International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys.13, No.1 (Apr-23)

http://dx.doi.org/10.12785/ijcds/130180

Blockchain for Optimized Pattern Recognition: Comparative Study

Hakima Rym Rahal[1], Sihem Slatnia[1], Okba Kazar[1,2] and Ezedin Barka[2]

[1]*LINFI Laboratory, Computer Science Department, University of Mohammed Khider, Biskra, Algeria*
[2]*College of Information Technology, University of the United Arab Emirates Al Ain, United Arab Emirates*

Received 8 Jul. 2022, Revised 5 Nov. 2022, Accepted 6 Feb. 2023, Published 16 Apr. 2023

Abstract: Machine Learning and specifically Deep Learning systems have received much interest recently in many fields including the domain of Pattern Recognition. However, these systems usually relay on centralized data for training, which causes vulnerability and security issues. In addition, the complexity of the learning algorithms and the fast growth of the amounts of data in the recent years is another challenge for these systems. Blockchain has the ability to solve the data security problems due to its decentralized nature. In the other hand, parallel optimizations can reduce the training time of the learning algorithms. In this paper, we'll be discovering the benefits of integrating Blockchain technology and Parallelism with the Pattern Recognition systems. We'll be discussing papers that implemented the Blockchain technology, Learning systems, and also the Parallel Optimization methods in the medicine field. Most of these papers focus on the prediction, early diagnosis, and diagnosis of certain diseases such as Cancer, COVID-19...etc. We'll make a comparative study between these works based on several criteria: Blockchain technology, Learning algorithms, Parallel optimization methods, Datasets, and the Accuracy. Furthermore, We will present in-depth analyses of these papers outlining the advantages and limits of each one.

Keywords: Blockchain, Consensus Protocol, Smart Contract, Medical Data, Machine Learning (ML), Deep Learning (DL), Parallel Optimization, Pattern Recognition.

1. INTRODUCTION

Machine learning is a technique of data analysis that automates the creation of analytical models. It's a field of artificial intelligence based on the premise that computers can learn from data, recognize patterns, and make judgments with little or no human involvement. Deep learning is a subcategory of machine learning that is a set of algorithms capable of mimicking the actions of the human brain using artificial neural networks. Giving computers the ability to reason has led to the development of this approach. The major issue with these learning algorithms is represented in the centralized data required for training. The risk of distribution, manipulation, and forgery associated with centralized data is a significant security issue, particularly in delicate industries like medicine. This problem can be solved by implementing the Blockchain technology. The Blockchain is a distributed database that records transactions between different parties and enables decentralized management. The Blockchain technology authenticates the data and federated learning. In addition to its own capabilities, it can help in handling many limitations that Machine Learning and Deep Learning-based systems have. Another challenge facing the learning systems is the complexity of those learning algorithms in addition to the growth of the size of data necessary for training which causes a temporal issue. Parallel optimization techniques can help reduce the time a system takes to reach a solution. The combination of these techniques can provide high-performing and useful Intelligent Pattern Recognition system.

The current paper focuses generally on works that make use of the Blockchain technology, the Learning algorithms, and the the Parallel optimization methods in the healthcare domain, from disease diagnosis to building Personal Health Records. This paper also presents a comprehensive review of the related literature to these three approaches to identify the existing systems' strengths and weaknesses. We also created a classification for categorizing these state of the art works according to a number of criteria: Blockchain platform, Blockchain type, Consensus protocol, Learning algorithm, Accuracy, existence of the Parallel optimization, Datasets, and the advantages and limits of each paper.

The following is how the rest of the paper is structured: Section 2 presents a small background about the main concepts of this paper. A brief discussion of the literature and the developed works is in Section 3. Section 4 contains a comparative study between the works according to several criteria. In the end, the paper comes to a close in Section

5.

2. BACKGROUND

We provide in this section some background information on blockchain technology, machine learning, deep learning, pattern recognition, and parallelism. To give readers a basic understanding, we first discuss blockchain concepts briefly, where we focus on the important fundamentals that'll be seen in this comparative study. Next, we'll introduce shortly machine learning and deep learning. We'll focus on their definitions and their important role in pattern recognition. Finally, we'll discuss parallelism and its capability to optimize the learning algorithms.

A. Blockchain Concepts

A blockchain is a collection of records that have been cryptographically connected together in a specific order. Blockchains can be classified into two categories: public blockchains and private blockchains. A public blockchain is comparable to the Internet in certain ways. It is accessible to all users of the record system. While a private blockchain restricts access to records (reading and editing) to authorized entities only. A consortium blockchain is also a type of blockchain that is a combination of public and private blockchains. Once a block is created in a blockchain, nothing in the previous blocks can be changed. The chain architecture ensures that blockchain records cannot be changed. Figure 1 represents the architecture of the blockchain:

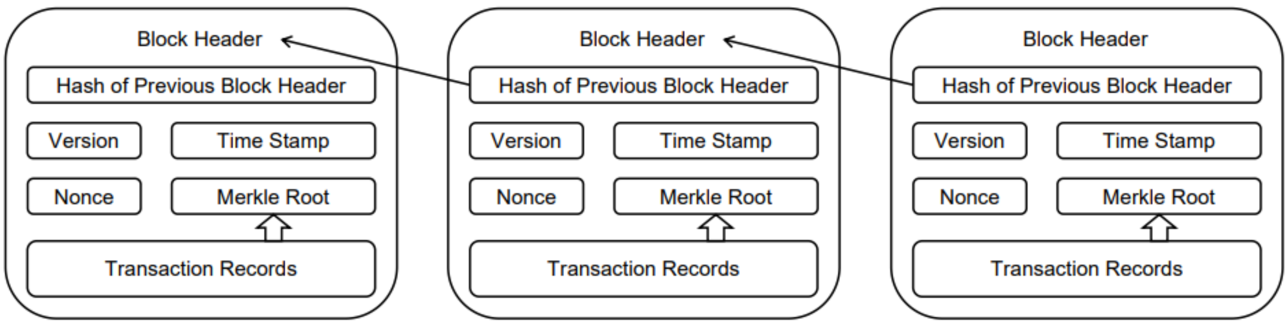

Figure 1. The blockchain's architecture [1]

- Block Header:
  A block header, which is used to distinguish a specific block on a whole blockchain, is repeatedly hashed to produce proof of work for mining rewards. A blockchain is made up of several different types of blocks that are used to store data about transactions that take place on a blockchain network. Each block has a distinct header, and the block header hash is used to uniquely identify each such block.
  The block header contains Metadata about the block. 4-byte for the blockchain version number, 32-byte for the previous block hash, 32-byte for Merkle root, 4-byte for the timestamp of the block, 4-byte for the difficulty target of the block, and 4-byte for miner nonce make up this 80-byte long string.

- Hash of the previous block header:
  The chain is essentially secured by the previous block's hash, which connects to the previous block or its parent block.

- Version:
  The Blockchain version number is helpful for tracking protocol-wide modifications and updates.

- Time Stamp:
  The timestamp is a short piece of information that is uniquely serialized and saved in each block. Its primary purpose is to pinpoint the precise time that the block was mined and validated by the network.

- Difficulty target:
  The difficulty target is used to modify the level of difficulty at which the miners must work to solve the block.

- Nonce:
  Blockchain miners are solving for the nonce number. The blockchain miner who finds the answer receives the block reward after it is solved [2].

