International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Kubernetes-based Deep Learning with Blockchain for Cancer Image Prediction

Arun Algude[1], Dr. Nihar Ranjan[2], Mayur Panpaliya[3]
[1]Research Scholar, Department of Computer Engineering,
JSPM's Rajarshi Shahu College of Engineering,
Pune, Maharashtra, India
arunalgude@gmail.com
[2]Professor, Department of Information Technology,
JSPM's Rajarshi Shahu College of Engineering,
Pune, Maharashtra, India
nihar.pune@gmail.com
[3]Research Scholar, Department of Computer Engineering,
JSPM's Rajarshi Shahu College of Engineering,
Pune, Maharashtra, India
mjpanpaliya@gmail.com

Abstract—The ability of modern artificial intelligence techniques to analyze complex data, including medical images, with greater precision and depth has significantly increased the amount of real-world data that can be collected for scientific research. This has led to the development of a global library and elevated computing choices to meet the demand for training model calculation. The development of medical image biomarkers involves additional processes that don't necessitate spending so much money on expensive computer resources. To collect and manage cancer data, this study presents a Blockchain-enabled deep learning framework for the Kubernetes cluster which replicates the training set of parameters of DL techniques across remote cluster nodes. Healthcare providers' data will be synchronized thanks to the distributed storage system known as the blockchain. Data are gathered for this suggested framework utilizing the Kaggle Dataset from several hospitals. Data preprocessing shows pixel correction and data augmentation. Following data preprocessing, U2-Net segmentation is used to segment up the images. This article uses blockchain to decentralize data, security, and privacy. A system called Kubernetes cluster is used to manage diverse containerized applications, Workload Distribution, Scalability, and Storage Management. The Kubernetes cluster is revealed in the study when investigating the cancer prediction capabilities of the hybrid learning algorithms VGG19 and BiLSTM. This proposed paper forecasts the article and explains the various applications for breast, kidney, and oral cancer predicated on the Blockchain-Enabled deep Learning model on Kubernetes and providing authentication for the use of the majority of the platform's features using a provided case.

Keywords-Deep Learning,Blockchain technology, KubernetesCluster, Healthcare industries, U2-Net segmentation, Privacy protection,VGG19 and BiLSTM,cancer images.

I. INTRODUCTION

In the most recent years, DL has emerged as the most promising strategy for next generation. The DL technique serves as a powerful analytical tool for reliable mining (feature extraction and representation) of large amounts of unstructured IoT data, including audio and image data produced by various IoT devices [1]. Many activities, including object detection, speech recognition, image classification, and other processes, utilize the mined data. Application-aware data created by terminal devices must be continuously provided for big data analysis utilizing DL to be accurate and timely.

Data ownership and security are the two areas in which healthcare systems are concentrated. Due to the patient health records system's poor security network, sensitive data may be exposed and utilized inappropriately by a third party [2, 3]. Data ownership is another factor since patients do not control their health records. Patient health records are now stored electronically rather than manually as they once are. The application of blockchain techniques in the medical field is a new exploit every day fresh research is available in this area to support this transformation and to induce new techniques for running and maintaining healthcare information.

Blockchain technology has drawn the attention of many academics, organizations, and businesses, particularly in the context of the use of the digital currency Bitcoin. A peer-to-peer network's transactions can be safely recorded on a blockchain, a decentralized ledger. Additionally, it makes transactions transparent and verifiable. Blockchain-based systems' main objective is to make it possible for two parties to conduct secure transactions without the use of an intermediary. Blockchain is capable of running on millions of devices and aids in protecting the anonymity of numerous reports. One of the top sectors in today's technology for communications and information market is the healthcare sector. The IoT within the healthcare sector has now made it possible to keep electronic medical records and monitor patients from a distance. Because of the quantity and variety of healthcare data produced by multiple sources, questions regarding the data's accuracy are also raised [4-6].

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Medical data can also be used for a variety of purposes, such as cancer prediction. As a result, it is difficult to verify data quality when combining data from different devices.

Deploying blockchain nodes on Kubernetes (such as Cryptocurrency or Hyperledger Fabric) is one technique to be using Kubernetes with blockchain technology. These nodes' deployment, scalability, and availability can be managed via Kubernetes, which can enhance the blockchain network's dependability and scalability. Blockchain data will be distributed and scaled using Kubernetes. To store the blockchain data, Kubernetes offers a variety of storage choices, including permanent volumes and object storage [7, 8].

A. Blockchain in Healthcare

Patient health data can be exchanged and stored on a very secure platform using blockchain technology without worrying about data loss or misuse. It is possible to use an EHR architecture built on a blockchain, which fosters interoperability and trusts between all parties. To store medical data, it has an immutable, distributed auditable, chronological, and time-stamped ledger. It's critical to comprehend how a secured blockchain function before deploying blockchain technology in healthcare businesses [9, 10]. The distributed storage system known as blockchain promises to synchronize data among health experts. Each block contains personal medical data that can only be accessed with permission. The benefits of blockchain include decentralized storage, permission, immutability, and enhanced capacity.

