ISSN: 2321-2152

IJMECE

International Journal of modern electronics and communication engineering

E-Mail
editor.ijmece@gmail.com
editor@ijmece.com

www.ijmece.com

ISSN2321-2152
www.ijmece .com
Vol 12 Issue1 Feb 2024

PROTECTING PERSONAL HEALTHCARE RECORD USING BLOCKCHAIN & FEDERATED LEARNING TECHNOLOGIES

RAVI KIRAN KUMAR TERA 1, MOHURALA SHREYAS GOUD 2, MOHAMMED AWAIS 3, DOULAPURAM MOHAN 4, TEKMAL PRASHANTH 5,

ABSTRACT

Healthcare institutions are progressively integrating artificial intelligence (AI) into their operations. The extraordinary potential of AI is restricted by insufficient medical data for AI model training and adversarial attacks wherein attackers perturb the dataset by adding some noise to it, which leads to the malfunctioning of the AI models, and a lack of trust caused by the opaque operational approach it employs. This Systematic Literature Review (SLR) is a state-of-the-art survey of the research on blockchain technology for securing AI-integrated healthcare applications. The most relevant articles from the Scopus and Web of Science (WoS) databases were identified using the PRISMA model. Most of the existing literature is about protecting the healthcare data used by AI-based healthcare systems using blockchain technology, but the modality of data (text, images, audio, and sound) was not specifically mentioned. Information on protecting the training phase and model deployment for AI-based healthcare systems considering the variations in feature extraction based on the modality of data was also not clearly specified. Hence, the three subfields of AI, namely, natural language processing (NLP), computer vision, and acoustic AI are further studied to identify security loopholes in its implementation pipeline. The three phases, namely the dataset, the training phase, and the trained models need to be protected from adversaries to avoid malfunctioning of the deployed AI models. The nature of the data processed by NLP, computer vision, and acoustic AI, underlying deep neural network (DNN) architectures, the complexity of attacks, and the perceivability of attacks by humans are analyzed to identify the need for security. A blockchain solution for AI-based healthcare systems is synthesized based on the findings that have demonstrated the distinctive technological features of blockchains. It offers a solution for the privacy and security issues encountered by NLP, computer vision, and acoustic AI to boost the widespread adoption of AI applications in healthcare.

INTRODUCTION The healthcare sector needs technology handholding from infectious diseases to cancer disease management. There are countless ways to use technologies to provide more accurate, reliable, and effective treatments. These treatments can be precise at the right time in a clinical decision. Artificial Intelligence uses a computer program with precise commands to execute functions that usually require human intelligence. Algorithms are coded programming rules. Machine Learning is a method of the constant improvement of an algorithm. The improvement process utilizes vast volumes of data and is performed dynamically, enabling the algorithm to adjust and improve the accuracy of the said Artificial Intelligence. AI can understand and interpret language, identify objects, detect sounds, and learn patterns to execute problem-solving operations. In this review, insight is provided into three main domains of artificial

ASSISTANT PROFESSOR 1, UG SCHOLAR 2,3,4&5
DEPARTMENT OF CSE, MNR COLLEGE OF ENGINEERING AND TECHNOLOGY, MOHD.SHAPUR, TELANGANA 502285

ISSN2321-2152
www.ijmece .com
Vol 12 Issue1 Feb 2024

intelligence (AI), namely, Natural Language Processing (NLP), computer vision (CV), and acoustic AI, and their specific challenges in healthcare, as shown in Figure 1. The primary objective of Natural Language Processing for computers is to comprehend texts and languages as grasped by humans. Computer systems can interpret, deduce, summarize, translate, and synthesize exact text and language. A vast amount of textual data is generated in healthcare systems in the form of clinical reports, lab reports, handwritten notes, and other documents like admission, discharge notes, and many more. The overweighing task for clinical experts is to handle and manually analyze this enormous data. The primary tasks that can be driven through NLP are extracting important facts from text, classification of information, and opinion mining. NLP helps the analysis and conversion of these growing data to a manageable computer format. It can help to assist clinical decisions, the identification of critical patients, and the classification of diseases and disorders.