1) Blockchain technology

All transactions, both those that have already happened and those that will soon, are maintained as a public record on the blockchain. This historical ledger of transactions is referred to as a "block chain" since it is a chain of blocks. It continues to expand as miners continue to add new blocks to it in order to keep track of all recent transactions. The blocks are consistently added to the blockchain in a sequential and linear manner [3].

2) Miner

A person who has the computer hardware and software necessary to validate new blockchain transactions before being added to the blockchain is called a "blockchain miner". Miners compete using a cryptographic hash method to determine the best solution to a difficult mathematical problem [4].

3) Consensus

Miners use "consensus" to add new data records to a blockchain. Proof-of-work (PoW) is employed to achieve this consensus in the Bitcoin system, which is the most well-known implementation of public blockchain. PoW resembles a mathematical "puzzle." Although it is difficult to discover, this puzzle's solution is simple to confirm. The search for the solution is referred to as "mining". The first miner to figure out the solution can add the block to the longest chain and receive a Bitcoin as compensation. The full copies of transaction records in this decentralized system are spread across numerous networked miners. The consensus algorithm is used to handle the verification and confirmation of each transaction [1].

4) Smart Contract

It is a program created to adhere to particular computer enforcement of the contract negotiation or execution. Smart contracts typically permit transactions that are trustworthy, irrevocable, and traceable without the involvement of outside parties [3]. Five stages—negotiation, development, deployment, maintenance, and learning—can be used to summarize the full life cycle of a smart contract based on how it functions. A 6-layered architecture made up of an infrastructure layer, contract layer, operation layer, intelligence layer, manifestation layer, and application layer makes up the structure for exploring smart contracts, in accordance with this life cycle [5]. These layers are shown in Figure 2.

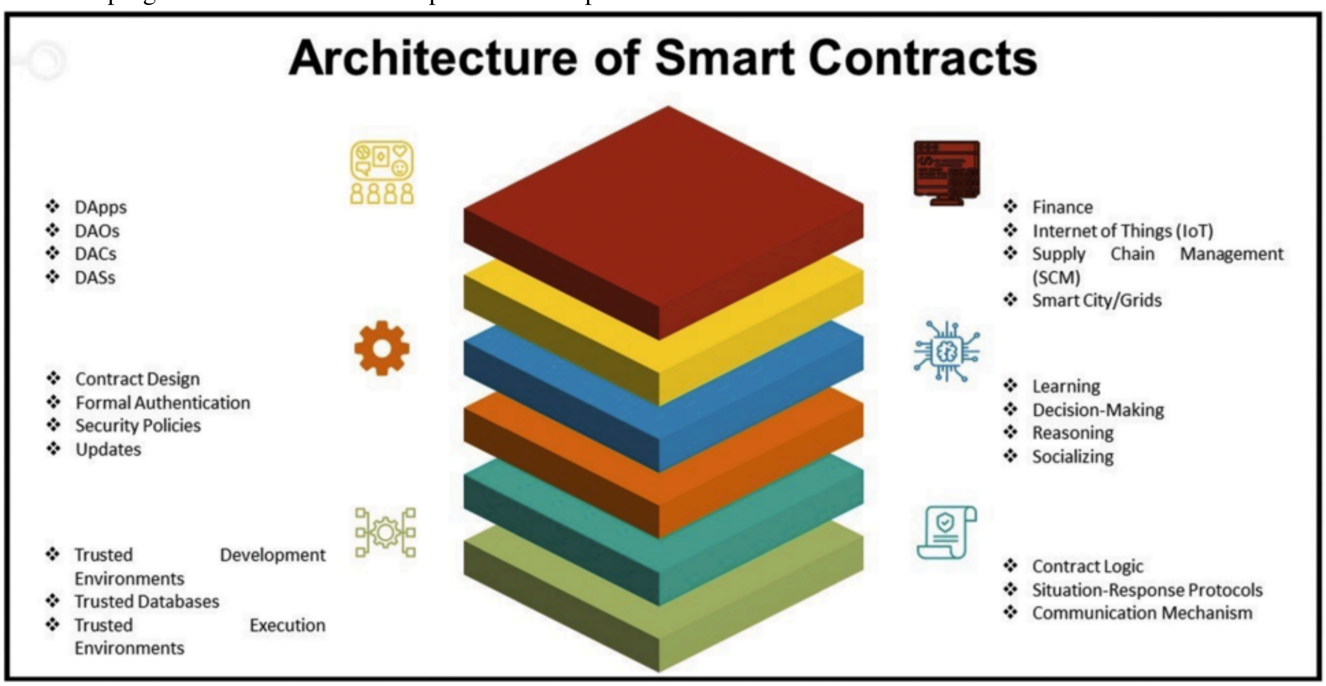

Figure 2. Smart contracts' architecture [5]

5) Characteristics of Blockchain

- Decentralization: Blockchain allows for the independent confirmation of transactions between two nodes without the need for outside interference or control from a central authority. Consequently, it lowers the total cost of the service, the performance bottleneck, and the danger of single-point failure.
- Traceability: A timestamp that was recorded at the moment of each transaction is also included with every entry in the distributed ledger. After analyzing blockchain data with accompanying timestamps, users may readily verify and track the origin of transactions as well as any revisions.
- Transparency: In general, users have equal access and privileges when interacting with public blockchain systems like Ethereum and Bitcoin. Additionally, every transaction is verified before being added to the global ledger and made instantly accessible to all users. In order to verify the committed transaction in the blockchain, data on the blockchain is transparent to each node.
- Immutability: Any type of data manipulation renders every subsequently formed block invalid since the

blockchain's structure relies on blocks that are connected one after the other, with each link being a hash of the header of the block before it. Any mutation can be easily detected because any transaction that is even slightly modified leads in the production of a new Merkle tree data structure.

- Non-Repudiation: Each node is given a private key to use in the blockchain's transactional system. The matching public key of that node can then be used to access and validate this by other nodes. As a result, the source node of the transaction cannot reject any transactions that have been digitally signed using cryptography.
- Pseudonymity: Even though blockchain transactions are transparent, the system still can maintain some level of privacy by giving users anonymous addresses. However, because these addresses can be found, blockchain systems can only keep the confidentiality up to a certain level. As a result, blockchain can only preserve pseudonymity, not complete confidentiality [5].

B. Machine Learning

Machine learning is the process of automatically finding significant patterns in data. With the use of machine learning (ML), which is a form of artificial intelligence (AI), Computers can now act independently without explicit programming [6]. In order to forecast new output values, machine learning algorithms use historical data as an input [7]. Building computer programs that can learn from data is the objective of machine learning, where it evolved into a typical tool for practically any task requiring data extraction from large data sets.

C. Deep Learning

Deep Learning is a group of machine learning techniques that aim to model data using complicated architectures that combine several non-linear transformations.Today, it is regarded as a core technology. Due to its ability to learn from data [8]. The artificial neural layers that are combined to create deep neural networks are the fundamental building blocks of deep learning. Neural networks can have many hidden layers. Deep learning has recently gained wide acceptance as a way of increasing the solution of specific kinds of challenging computer problems, particularly in the computer vision and natural language processing domains. Deep learning models outperform conventional machine learning techniques in terms of speed by extracting high-level, complicated abstractions as data representations. In other words, instead of having the data analyst to manually choose the relevant attributes, a deep learning model will learn the properties that are crucial by itself. Figure 3 represents the structure of deep learning.

