interpretation System,” *Sriwij. Int. Conf. Med. Sci.*, 2018.
[15] J. W. Goldfarb, J. Craft, and J. J. Cao, “Water--fat separation and parameter mapping in cardiac MRI via deep learning with a convolutional neural network,” *J. Magn. Reson. Imaging*, 2019.
[17] F. Milletari, N. Navab, and S.-A. Ahmadi, “V-net: Fully convolutional neural networks for volumetric medical image segmentation,” in *2016 Fourth International Conference on 3D Vision (3DV)*, 2016, pp. 565–571.
[18] U. R. Acharya, H. Fujita, O. S. Lih, Y. Hagiwara, J. H. Tan, and M. Adam, “Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network,” *Inf. Sci. (Ny).*, vol. 405, pp. 81–90, 2017.
[19] Y. Xia, N. Wulan, K. Wang, and H. Zhang, “Detecting atrial fibrillation by deep convolutional neural networks,” *Comput. Biol. Med.*, vol. 93, pp. 84–92, 2018.
[20] A. N. Jagannatha and H. Yu, “Bidirectional RNN for medical event detection in electronic health records,” in *Proceedings of the conference. Association for Computational Linguistics. North American Chapter. Meeting*, 2016, vol. 2016, p. 473.
[21] M. Aczon *et al.*, “Dynamic mortality risk predictions in pediatric critical care

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)

using recurrent neural networks," *arXiv Prepr. arXiv1701.06675*, 2017.

[22] H. W. Lui and K. L. Chow, "Multiclass classification of myocardial infarction with convolutional and recurrent neural networks for portable ECG devices," *Informatics Med. Unlocked*, vol. 13, pp. 26–33, 2018.

[23] P. Xiong, H. Wang, M. Liu, S. Zhou, Z. Hou, and X. Liu, "ECG signal enhancement based on improved denoising auto-encoder," *Eng. Appl. Artif. Intell.*, vol. 52, pp. 194–202, 2016.

[24] Z. Wang and Y. Wang, "Exploring DNA Methylation Data of Lung Cancer Samples with Variational Autoencoders," in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2018, pp. 1286–1289.

[25] A. J. Titus, O. M. Wilkins, C. A. Bobak, and B. C. Christensen, "Unsupervised deep learning with variational autoencoders applied to breast tumor genome-wide DNA methylation data with biologic feature extraction," *bioRxiv*, p. 433763, 2018.

[26] Q. Liu and P. Hu, "Association Analysis of Deep Genomic Features Extracted by Denoising Autoencoders in Breast Cancer," *Cancers (Basel).*, vol. 11, no. 4, p. 494, 2019.

[27] B. Taji, A. D. C. Chan, and S. Shirmohammadi, "Classifying measured electrocardiogram signal quality using deep belief networks," in *2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*, 2017, pp. 1–6.

[28] W. Li *et al.*, "Improved deep belief network model and its application in named entity recognition of Chinese electronic medical records," in *2018 IEEE 3rd International Conference on Big Data Analysis (ICBDA)*, 2018, pp. 356–360.

[29] H. Shao, H. Jiang, X. Li, and T. Liang, "Rolling bearing fault detection using continuous deep belief network with locally linear embedding," *Comput. Ind.*, vol. 96, pp. 27–39, 2018.

[30] T. Brosch, R. Tam, A. D. N. Initiative, and others, "Manifold learning of brain MRIs by deep learning," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2013, pp. 633–640.

[31] V. Jindal, J. Birjandtalab, M. B. Pouyan, and M. Nourani, "An adaptive deep learning approach for PPG-based identification," in *2016 38th Annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, 2016, pp. 6401–6404.

[32] C. Mockford, S. Staniszewska, F. Griffiths, and S. Herron-Marx, "The impact of patient and public involvement on UK NHS health care: a systematic review," *Int. J. Qual. Heal. care*, vol. 24, no. 1, pp. 28–38, 2011.

[33] J. Lee, J. S. Lee, S.-H. Park, S. A. Shin, and K. Kim, "Cohort profile: The national health insurance service--national sample cohort (NHIS-NSC), South Korea," *Int. J. Epidemiol.*, vol. 46, no. 2, pp. e15--e15, 2016.

[34] N. Sarpong *et al.*, "National health insurance coverage and socio-economic status in a rural district of Ghana," *Trop. Med. Int. Heal.*, vol. 15, no. 2, pp. 191–197, 2010.

ISSN: 2252-4274 (Print)
ISSN: 2252-5459 (Online)