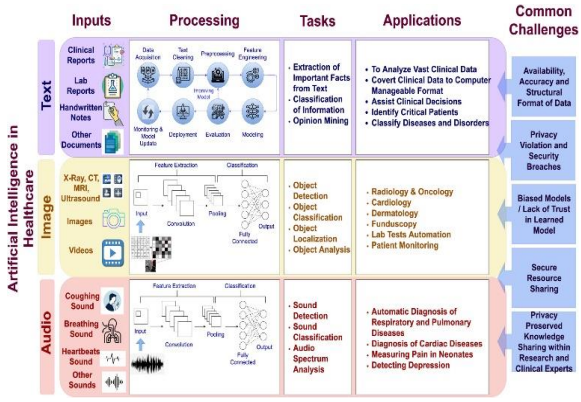

Fig: Artificial intelligence in healthcare.

The rapidly growing area of computer vision is concerned with training computers to mimic human vision and understand the items opposite to them. Computer vision fixes this by leveraging Artificial Intelligence algorithms, which aid in the analysis of images. X-ray, Computerized Tomography (CT), Magnetic Resonance Imaging (MRI), Fluorescence-MRI, ultrasound images and videos have been proven to be among the most vital tools in deciding on the diagnosis for a patient.[1] Computer vision can promote remote patient monitoring, automated diagnosis, and automated lab reports through different tasks like object detection, classification, localization, and analysis from images or videos. It can promote the emergence of numerous applications that can be lifesaving for patients in radiology, oncology, cardiology, dermatology, and fundoscopy.

Certain sounds such as coughing, breathing, heartbeats, and crying play a major role in diagnosing respiratory, pulmonary, and cardiac diseases, as well as pain in neonates and detecting depression in human beings. AI assists in the automation of these diagnoses by sound detection, performing classification, and their analysis through the audio spectrum. Various state-of-the-art deep learning algorithms are available for audio signal processing, which can be helpful in the healthcare industry.

There are certain common challenges faced by AI models for their wider adoption in the healthcare industry, which is mentioned in Figure 1. When AI models are trained on sufficiently large datasets, they can work with precision. Therefore, the availability of vast, accurate, and trusted datasets for the training is one of the major challenges. This is achievable by aggregating data from different resources. But the data should be protected from privacy violations and security breaches as the organizations continue to collect, store, and transport the sensitive health vitals of the individual. It is difficult to identify biased models as AI models are black boxes in nature. There must be the provenance of prediction or classification resulting in specific healthcare input to overcome the lack of trust in the learned model. Human lives are at stake if the wrong treatment is followed based on the AI results. There should be secure resource sharing to overcome the threat of rogue devices. Knowledge sharing among researchers and clinical experts may be prone to information privacy issues. Hence, a proven strategy is a prerequisite to overcoming these challenges and establishing the dominance of AI over the healthcare industry in the future.[2]

1.1 Background of study

Blockchain technology can address the challenges faced by AI in several ways. A blockchain is a distributed ledger with transactions that are replicated throughout the Blockchain ecosystem. The security and privacy feature of Blockchain is enriched with the cryptographic linkage of information in chronicle order, consensus protocol within the network, and smart contracts. Moreover, it builds strong trust among the users. Hence it can also establish trust, organize data, and allow sharing of resources while supporting interoperability in AI-based healthcare.[3] This study mainly targets the applicability of Blockchain in AI-based healthcare systems taking into consideration security and privacy issues in three vertical aspects, namely natural language processing (NLP), computer vision, and acoustic AI.

Figure 2 depicts the flow of activities in the Blockchain network. Blockchain technology features a distributed ledger in the peer-to-peer network. The distributed ledger securely maintains the transactional records. This feature promotes secured Distributed Learning or Federated Learning on heterogeneous data by recording local gradients on Blockchain. Moreover, a smart contract automates the execution of a transaction in the distributed network without any third-party or centralized authority. The smart contract is an executable code available at every node, which gets triggered on transaction initialization. Smart contracts validate the transaction. Access control rules can be imposed for data access through smart contracts. User provenance is possible with a smart contract. A block is

ISSN2321-2152
www.ijmece .com
Vol 12 Issue1 Feb 2024

generated for the transactional data. The miners are responsible for committing the block in the Blockchain by using consensus algorithms which are responsible for mining the block. It makes miners solve difficult cryptographic puzzles and share their results with a group of miners. The miner who first solves the puzzle gets a chance to mine a block of the transactions into the existing chain of blocks and replicate the new chain at every node. Consensus algorithms is the proven technique for collective decision-making on the diagnosis and treatment in AI-based healthcare systems. The blocks are linked with each other cryptographically, which makes them immutable and auditable. The same copy of the ledger is replicated at all nodes in the network, henceforth it achieves the highest degree of availability and transparency. Cryptographic linkage can validate the medical data and support its tamperproof copy. There are three types of Blockchain available, namely Public Blockchain, Private Blockchain, and Consortium Blockchain. In Public Blockchain, anyone can enter the network and participate in the transaction process. In contrast, Private Blockchain restricts entry without proper authentication and verification. Consortium Blockchain combines the features of public and private Blockchain.