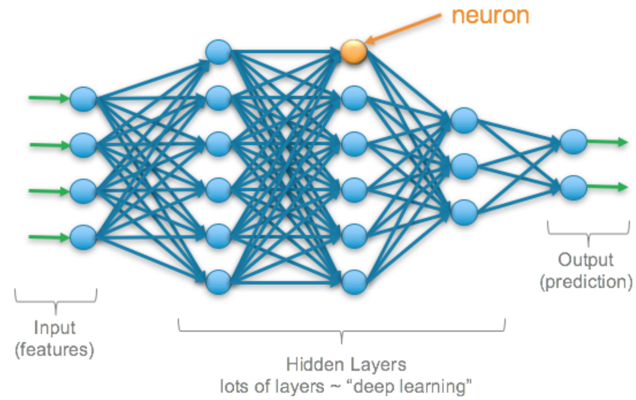

Figure 3. Deep Learning's architecture [9]

Convolutional Neural Network (CNN) is a type of deep learning algorithms built on multi-layer neural networks. CNNs are capable of classifying, detecting, and segmenting objects as well as learning significant properties from images.

D. Pattern Recognition

Pattern recognition is the process of identifying patterns in data involves using machine learning algorithms. It categorizes data using statistical information or knowledge of patterns and how they are represented [10]. The idea of learning is used to produce pattern recognition. The system can be trained and made adaptive through learning in order to produce results that are more precise. Pattern recognition systems are trained using labeled training data. Each input value that is utilized to generate a pattern-based output has a label associated with it.

E. Parallelism

Basically, parallelism is a sort of processing in which numerous computations or actions are executed concurrently to increase computing speed [11]. In the machine learning approach, due to the growth of data volumes, and the complexity of the algorithms, data processing became a challenging task in many fields. Parallelism can be effective in dealing with those problems in order to reduce the calculating time.

F. Blockchain and Healthcare

In the healthcare systems, a blockchain network is utilized to store and share patients data amongst hospitals, diagnostic labs, medical companies, and doctors. Blockchain applications can precisely detect serious mistakes, including potentially deadly ones, in the medical industry. Therefore, it can enhance the efficiency, security, and transparency of exchanging medical data between the healthcare sectors. Healthcare centers can acquire knowledge and improve the analysis of patients information with the use of this technology [12].

G. Deep Learning and Healthcare

Deep learning, a cutting-edge machine learning discipline, can automatically learn strong, complex features from raw data in contrast to conventional methods even without feature engineering. It is suggested that deep learning has clear advantages in maximizing the use of biomedical data and enhancing medical health level in the field of computer vision in medical image, electronic health record, genomics, and medication discovery. Deep learning is becoming more and more significant in the realm of medicine and has a wide range of potential applications [13].

3. Literature Review

In this section, we will select some of the existing in section 1 and section 2. Then, we'll be delivering a brief discussion on the content, the methods, and techniques used in each one of them. First, we made a keywords map of all the existing papers from 2019 to 2022 that include the techniques we talked about earlier applied in the healthcare domain. The keywords map is represented in Figure 4.

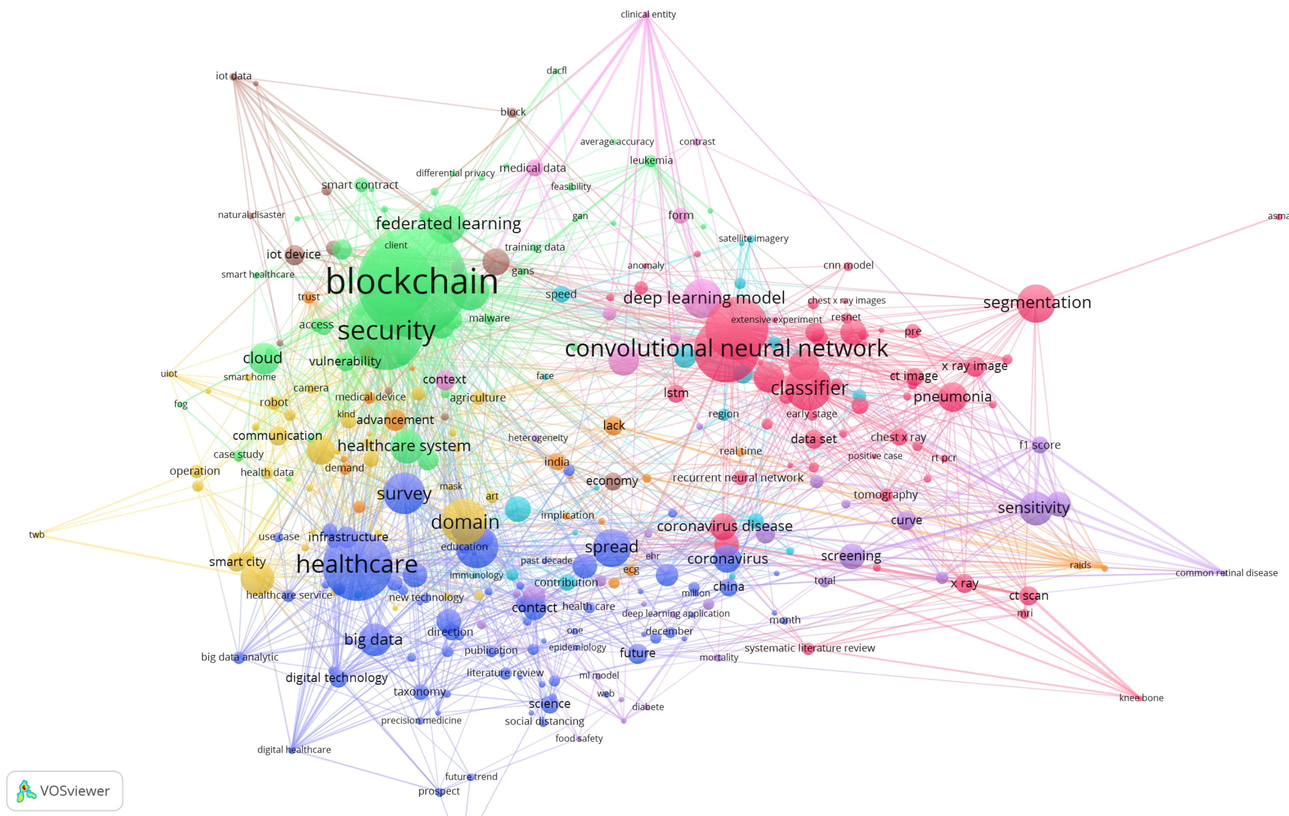

Figure 4. Keywords map of the existing papers (2019-2022)

The keywords map can help us define the main focus of the papers in a certain period to better understand them and further analyze them easily. As we can see in the map, the main focus topics of the papers are in a big size and a unique color where we can see that Blockchain, Convolutional neural network (CNN), and Healthcare are the main terms in these works. Then we can see that the secondary terms are security, and deep learning.

Next, we'll be discussing 15 selected papers that focus on the Blockchain, Learning algorithms, and parallel optimization methods. We'll be focusing on the problems solved by those works and the techniques that they have used.