B. Kubernetes Cluster

An open-source container orchestration framework called Kubernetes automates the scaling, administration, and deployment of the containerized program. The smallest execution unit in a pod on the Kubernetes framework is a grouping from one or more containers. Pods share storage, have a specific cluster IP, and are closely connected. Figure 1 depicts the master and worker nodes that make up the Kubernetes cluster. Every node can run on a physical system or a virtual machine. By default, only one master node supervises the cluster, however, to achieve high availability, many master nodes might be employed. The kube-controller, kube-apiserver, ectd, and kube-scheduler are just a few of the parts that make up the master node. Using the kube-apiserver, the entire Kubernetes cluster is accessible to the Kubernetes control plane. It accepts all requests from the client and every other cluster member, authenticates them, and changes the related Kubernetes database objects [11-13]. Using apiserver, the kube-controller continuously monitors the attempts to alter the shared state of a cluster from its present state to its preferred state. Unscheduled pods are found by the kube-scheduler, which then deploys them to the proper node in the cluster. Affinity and anti-affinity specifications, resource needs, hardware/software/policy constraints, and policy considerations all have an impact on the scheduling choice. Every cluster of data, including configuration information and the cluster's status, is kept in a distributed, etcd, and a consistent key-value store.

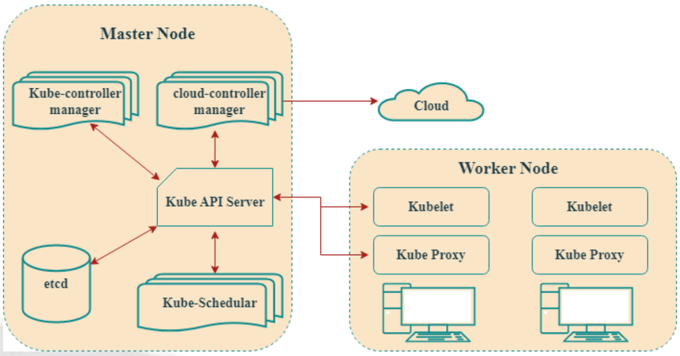


Figure 1.Architecture of Kubernetes.

C. Contribution

The following are the work's main contributions:

- To combine heterogeneous compute backends customized for varied workloads on a deep learning and synchronized distributed data storage system.
- In the proposed framework, images are segmented by U2-Net segmentation.
- To propose a hybrid model of VGG19 and BiLSTM for cancer prediction.
- To deploy permissioned blockchain using Hyperledger fabric on Kubernetes cluster to enable scalable networking across blockchains.
- To efficiently transmit disease prediction information utilizing the safe transactions of blockchains.

The remaining information is structured as follows. Part II provides a summary of similar work. Part III explains the foundational components of the Blockchain-Enabled deep Learning on the Kubernetes system for cancer image prediction. Part IV discusses the framework's performance evaluation and implementation results. Lastly, Section V includes conclusions and recommendations.

II. LITERATURE SURVEY

To collect and manage cancer data, this study presents a deep learning framework for illness prediction on the Kubernetes cluster that is Blockchain-enabled. This framework distributes the learning model parameters of DL algorithms among remote cluster nodes. Artificial intelligence (AI) algorithms can be used in the procedure for gleaning valuable lessons learned through medical imaging (i.e., deep learning or machine learning). These algorithms have been used in a variety of industries, such as healthcare and medicine. In recent efforts, a blockchain-based approach to decentralized learning has been proposed. A technique that encourages multi-center machine learning is distributed learning. After that, the scientific community began to pay close attention to blockchain technology. Because of this early success, blockchain technology is now being used in the healthcare sector. Now, among other applications in healthcare and other fields, Blockchains can be utilized to assure secure data exchange, compliance with license restrictions, and the prevention of medicine counterfeiting.

Ramanan et al. [16] provided a solution to the issue in the health sector using a secure, invasive, blockchain-based data transfer and the CPS classification methodology. This study provides an appropriate framework for analyzing the mammography image to identify and categorize distinct cancer stages. Images are extracted from the mammography and put through several processing processes to detect breast cancer. To protect patients' privacy, healthcare institutions as well as other organizations are unwilling to disclose patient information. Here, secure data transmission based on the blockchain is used. A smart gateway using blockchain has three layers: a gateway layer and a device level. To provide consumers with knowledge anywhere and in every situation, blocks are traded. They possess a fixed ID as well as the computing power to carry out PKI and SHA2 encoding and decoding operations.

Noorbakhsh et al. [17] presented to demonstrate how CNNs can be consistently used among numerous types of cancer, allowing evaluations to discover common spatial characteristics. We create a CNN framework to classify tumor/normal, cancer subtype, and mutations in 27,825 eosin, hematoxylin, & scanned pictures from the CGA. Breast, bladder, and uterine cancers all exhibit spatial characteristics that are very simple to spot, indicating that these malignancies could serve as standard kinds for image analysis. We used the LUSC and LUAD slides from the CPTAC dataset to test the trained CNNs' cross-classification accuracy. Finally, using 170 colon and breast cancer pictures annotated by pathologist tissues, we compare and contrast CNNs and discover that both intercellular and cellular regions influence CNN reliability.