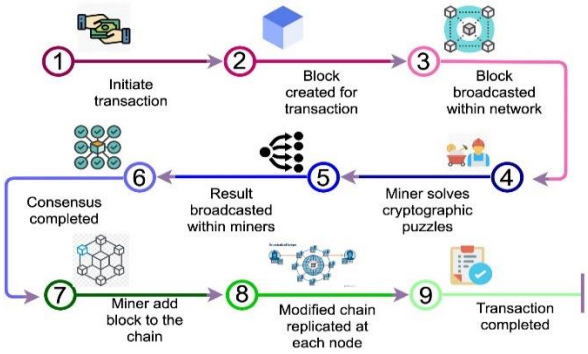

Fig: Blockchain implementation process.

EXISTING SYSTEM

The artificial intelligence-based technique already providing promising result in detecting various critical disease using medical images. Mainly the role of the AI based technique in diagnosis of various diseases using medical images is to act as a decision support system for the doctors so that the decision taking process will be faster and at the same time maintaining the precision and accuracy. One of the best techniques is deep learning that has been successful in detecting various diseases using medical images. Sekeroglu and Ozsahin proposed an approach that can detect Covid-19 based on chest X-ray using Convolutional Neural Network. The proposed method performed the classification of images in three groups such as healthy, pneumonia, and Covid-19 with an accuracy of 98.50% with a small imbalanced dataset [10]. Jain et al., proposed a method, that is used to classify two groups such as healthy group and Covid-19 group using chest X-rays. They have used 6432 samples and used various deep learning techniques for classification of these groups.

They have tried different deep learning techniques and found that Xception model provides high accuracy of 97.97%. This model was found to be effective for detecting covid-19 patients [11]. Ozturk et al., proposed an approach that uses binary classification for differentiating two group such as Covid-19 vs No-Findings and Multiclass classification for differentiating three groups such as Covid-19 vs No-Findings vs Pneumonia. They have tried YOLO object detection system using the darknet model. It was found that binary classification provides an accuracy of 98.08% and multiclass classification provides an accuracy of 87.02%. This model would able to help the radiologists for initial screening [12].

Linn and Koo identify quite simple yet very robust use of blockchain for the data storage purpose and leveraging it to the healthcare sector for storing the data of the patients. The proposed system tends to motivate the storage of the complete data starting from the data of birth to the wearable data to the lipid profile data to the MRI data everything on the patients Blockchain [14]. Liangetal mentioned that user can share data with even the healthcare providers and insurance companies for either getting the services or insurance quotes respectively. The system also employs to be a user centric system where the user has got complete rights regarding the data to be shared or not. This kind of systems can be very helpful in terms of medical research and also preservation of personal medical records which are a real matter of privacy can be very much achieved with Blockchain as an architecture [15].

FL based technique would able to help the healthcare providers because it can improve the accuracy and robustness of the AI model and that helps to make the model more generalizable so that it could be used in real time environment. At the same time this approach would able to save time and cost. The most important part is that we would able to get all the benefits from FL without leakage of information.

Disadvantages

- Our solution is not a differential privacy mechanism that modifies patient data to confuse the hospital records.

- An existing system is not implements Blockchain-Federated Learning Architecture.

PROPOSED SYSTEM

In the beginning hospitals, universities, pharma work in the isolated place and run the model isolated to each other but later all the trained model are shared to the federated server where the models

learn among each other without leaking the private information. Usually all the devices connected to the centralized federated server used to send the data to the server and inside the server all the model averaged and made to a single model. This process keeps on repeating until a high- quality model has been generated. This kind of architecture provides lot of benefits as follows:

1) Since the learning process is collaborative in nature the models become smarter and smarter which helps in producing good result at the later stage with unknown data.

2) Since the prediction is happening locally in the new model, it lowered down the latency

3) In the present approach it is not necessary to be available near the data to get more insight about the data, insight could be obtained from anywhere.

4) The most important point is that the privacy remained through the process because all the data are isolated to that particular device, only the trained models are shared to the server.