Ovarian Cancer Blockchain based prediction tool

The accuracy, efficiency, and computing capability of deep learning models made them an essential element in the pharmacogenomics field. Pharmacogenomics studies how our DNA affects the way we respond to drugs. The work discussed in [14] suggested a prediction tool that can be used to predict ovarian cancer. This detection system has been developed based on the CNN Siamese network to predict the mutation behind cancer from microscopic images of protein expression analysis, which was retrieved from the database 'Human Protein Atlas'. This system also ensures the secure sharing of the healthcare data, patient records, and ovarian cancer predictions made by the model between organizations and research labs using the blockchain technique. The method's efficiency has been validated using the CRYPTO++ standards, and the model presented an accuracy of 86%. But, The model has been tested on a small number of classes and makes predictions based on unseen data since only the weights have been shared.

COVID-19 detection model based on Blockchain Federated Learning (CT Imaging)

The lack and the inaccuracy of testing kits have been a big problem disconcerting the medical practitioners in diagnosing COVID-19 patients, this article [15] presented an accurate collaborative model that recognizes the patterns of COVID-19 from the computed tomography (CT) images of the lungs using Capsule Network-based segmentation and classification. One of the challenges with this work is that there was no dataset, which required collecting real-life COVID-19 patients' data from different various hospitals to detect the positive cases while preserving the privacy of the organizations using blockchain-based federated learning. Due to the different kinds of CT scanners that those multiple hospitals have, the authors proposed a data normalization method to normalize the data collected from these different sources. In the end, this intelligent model was shared in a decentralized manner in the public network. This paper presented a new dataset named CC-19, related to the COVID-19. The model delivered a high detection performance through the Capsule Network. But it requires more data providers. The model's accuracy reached 98.68%.

COVID-19 diagnosis framework (X-Ray Images)

The fast increase of COVID-19 cases became a big issue that threatens the world. To help solve this issue, the work submitted in [16] proposed a detection framework based on deep learning that helps in the early diagnosis of COVID-19 cases from chest X-ray images. The pre-trained network DenseNet169 was used to extract features from the X-ray images, and those features were chosen by the analysis of variance (ANOVA) method (a feature selection method), to reduce the calculating time and improve the accuracy. The final features were classified by the eXtreme Gradient Boosting (XGBoost). The model was trained with ChestX-ray8 dataset, and the accuracy reached 98.72% for two-class classification (COVID-19, No-findings), and 92% for multiclass classification (COVID-19, No-findings, and Pneumonia). The proposed method outperforms other state-of-the-art methods. Yet, data privacy was totally neglected in this work.

Healthcare Emergency Expert system based on Blockchain

In the healthcare field, it is difficult to identify which are the emergency cases and which are not, valuable time might get wasted on unworthy cases instead of the real urgent ones. The authors in this article [17] presented a predictive system called "Emotional Medical System Administrator", this system can help determine a healthcare emergency using audiovisual emotion patterns. Emotion Recognition can help in deciding whether a patient has an emergency emotion profile to obtain fast medical assistance and lower the expenses. The system uses Convolutional Neural Networks (CNN) and Kalman filters for the identification and classification of audiovisual patterns, the model gained an accuracy up to 84.1%. The patient's emotions were captured using an acquisition device that consists of a camera, a microphone, and an oscilloscope. The system was trained using a centralized database that contains two types of emotion collections, one is for the regular emotion such as sadness, fear, anger. . . , and another one for the intense emotions like extreme sadness, extreme fear. . . etc. Another decentralized database was used to protect the patients' sensitive data and avoid hacking, which is the Blockchain. Although, The accuracy can be enhanced by employing a larger database, but this slows down the learning process.

Myopia Blockchain Deep Learning platform

Due to a variety of circumstances, Myopia has grown fast over the world in recent years and became a significant global health issue. Many people now have extreme Myopia, putting them at risk of sight-threatening consequences (Myopic macular degeneration as an example). In this article [18], the authors constructed a platform based on deep learning algorithms for diagnosing Myopia that could detect high Myopia and Myopic macular degeneration. The model was made up of three deep learning algorithms based on ResNet-101 (CNN), two of which are for identifying high myopia and the third for diagnosing myopic macular degeneration. The training was performed using different datasets from Singapore, and the testing was applied to datasets from India, Taiwan, Russia, China, and the United Kingdom. The main problem with deep learning algorithms is how to securely exchange data and models between the different hospitals and centers in different countries. This issue was solved by using blockchain technology to improve security and build a trusted platform. The model showed accuracy up to 91.3% for high myopia, and 96.9% for myopic macular degeneration. They also tested the deep learning algorithms against six human experts in assessing a randomly selected collection of 400 photos from the external datasets where the deep learning detected each situation better than all the six experts (97.8% accuracy for myopic macular degeneration and 97.3% accuracy for high myopia). Although that, the system has several limitations, it is only able to detect myopic macular degeneration as a binary (present or absent) result. Also, the underlying cause cannot be determined by the algorithm. In addition, using different definitions would require retraining of the deep learning algorithms. Moreover, the algorithm would not generalize well to very low-resolution images. Furthermore, the dataset can still be shared out by a collaborator which poses a data privacy issue.

Secure data sharing and detection of CT images using Blockchain

Nowadays, deep learning for image analysis techniques has become commonly employed to solve many difficulties in medical procedures, especially in diagnosing cancer, due to its capability of treating huge quantities of data. However, data privacy issues obstructed hospitals and medical organizations from cooperating to train deep learning models.

The authors of this work [19] suggested a way to secure data transmissions between various associations based on Blockchain technology. The method was represented in using the Blockchain to merge locally trained deep learning models by exchanging only the weights through a smart contract. This last, provides a safe sizable real-time data exchange platform among many data sources, in order to enhance lung cancer prediction in the early stage. Also, they proposed the Bat algorithm and data augmentation to solve the problem of the various sizes of CT images. The global model used Recurrent Convolutional Neural Network (RCNN) to detect the Region Of Interest for auto training over the distributed network using weight distribution. This model had a prediction accuracy of 99.2% for lung cancer. Yet, it makes predictions based on unseen data.

Malaria diagnosis smartphone platform based on Blockchain

When testing in remote rural regions with limited resources, where infectious diseases frequently cause the most burden, prompt communication of the testing results to healthcare professionals poses a significant problem in the process of the detection of these diseases. This article [20] proposed a smartphone platform for multiplexed DNA malaria diagnostics. The method makes use of a low-cost paper-based microfluidic diagnostic test, along with a convolutional neural network (CNN) for supporting local decisions and blockchain technology for managing and securely connecting data. The method was tested in the field in remote areas of Uganda, in which it successfully detected more than 98% of the cases. In addition to safe geotagged diagnostic data, this technology also makes it possible to incorporate infectious disease data into surveillance systems. The deep learning decision support system was trained using a dataset with five categories and 92 test images collected from the loop-mediated isothermal amplification (LAMP) diagnostic tests. Although, Security is at risk because the data is temporarily retained on the phone (for later propagation to the cloud).

Secure data sharing using Blockchain in healthcare systems

Through data exchange among intelligent wearables and devices, the industrial healthcare system has opened up the possibility of implementing enhanced and improved medical services. This work [21] presented a combination of Deep Learning (DL) techniques with Private Blockchain and smart contracts to create the new, safe, and effective data sharing framework called PBDL. Prior to adopting a smart contract-based consensus method, PBDL initially uses a blockchain technique to register, verify (using zero-knowledge proof), and authenticate the collaborating parties. Next, a novel DL technique combining Stacked Sparse Variational AutoEncoder (SSVAE) with Self Attention-based Bidirectional Long Short Term Memory (SA-BiLSTM) was proposed using the verified data. In this system, SA-BiLSTM recognizes and enhances the attack detection method while SSVAE converts the healthcare data into a new format. The security analysis and experimental findings utilizing the IoT-Botnet and ToN-IoT datasets demonstrated that the PBDL system outperformed other methods with an accuracy up to 99.89% for ToN-IoT and 99.98% for IoT-Botnet. This work needs a software-defined network version of the approach to test its performance and scalability.

Blockchain Personal Health Record (PHR) application

Personal health records (PHRs) are vital medical records that can help in improving the healthcare field and patients' safety, by providing the health data such as allergies, medicine doses, test results…etc, especially in emergency cases, where the patient might be unconscious, amnesiac… etc, or worthy time might be wasted on extra unnecessary tests. Although PHRs are very helpful, they are very vulnerable to fraud and data manipulation due to security issues, which makes the patients worried about their personal informations. This work [22] presented a new blockchain-based personal health records (PHR) application to store and share personal medical data safely. The application retrieved the component containing the patient's personal information off-chain and stores the encrypted data on-chain to avoid data fraud and manipulation. Moreover, this study also focused on the user's perspective by creating a mobile PHR application that uses blockchain technology. Unfortunately, the number of participants was limited, and the usability or practicality of each function of the application was not tested due to the short time.

Sports Injuries detection and assessment Deep Learning model

Sports medical data is currently an important area in the medical field and is in charge of guaranteeing sports safety depending on the degree of recovery following an injury sustained during a sporting activity. In order to successfully detect and assess the risk of sports medicine diseases, this paper [23] used an optimized convolutional neural network (OCNN) that can be used to categorize medical data relating to sports. And adopted the Self-Adjustment Resizing algorithm (SAR), which is enhanced by the convolutional neural network's self-coding method (SCM). The CNN makes it possible to analyze data from sports medicine in several dimensions and come to a cloud-based loop model to build a cutting-edge medical data network for sports medicine. A professional sports laboratory provided the majority of the medical data used. The OCNN gained approximately 80% accuracy. The study as an enhancement has to focus on getting a good evaluation of the sport-injury data with improved neural touch in the time series data function testing.

Parallel optimization algorithm for deep CNN

Deep Convolutional Neural Network (DCNN) is commonly used in Pattern Recognition and Image classification, it has proven its worth in many fields, including Medicine, Security, Advertising, and plenty of others. But, considering the growth of the network complexity and data size, some challenges appeared with this technique, such as the numerous network parameters, the lack of parameter optimization capabilities, and ineffective parallelism. In the MapReduce framework, this research [24] provided an optimizing algorithm for deep convolutional neural networks (FPDCNN). To minimize superfluous parameters, a clipping method was used; this method was proposed based on Taylor's loss (FMPTL), It doesn't only reduce the structure of DCNN but also lowers the training cost. Also, the ability to optimize parameters is improved by altering the initialization of weights in a glowworm swarm optimization method based on the data sharing method (IFAS). Finally, to ensure an equitable data distribution and hence increase the parallel efficiency of the cluster, a dynamic load balancing technique based on parallel computing entropy (DLBPCE) was developed. The authors' findings revealed that, when compared to previous parallelized methods, this technique not only lowers the computational cost of network training but also improves processing speed. Unfortunately, the authors focused only on the computation speed and neglected the accuracy.

Parallel Learning algorithm for pattern classification

Margin Setting algorithm (MSA) has proven its effectiveness in classifying images and recognizing patterns, but due to the growth and scalability of the data size nowadays, as well as the high complexity of the algorithms, parallelism is needed to reduce the execution time. This study [25] aimed to speed up the classification process, by applying a parallel version of MSA called PMSA. The model was developed using multiple datasets, and the findings revealed that the suggested PMSA improved the execution time significantly, with a promising speed up over the single-threaded CPU counterpart. The only problem with this research is that PMSA's thread-level parallelism is limited to multicore and multiprocessor systems. The algorithm showed accuracy up to 97.80%.

Parallel CNN model for Lung Sounds Classification

Recognizing different lung sounds captured by electronic stethoscopes is crucial for making an early diagnosis of respiratory disorders. in order to extract deep features, the authors of this study [26] proposed a novel pre-trained Convolutional Neural Network (CNN) model in which to improve the classification performance, a max-pooling layer and an average-pooling layer are connected in parallel. The Linear Discriminant Analysis (LDA) classifier is fed data from the deep features using the Random Subspace Ensembles (RSE) technique. The suggested approach was examined using the ICBHI 2017 challenge dataset. When compared to other existing approaches utilizing the same dataset, the deep features and the LDA with the RSE method delivered the best classification accuracy, increasing it by 5.75%. The model showed accuracy up to 71.15%. Although, the work did not address the issue of data protection.

A Blockchain CNN model for analyzing the Food Quality Well-Being for Lung Cancer

Deep Learning has been commonly used recently in the medical field, due to its capability of enabling extraordinary speed and accuracy in healthcare tasks. CNN and Blockchain are two critical components that, when combined, can ensure that diagnostic procedures are completed quickly and safely. Within Deep Learning, CNN is a form of artificial neural network frequently used for image/object identification and classification. On the other hand, Blockchain can ensure data confidentiality. This study [27] aimed to figure out how separate factors (such as features, filters, resolution, kernel size, epoch values, and padding value) influence the accuracy of a lung cancer prediction model (CNN + Blockchain). This study also focused on determining food quality. The authors ended up with a result that when features and filters are applied effectively, image augmentation and using a large number of images improve CNN accuracy in lung cancer prediction and food safety assessment, also 10–12 epochs were required for the model to achieve 99% accuracy with 1 padding, but when exceeding the number of epochs, the accuracy decreases. In addition, the authors found out that the pixel size of the images also affects the accuracy, whenever the image resolution decreases, the accuracy increases. The overall model was 92.5% accurate. However, the study may not be completely accurate, that's why the researcher conducted secondary research, which yielded data both in favor of and against the analysis.

Utilizing a Blockchain based Digital Pathology to improve diagnosis

It is getting harder to apply artificial intelligence algorithms for machine-assisted disease detection since data collecting and diagnosis have become more complicated in terms of storage, transmission, and security. A decentralized, safe, and individual privacy respectful digital pathology system was created and prototyped by the authors of this paper [28] utilizing Ethereum-based smart contracts, the nonfungible token (NFT) standard, and the Interplanetary File System for storing data. The suggested solution that was applied in data systems can speed up diagnosis and decrease process time while improving service quality and access to specialized pathological diagnostics. Also, other medical specialties that need high-fidelity imaging and data storage can use the suggested approach. But, the risk of information leakage by one of the collaborators remains.

4. COMPARATIVE STUDY

In this section we'll be presenting a comparative study between the previously discussed state of the art papers in section 3 based on several criteria and points of view including the advantages and limits of each work.

Table 1 : is a summary of each of the health problems, the Learning methods, the accuracy, and the presence or absence of the parallelism in the literature review papers. the gap (-) represents the absence of the information; the check

mark (✓) in the Parallelism case represents the presence of the Parallelism.

Figure 5 : represents the covered health problems rate in the state of the art papers, where we can see that Lung problems were the most covered diseases in the selected works.