The blockchain service allows the users (researchers, clinicians, institutions, etc) to access the data by providing a small transaction fee. The non-repudiation and patient's private data are guaranteed. By using the blockchain system, we can know who is accessing the data, when the data was requested, for what reasons. The data access mechanism is described as follow:

1) The user creates a user account on the platform,

2) The user requests access to a certain resource/asset,

3) The smart contract checks if the resource is available in the ledger. If the requested data is available, the smart contract reserves it and notifies the user. The user checks the requirements to access the data and signs the contract. Once the validation is done the user will be charged according to the smart contract's rules. Accordingly, a usage token is issued to the user.

Advantages

- Privacy is most important, especially in the case of medical

imaging data. As the AI based model is getting traction these days especially in the field of medical imaging, protecting the private information is necessary before used in the real time environment.

- The healthcare records are maintaining using blockchain which are very safe and secure.

LITERATURE REVIEW

This section presents a thorough review of the recent and relevant literature related to the core areas of this review.

Blockchain for AI-based healthcare

The deep models have been extensively utilized to tackle several difficulties in medical treatments in image analysis. Deep learning often works more efficiently when trained on large volumes of data. Hospitals, diagnostic laboratories, research institutions, and patients may exchange valuable findings and work together to improve the AI model. However, they face challenges in sharing important, confidential data with third parties due to privacy and security concerns. Hence, secure data sharing becomes an obstacle in improving the quality of AI-based healthcare systems. Neelakandan et al.[41] proposed a BDL-SMDTD model for secure image transmission to provide a solution to the above-mentioned obstacles, wherein Blockchain is used to store encrypted images. Kumar et al.[42] proposed a strategy that involves sharing local models via the Blockchain network which was leveraged to collectively develop a global model for an improved prediction of lung cancer using CT scan images. As a result, the collectively updated model aids in the accurate diagnosis of patients' diseases, resulting in improved treatment and therapy. This will avoid the actual sharing of data and hence maintains the privacy of patients. The organizations will upload their data over the Interplanetary File System (IPFS) and share local gradients through smart contracts. Delegated Proof-of-Stake consensus algorithm is used to train the global model. Confidence in the data is established through a smart contract, with the Blockchain retaining the hash of the nearby gradient. Kim and Huh[43] introduced a Blockchain-based algorithm for validating the enhanced data, utilizing the HyperPOR consensus algorithm within the Blockchain framework. The HyperPOR algorithm functions by confirming the identity of the business partner. The generation block then validates and accomplishes distributed computing and adds sharding technology for protecting the Patient Health Records (PHR). Nguyen et al.[44] proposed an intrusion detection system to protect data transmission in the Cyber-Physical system for healthcare.

Most of the time, patients have little control over who can access their medical records and are ignorant of the full worth of the information they possess. Mamoshina et al.[11] presented a Blockchain and AI-based solution to speed up biomedical research to provide patients with new technologies for controlling and profiting from their personal information and incentives to undertake periodic health checkups. They have proposed Exonum as a permissioned Blockchain framework wherein patients can sell their health records using tokens. Nonetheless, this structure lacks authority once the data is transferred to regulatory entities. Jennath et al.[45] put forward a dependable hybrid AI-Blockchain framework for e-health. They employed an unchangeable distributed ledger to document the origin of individual permissions and the credibility of data origins, serving the purpose of constructing and refining the AI model.

Rahman et al.[46] used Blockchain and off-chain to safeguard from manipulation and illegal access, bringing confidence to the provenance of datasets and distributed models to protect the privacy and security of the Internet of Health Things (IoHT) data. The insecure central gradient aggregator is replaced with a secure, tamper-proof gradient mining and distributed consensus-based aggregator in the Blockchain. The edge training, trust management, and authentication of participating federated nodes, the dissemination of globally or locally trained models, and the identity of edge nodes and their contributed datasets or models are managed by Smart Contracts. This system provides the complete encryption of both, a dataset and a trained model. Puri et al.[47] implemented a decentralized healthcare framework powered by AI that accesses and authenticates Internet of Things (IoT) devices while instilling confidence and transparency in the PHR. The technique is based on AI-enabled Smart Contracts and the development of a Public Blockchain network. In addition, this framework detects potentially dangerous IoT nodes in the network. Gupta et al.[48] offer BITS, a unique intelligent TS system based on Blockchain. They provide thorough insights into the Cloud-based and Blockchain-based smart TS frameworks, emphasizing the challenges of security, dependability, confidentiality, and data management. If rogue devices start communicating erroneous local model updates, the global model's accuracy will be skewed. Here, Blockchain can assure that the local updates in Federated Learning come from trusted devices. The presence of local modifications within the Blockchain aids in additional validation of the precision of the acquired model. Polap et al.[49] introduced a Federated Learning approach that merges decentralized learning with Blockchain-driven security, offering a resolution for constructing intelligent systems using locally stored data in