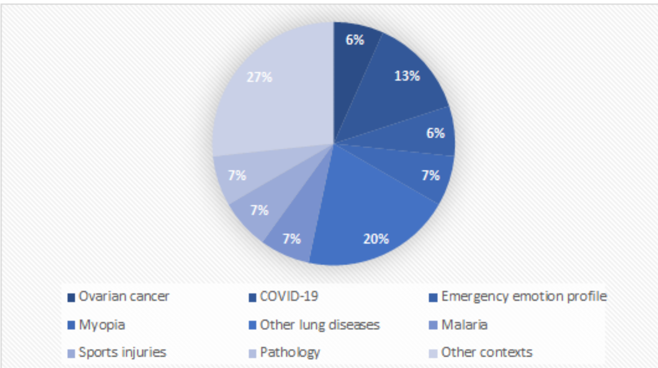

Figure 5. Health problems diagnosis rate in the state of the art papers

Figure 6 : represents the used Learning methods rate in the state of the art papers, where we can see that CNN was the most used Learning method in the selected works with 76%, while 24% is for the rest of the methods equally.

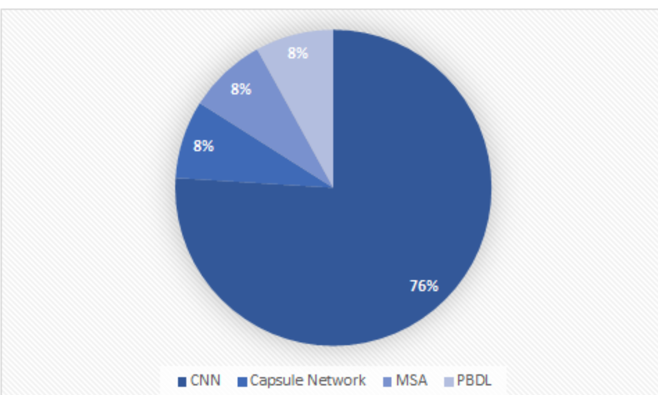

Figure 6. Pattern recognition methods rate in the state of the art papers

Figure 7 : represents the rate of articles that used the parallel optimization in their work, and the articles that haven't. As we can see 20% of the selected papers used the parallelism, while 67% didn't. As for the rest 13%, they didn't focus on Pattern Recognition in their work, that's why the parallelism wasn't required.

Table 2 : represents a comparison between the different blockchain platforms, types, and the consensus protocols used in the state of the art papers. The gap (-) represents the absence of the information.

Figure 8 : represents the rate of articles that used the Blockchain technology in their work, and the articles that

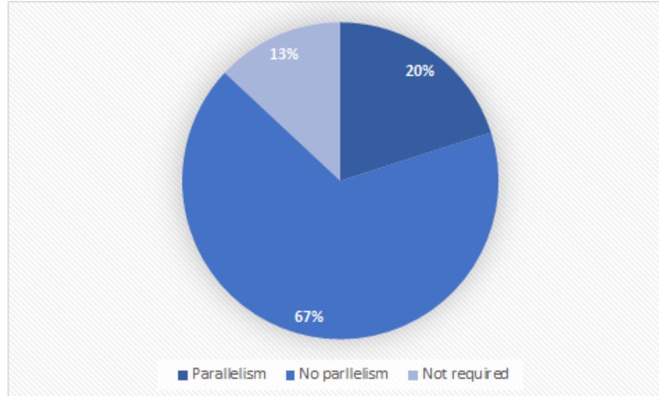

Figure 7. Presence rate of parallelism

haven't. As we can see 67% of the selected papers used the parallelism, while 33% didn't.

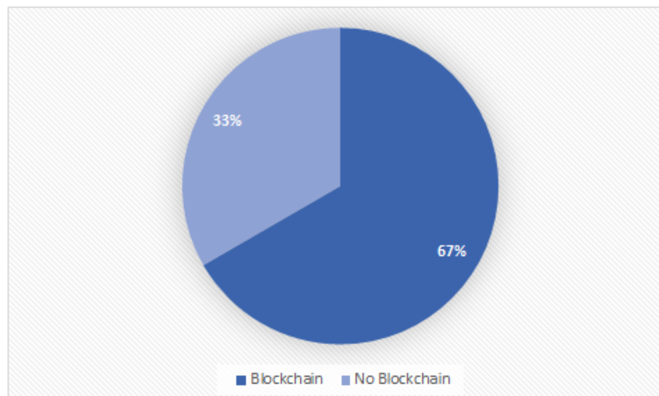

Figure 8. Presence rate of Blockchain

Figure 9 : represents the used Blockchain platforms rate in the state of the art papers, where we can see that 20% of the articles used Ethereum, 20% used HyperLedger Fabric, 20% used IPFS, 10% used Multichain, and the rest 30% didn't mention what platform they have used.

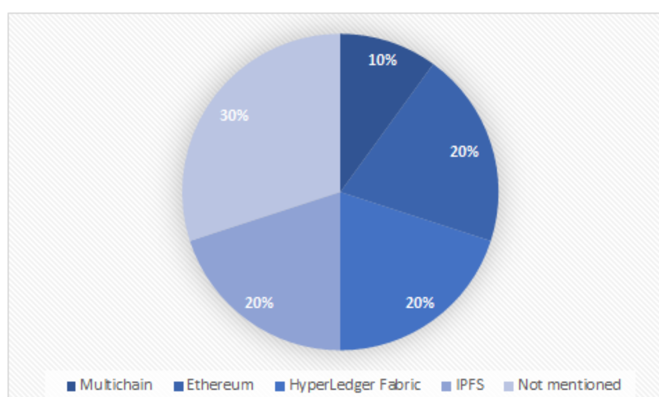

Figure 9. Blockchain platforms rate

TABLE I. A comparison between the methods of the state of the art papers.

Ref	Authors	Health Problem	Learning Method	Accuracy	Parallelism
[14]	M. Abraham et al	Ovarian Cancer	CNN Siamese network	86%	-
[15]	R. Kumar et al	COVID-19	Capsule Network + Federated Learning + Data normalization method	98.68%	-
[16]	H. Nasiri et al	COVID-19	DenseNet 169 + Analysis of Variance (ANOVA) + eXtreme Gradient Boosting (XGBoost)	98.72% for two-class classification and 92% for multiclass classification.	-
[17]	R.C. Aguilera et al	Emergency emotion profile	CNN + Kalman filters	84.1%	-
[18]	T.E. Tan et al	Myopia	ResNet-101 (CNN)	96.9% for myopic macular degeneration and 91.3% for high myopia.	-
[19]	R. Kumar et al	Lung Cancer	RCNN (CNN + Region of Interest (ROI)) + Bat algorithm.	99.2%	-
[20]	X. Guo et al	Malaria	CNN (using paper-based microfluidic diagnostic test)	98%	-
[21]	R. Kumar et al	-	PBDL (SSVAE + SA-BiLSTM)	99.89% for ToN-IoT and 99.98% for IoT-Botnet	-
[23]	H. Song et al	Sports injuries	OCNN + SAR + SCM	approximately 80%	-
[24]	Y. Le et al	-	FPDCNN + glowworm swarm + DLBPCE	-	✓
[25]	Y. Wang et al	-	PMSA	97.80%	✓
[26]	F. Demir et al	Lung sounds	CNN (max-pooling layer and an average-pooling layer were connected in parallel)	71.15%	✓
[27]	M.A. Aboamer et al	Lung Cancer	CNN	92.5%	-

TABLE II. A comparison between the Blockchain technology used in the state of the art papers.

Ref	Authors	Blockchain platform	Blockchain type	Consensus protocol
[14]	M. Abraham et al	Multichain	Private	-
[15]	R. Kumar et al	-	Private	Proof of Work
[17]	R.C. Aguilera et al	Ethereum	Private	-
[18]	T.E. Tan et al	HyperLedger Fabric	Private	Proof of Concept
[19]	R. Kumar et al	IPFS	Permissioned	Proof of work, Proof of stake, Delegated of work and Delegated Proof of Stake
[20]	X. Guo et al	HyperLedger Fabric	Permissioned	-
[21]	R. Kumar et al	IPFS	Permissioned	Enhanced Proof of Work (ePoW)
[22]	J.W. Kim et al	-	Private	-
[27]	M.A. Aboamer et al	-	Private	-
[28]	H. Subramanian et al	Ethereum	Permissioned	Proof of Concept

TABLE III. The different datasets used in the state of the art papers.

Ref	Authors	Datasets
[14]	M. Abraham et al	Human Protein Atlas.
[15]	R. Kumar et al	CC-19
[16]	H. Nasiri et al	ChestX-ray8
[17]	R.C. Aguilera et al	Database with two types of emotions: Regular and Intense.
[18]	T.E. Tan et al	from: Singapore Epidemio-logy of Eye Disease (SEED), Singapore National Eye Center High Myopia Clinic (SNEC-HMC).
[19]	R. Kumar et al	Local data of Lung Cancer from the Sichuan Cancer Hospital.
[20]	X. Guo et al	From loop-mediated isothermal amplifi-cation (LAMP) diagnostic tests
[21]	R. Kumar et al	IoT-Botnet, ToN-IoT
[23]	H. Song et al	from a professional sports laboratory
[24]	Y. Le et al	CIFAR-10, Fashion-MNIST, Patch-Camelyon bench-mark, Emnist-Bymerge.
[25]	Y. Wang et al	Standard image datasets, Bench-mark datasets, Irvine Machine Learning Repository, Pima Indian Diabetes, Wisconsin Breast Cancer, Australian Credit Approval, Wine, Svmguide2.
[26]	F. Demir et al	ICBHI 2017 challenge dataset.
[27]	M.A. Aboamer et al	from Healthcare sectors, Food images from social media and restaurants.

Table 3 : represents the names and sources of the different datasets used in the state of the art papers. **Table 4 :** represents the points of strength and weakness of each article respectfully.

5. CONCLUSION

In this paper, we have made a literature review of several articles that focus on three techniques: Blockchain technology, Learning algorithms, and Parallel optimization methods. We discussed the used Learning methods and their combination with Blockchain technology to solve the existing security problems. Also, we reviewed some Parallel Optimizations of these methods and how they were effective in reducing the calculating time. In addition, we made a comparative study for these state of the art papers according to several criteria to better identify their strengths and weaknesses. Through this study, we have

TABLE IV. The strengths and weaknesses of each paper.

Ref	Advantages	Limits
[14]	✓ High accuracy. ✓ Requires less calculating time.	✓ Tested on a small number of classes. ✓ Makes predictions based on unseen data.
[15]	✓ High detection performance. ✓ Data privacy is guaranteed. ✓ Data normalisation.	✓ Requires more data providers.
[16]	✓ Calculating time was reduced. ✓ Accuracy was improved.	✓ Data privacy was totally neglected in this work.
[17]	✓ High accuracy. ✓ Data privacy was secured.	✓ A more extensive database is required to improve the accuracy.
[18]	✓ Robust Perfor-mance. ✓ High accuracy.	✓ Only able to detect myopic macular degeneration as a binary (present or absent) result. ✓ The underlying cause cannot be determined by the algorithm. ✓ Using different definitions would require retraining of the deep learning algorithms. ✓ The algorithm would not generalize well to very low-resolution images. ✓ Dataset can still be shared out by a collaborator which poses a data privacy issue.
[19]	✓ More secure and accurate. ✓ Better Perfor-mance. ✓ Data Normalisation.	✓ It makes predictions based on unseen data.
[20]	✓ High security. ✓ Makes it possible to incorporate infectious disease data into surveillance systems.	✓ Security is at risk because the data is temporarily retained on the phone.
[21]	✓ High performance measures.	✓ This work needs a software-defined network version of the approach to test its performance and scalability.
[22]	✓ Secure data sharing. ✓ Focusing on the user perspective.	✓ The number of participants was limited. ✓ The usability or practicality of each function of the application wasn't tested.
[23]	✓ Strengthened algorithms of neural networks. ✓ High accuracy.	✓ Has to focus on getting a good evaluation of the sport-injury data with improved neural touch in the time series data function testing. ✓ Has to care more about data safety.
[24]	✓ Lowers the computational cost of network training. ✓ Improves processing speed. ✓ Suitable for processing large-scale datasets.	✓ Focused only on the computation speed and neglected the accuracy.
[25]	✓ Improves execution time significantly. ✓ High accuracy.	✓ Limited to multi-core and multi-processor systems.
[26]	✓ The performance was improved. ✓ Overcame other methods.	✓ The work did not address the issue of data protection.
[27]	✓ High accuracy. ✓ Data security was provided.	✓ The study may not be completely accurate.
[28]	✓ Decreased process time and improved service quality and access.	✓ The risk of information leakage by one of the collaborators remains.

seen the big importance of Blockchain technology in the Pattern Recognition systems, and how it is capable to secure the sensitive data, and even make trusted data exchange between different organizations to make developing systems safer and easier. Furthermore, we acknowledged the efficiency of the parallelism in speeding up the Learning systems and reducing the training time.

References

[1] Y. Du, Z. Wang, and V. C. Leung, "Blockchain-enabled edge intelligence for iot: Background, emerging trends and open issues," *Future Internet*, vol. 13, no. 2, p. 48, 2021.

[2] J. Frankenfield, "Block header (cryptocurrency)," 2021, https://www.investopedia.com/terms/b/block-header-cryptocurrency.asp Accessed: 30.10.2022.

[3] P. P. Ray, D. Dash, K. Salah, and N. Kumar, "Blockchain for iot-based healthcare: background, consensus, platforms, and use cases," *IEEE Systems Journal*, vol. 15, no. 1, pp. 85–94, 2020.

[4] K. Y. Bandara, S. Thakur, and J. G. Breslin, "End-to-end tracing and congestion in a blockchain: A supply chain use case in hyperledger fabric," in *Industry Use Cases on Blockchain Technology Applications in IoT and the Financial Sector*. IGI Global, 2021, pp. 68–91.

[5] S. Saxena, B. Bhushan, and M. A. Ahad, "Blockchain based solutions to secure iot: background, integration trends and a way forward," *Journal of Network and Computer Applications*, vol. 181, p. 103050, 2021.

[6] V. Kanade, "What is machine learning? definition, types, applications, and trends for 2022," 2022, https://www.spiceworks.com/tech/artificial-intelligence/articles/what-is-ml/ Accessed: 30.10.2022.

[7] E. Burns, "In-depth guide to machine learning in the enterprise," 2021, https://www.techtarget.com/searchenterpriseai/definition/machine-learning-ML Accessed: 30.10.2022.

[8] I. H. Sarker, "Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions," *SN Computer Science*, vol. 2, no. 6, pp. 1–20, 2021.

[9] S. Ronaghan, "Deep learning: Common architectures," 2018, https://srnghn.medium.com/deep-learning-common-architectures-6071d47cb383 Accessed: 30.10.2022.