a decentralized manner to enhance the security and confidentiality of the Internet of Medical Things. This approach serves as a countermeasure against model poisoning attacks. The study in Reference 50 stated a way for training a global model cooperatively utilizing Blockchain technology and Federated Learning while maintaining anonymity in detecting Covid-19 patients using CT images.

Technological advances, such as distributed learning, provide a road ahead, but they are plagued by a lack of openness, thereby reducing trust in the data utilized for analysis. To solve these challenges, Zerka et al.[51] have projected that Chained Distributed Machine Learning (C-DistriM), a novel distributed learning that blends sequential distributed learning with a Blockchain framework would be developed in medical imaging. Blockchain is used to record the immutable history of computation and protect from the threat of model poisoning. After training, it encrypted the local models and uploaded them on the cloud simultaneously, by the removal of all local copies of the model. Subsequent to this, unapproved users are prevented from accessing the cloud by the Smart Contract. Kuo et al.[52] introduced the Explorer Chain framework, which integrates two cutting-edge technologies, Online Machine Learning and a decentralized Blockchain, to develop a predictive model across multiple institutions within a distributed structure, eliminating the necessity for sharing patient-level data or relying on a central coordinating node.

Many published deep learning systems lack clarity about model validation and testing outcomes. Blockchain technology could potentially provide a suitable solution to these issues by functioning as a decentralized, secure, and reliable distributed ledger for data administration. It also offers the ability to track and ensure accountability for the reporting of testing results. Schmetterer et al.[53] implemented a Blockchain-driven AI platform to establish real-world data transmission, model transfer, and model testing across three locations in Singapore and China, demonstrating the proof of concept. The researchers aimed to develop and assess deep learning algorithms for the identification of myopic macular degeneration and extreme myopia, utilizing retinal images from diverse multiethnic populations across various countries. They leveraged a blockchain-enabled AI infrastructure that helps to have secure, persistent, and verifiable data transmission, model transfer, and transparency in the diagnostic performance of deep learning algorithms. However, it does not maintain the privacy of data.

Khan et al.[54] explored a wireless capsule endoscopy frame-based automated method for detecting stomach infections. A Blockchain-based technique is used in a convolutional neural network (CNN) model to secure the

network for the precise identification of stomach ailments such as ulcers and bleeding. Each layer comprises an additional block that keeps certain information to resist all tempering and modification attacks. Pilozzi et al.[55] state that AI technologies, particularly NLP, are effective tools for classifying the emotions and tonality of texts, like in social media posts. These approaches could be used to investigate the public perception of Alzheimer's disease. The incorporation of secure and decentralized data transfer and storage methods like Blockchain will give patients greater control over their data. It will help to relieve most of the insecurities of mistakenly revealing the personal information of a patient to an entity that may discriminate against the patient

CONCLUSION

This systematic literature review underscores the significance of Blockchain technology in enhancing the security of AI-based healthcare applications. The primary objective of this investigation was to illuminate the various avenues of attack that threaten AI-based healthcare applications. These include adversarial attacks on datasets, spoofing, backdoor/Trojan attacks, and timing side-channel attacks, all of which have the potential to endanger patients' lives. The heightened susceptibility of AI models to even minor input alterations underscore the need to address the models' aberrant behavior. The findings gleaned from this survey reveal that existing solutions aimed at countering AI attacks are often specialized, focusing on distinct attack vectors. However, a majority of these solutions are themselves AI-based, rendering them susceptible to adversarial exploits. Consequently, this survey underscores the value of Blockchain technology. By facilitating real-time data collection, distributed storage across multiple servers to avert hacking, access restriction to authorized personnel, and the preservation of current file versions with each interaction, Blockchain serves as a pivotal solution. This study introduces a Blockchain-centered approach to fortify the entirety of the AI development pipeline in healthcare, encompassing dataset creation, training phases, and post-training stages for NLP, computer vision, and acoustic AI. This proposal recognizes the potency of Blockchain technology in addressing the domain's unique challenges. It is important to note that the healthcare sector, due to its rigorous regulatory framework and aversion to risk, often adopts new technologies at a gradual pace. The lack of a standardized IT infrastructure and interoperability further compounds the challenges in this field. Future research avenues could involve the implementation of sophisticated lightweight Blockchain solutions tailored to the demands of AI-based healthcare applications. Such endeavors hold

the promise of ushering in a new era of secure and resilient healthcare technology.