[10] A. Waweru, "Understanding pattern recognition in machine learning," 2021, https://www.section.io/engineering-education/understanding-pattern-recognition-in-machine-learning/ Accessed: 30.10.2022.

[11] C. BasuMallick, "What is parallel processing? definition, types, and examples," 2022, https://www.spiceworks.com/tech/iot/articles/what-is-parallel-processing/ Accessed: 30.10.2022.

[12] A. Haleem, M. Javaid, R. P. Singh, R. Suman, and S. Rab, "Blockchain technology applications in healthcare: An overview," *International Journal of Intelligent Networks*, vol. 2, pp. 130–139, 2021.

[13] S. Yang, F. Zhu, X. Ling, Q. Liu, and P. Zhao, "Intelligent health care: Applications of deep learning in computational medicine," *Frontiers in Genetics*, p. 444, 2021.

[14] M. Abraham, A. Vyshnavi, C. Srinivasan, and P. Namboori, "Healthcare security using blockchain for pharmacogenomics," *Journal of International Pharmaceutical Research*, vol. 46, pp. 529–533, 2019.

[15] R. Kumar, A. A. Khan, J. Kumar, N. A. Golilarz, S. Zhang, Y. Ting, C. Zheng, W. Wang *et al.*, "Blockchain-federated-learning and deep learning models for covid-19 detection using ct imaging," *IEEE Sensors Journal*, vol. 21, no. 14, pp. 16 301–16 314, 2021.

[16] H. Nasiri and S. A. Alavi, "A novel framework based on deep learning and anova feature selection method for diagnosis of covid-19 cases from chest x-ray images," *Computational intelligence and neuroscience*, vol. 2022, 2021.

[17] R. C. Aguilera, M. P. Ortiz, A. A. Banda, and L. E. C. Aguilera, "Blockchain cnn deep learning expert system for healthcare emergency," *Fractals*, vol. 29, no. 06, p. 2150227, 2021.

[18] T.-E. Tan, A. Anees, C. Chen, S. Li, X. Xu, Z. Li, Z. Xiao, Y. Yang, X. Lei, M. Ang *et al.*, "Retinal photograph-based deep learning algorithms for myopia and a blockchain platform to facilitate artificial intelligence medical research: a retrospective multicohort study," *The Lancet Digital Health*, vol. 3, no. 5, pp. e317–e329, 2021.

[19] R. Kumar, W. Wang, J. Kumar, T. Yang, A. Khan, W. Ali, and I. Ali, "An integration of blockchain and ai for secure data sharing and detection of ct images for the hospitals," *Computerized Medical Imaging and Graphics*, vol. 87, p. 101812, 2021.

[20] X. Guo, M. A. Khalid, I. Domingos, A. L. Michala, M. Adriko, C. Rowel, D. Ajambo, A. Garrett, S. Kar, X. Yan *et al.*, "Smartphone-based dna diagnostics for malaria detection using deep learning for local decision support and blockchain technology for security," *Nature Electronics*, vol. 4, no. 8, pp. 615–624, 2021.

[21] R. Kumar, P. Kumar, R. Tripathi, G. P. Gupta, A. N. Islam, and M. Shorfuzzaman, "Permissioned blockchain and deep-learning for secure and efficient data sharing in industrial healthcare systems," *IEEE Transactions on Industrial Informatics*, 2022.

[22] J. W. Kim, S. J. Kim, W. C. Cha, and T. Kim, "A blockchain-applied personal health record application: Development and user experience," *Applied Sciences*, vol. 12, no. 4, p. 1847, 2022.

[23] H. Song, C. E. Montenegro-Marin *et al.*, "Secure prediction and assessment of sports injuries using deep learning based convolutional neural network," *Journal of Ambient Intelligence and Humanized Computing*, vol. 12, no. 3, pp. 3399–3410, 2021.

[24] Y. Le, Y. Nanehkaran, D. S. Mwakapesa, R. Zhang, J. Yi, and Y. Mao, "Fp-dcnn: a parallel optimization algorithm for deep convolutional neural network," *The Journal of Supercomputing*, pp. 1–23, 2021.

[25] Y. Wang, J. Fu, and B. Wei, "A novel parallel learning algorithm for pattern classification," *SN Applied Sciences*, vol. 1, no. 12, pp. 1–12, 2019.

[26] F. Demir, A. M. Ismael, and A. Sengur, "Classification of lung sounds with cnn model using parallel pooling structure," *IEEE Access*, vol. 8, pp. 105 376–105 383, 2020.

[27] M. A. Aboamer, M. Y. Sikkandar, S. Gupta, L. Vives, K. Joshi, B. Omarov, and S. K. Singh, "An investigation in analyzing the food quality well-being for lung cancer using blockchain through cnn," *Journal of Food Quality*, vol. 2022, 2022.

[28] H. Subramanian, S. Subramanian *et al.*, "Improving diagnosis through digital pathology: Proof-of-concept implementation using smart contracts and decentralized file storage," *Journal of medical Internet research*, vol. 24, no. 3, p. e34207, 2022.

Hakima Rym Rahal was born in the Algerian city of Biskra. She continued her college studies at the University of Biskra in the Department of Computer Science, where she received her license diploma for the project "Implementation of the different types of shading via shaders" in 2019. She then proceeded to receive her master's degree for the project "Direct illumination using cube mapping technique on GPU" in 2021 (option: Image and Artificial Life). Ms. Rahal, is a PhD student affiliated to the LINFI laboratory at the University of Biskra.

Sihem Slatnia was born at the city of Biskra, Algeria. She has followed her highs schools studies at the university of Biskra, Algeria at the Computer Science Department and obtained the engineering diploma in 2004 on the work "Diagnostic based model by Black and White analyzing in Background Petri Nets", after that she has obtained Master diploma in 2007 (option: Artificial intelligence and advanced system's information), on the work "Evolutionary Cellular Automata Based-Approach for Edge Detection". Also, she has contributed by an article at WILF 2007 about the segmentation based Cellular automata. Ms. Slatnia, is a full Professor at the University of Biskra.

Okba Kazar Obtained his engineer diploma (1987) at Constantine University (Algeria) and Magister degree (1997) followed by PhD degree from the same university (2005). He is member of editorial board of some international journals and author of more than 370 publication in international journals and conferences. He participates as a member of program committee and co-chair for international conferences. He is interested and working in artificial intelligence field and multi-agents' systems with their applications and also advanced information systems, web services, semantic web, Bigdata, IoT, robotics, cloud computing and information security. He is a Full Professor since 2011 at Computer Science Department of Biskra University. He is visiting professor at the United Arab Emirate University

Ezedin Barka is currently an Associate Professor at the United Arab Emirate University. He received his Ph.D. in Information Technology from George Mason University, Fairfax, VA in 2002, where he was a member of the Laboratory for Information Security Technology (LIST). His current research interests include Access Control, where he published a number of papers addressing delegation of rights using RBAC. Other research areas include Digital Rights Management (DRM), Large-scale security architectures and models, Trust management, Security in UAVs, and Network "Wired & Wireless" and distributed systems security. Dr. Barka has published over 50 Journals and conference papers. Dr. Barka is an IEEE member, member of the IEEE Communications Society and member of the IEEE Communications & Information Security Technical Committee (CISTC). He serves on the technical program committees of many international IEEE conferences such as ACSAC, GLOBECOM, ICC, WIMOB, and WCNC. In addition, he has been a reviewer for several international journals and conferences.