REFERENCES

[1] Choe, J., Lee, S.M., Do, K.H., Lee, G., Lee, J.G., Lee, S.M. and Seo, J.B., 2019. Deep learning–based image conversion of CT reconstruction kernels improves radiomics reproducibility for pulmonary nodules or masses. Radiology, 292(2), pp.365-373.

[2] Kermany, D.S., Goldbaum, M., Cai, W., Valentim, C.C., Liang, H., Baxter, S.L., McKeown, A., Yang, G., Wu, X., Yan, F. and Dong, J., 2018. Identifying medical diagnoses and treatable diseases by image- based deep learning. Cell, 172(5), pp.1122-1131.

[3] Negassi, M., Suarez-Ibarrola, R., Hein, S., Miernik, A. and Reiterer, A., 2020. Application of artificial neural networks for automated analysis of cystoscopic images: a review of the current status and future prospects. World Journal of Urology, pp.1-10.

[4] Gulshan, V., Peng, L., Coram, M., Stumpe, M.C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J. and Kim, R., 2016. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama, 316(22), pp.2402-2410.

[5] Tandon, A., Dhir, A., Islam, N. and Mäntymäki, M., 2020. Blockchain in healthcare: A systematic literature review, synthesizing framework and future research agenda. Computers in Industry, 122, p.103290.

[6] Yang, Q., Liu, Y., Chen, T. and Tong, Y., 2019. Federated machine learning: Concept and applications. ACM Transactions on Intelligent Systems and Technology (TIST), 10(2), pp.1-19.

[7] Kairouz, P., McMahan, H.B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A.N., Bonawitz, K., Charles, Z., Cormode, G., Cummings, R. and d'Oliveira, R.G., 2019. Advances and open problems in federated learning. arXiv preprint arXiv:1912.04977.

[8] Li, W., Milletarì, F., Xu, D., Rieke, N., Hancox, J., Zhu, W., Baust, M., Cheng, Y., Ourselin, S., Cardoso, M.J. and Feng, A., 2019, October. Privacy-preserving federated brain tumour segmentation. In International Workshop on Machine Learning in Medical Imaging (pp. 133-141). Springer, Cham.

[9] Sheller, M.J., Reina, G.A., Edwards, B., Martin, J. and Bakas, S., 2018, September. Multi-institutional deep learning modeling without sharing patient data: A feasibility study on brain tumor segmentation. In International MICCAI Brainlesion Workshop (pp. 92-104). Springer, Cham.

[10] Sekeroglu, B. and Ozsahin, I., 2020. <? covid19?> Detection of COVID-
19 from Chest X-Ray Images Using Convolutional Neural Networks. SLAS TECHNOLOGY: Translating Life Sciences Innovation, p.2472630320958376.

[11] Jain, R., Gupta, M., Taneja, S. and Hemanth, D.J., 2020. Deep learning based detection and analysis of COVID-19 on chest X-ray images. Applied Intelligence, pp.1-11.

[12] Ozturk, T., Talo, M., Yildirim, E.A., Baloglu, U.B., Yildirim, O. and Acharya, U.R., 2020. Automated detection of COVID-19 cases using deep neural networks with X-ray images. Computers in Biology and Medicine, p.103792.

[13] Patel, V., 2019. A framework for secure and decentralized sharing of medical imaging data via blockchain consensus. Health informatics journal, 25(4), pp.1398-1411.

[14] Linn, L.A. and Koo, M.B., 2016. Blockchain for health data and its potential use in health it and health care related research. In ONC/NIST Use of Blockchain for Healthcare and Research Workshop. Gaithersburg, Maryland, United States: ONC/NIST (pp. 1-10).

[15] Ozturk, T., Talo, M., Yildirim, E.A., Baloglu, U.B., Yildirim, O. and Acharya, U.R., 2020. Automated detection of COVID-19 cases using deep neural networks with X-ray images. Computers in Biology and Medicine, p.103792.