Toka et al. [18] demonstrated a scaling engine for Kubernetes that can automatically decide how much capacity is needed to handle the actual unpredictability of incoming requests. Using a short-term assessment loop, this engine pits multiple machines learning forecasting techniques against one another to constantly favor the approach that best matches the dynamics of the real request. Also, they added a small management parameter that would make it simple for the provider of cloud-tenant applications to determine the optimal balance in the exchange between resource over-provisioning and SLA violations. Through analytical modeling and assessment of the existing Kubernetes behavior, they motivate the scaling solution.

Zerka et al. [19] introduced Machine learning with Chained Distributed, a ground-breaking sequential distributed learning approach combined with a blockchain platform for distributed learning. C-DistriM would be practical and produce results that are comparable to those of a traditional centralized strategy. Health centers can dynamically engage in distributed learning model training thanks to C-DistriM. We demonstrate how C-DistriM can forecast survival from lung cancer two years from now using the NSCLC-Radiomics open data. This proof-of-concept work is presented using the open Radiomics- NSCLC dataset. By storing unprovable data of the training procedure on the blockchain, C-DistriM promotes confidence between the partners. The reference standard is the model developed using a centralized approach, in which the training is conducted without integrating the blockchain, and all the data is kept in a single database.

Nasir et al. [20] suggested a solution that employs the early detection of kidney cancer using various deep learning algorithms and the Internet of Medical Things-based transfer learning technique. It also involves blockchain-based cloud services and transfer-learning trained models for patient data protection. Data is originally collected from several hospitals using the IoMT method, then sent to the data preparation layer of the suggested architecture. Pixel augmentation and data preparation are visible in the data preparation. After preparing the data, the suggested method separates it into training and validation data before storing it within a blockchain-based private, secure cloud. The next phase of the process involves importing information from a private cloud into the training layer, which then provides the deep learning methods Stochastic Gradient Descent Momentum (SDGM), Adaptive Momentum Estimation (ADME), and Root Mean Square Propagation (RMSP) to train the data samples. The most effective learning algorithm from private cloud Z should be imported in the last phase, coupled with samples of validating data from a private data cloud. Then, kidney cancer should be predicted using the testing protocols.

Nguyen et al. [21] provided a safe intrusion detection and classification methodology for CPS using blockchain-based data transmission in the healthcare industry. The method that is being given uses sensor devices to do data collecting, and the deep belief network (DBN) model is used to detect intrusions. Additionally, the proposed methodology accomplishes privacy and security by creating numerous shares of the acquired image using the multiple share creation (MSC) concepts. Also, the residual network (ResNet) based classification model, which uses blockchain technology, is implemented on the cloud server for safe data transmission to detect the existence of the disease.

Durga and Poovammal [22] created a cutting-edge technique based on the FL paradigm and the blockchain. Reduced complexity is taken care of by the federated learning paradigm, and distributed data with privacy preservation is made possible by blockchain. The suggested FL ensemble five deep learning blockchain (FLED-Block) model, collects data from various healthcare facilities, is developed using a hybrid capsule learning network, and data from multiple healthcare facilities are gathered, and produces precise predictions while preserving privacy and sharing among authorized parties. As the initial phases in the data normalizing process, spatial normalization and signal normalization are used. Then, COVID-19 patterns are found in lung CT scans using deep learning techniques. The ensembled capsule system is used for image segmentation & training to increase generalization. The ensemble learning machines are then used to classify the COVID-19 photos (ELMs). Lastly, to create both the globalized model and the privacy issue that must be addressed, we use the FL technique.

Although emerging technologies like distributed learning present a way ahead, they unfortunately frequently suffer from a lack of openness, which erodes confidence in the data utilized for analysis. The existing framework has high Latency and low throughput compared to the proposed method.

III. METHODLOGY

This research demonstrates a Blockchain-enabled deep learning framework for disease prediction on the Kubernetes cluster that distributes the learning model parameters of DL algorithms across remote cluster nodes to gather and handle cancer data. Using a database of contests from the online data science site Kaggle Dataset, data is initially gathered for the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

proposed framework from several hospitals. Data preparation shows pixel correction and data augmentation. The suggested model uses multiple image processing techniques for pixel correction to enhance the quality of the data sample for improved prediction efficiency. To overcome the limitations of imbalanced data, the suggested system employed data augmentation techniques to balance the data samples. Further data preprocessing, images are segmented by U2-Net segmentation. This will aid in streamlining the segmentation procedure. A blockchain-enabled Kubernetes cluster can offer several benefits in terms of security, scalability, and reliability to manage and deploy containerized applications. It can automate the scaling, deployment, and containerized applications management. Blockchain is a decentralized and tamper-proof digital ledger technology that can enable secure and transparent data sharing and management. In this proposed framework using hybrid learning algorithms VGG19 and BiLSTM to predict cancer such as breast, kidney, and oral cancer.

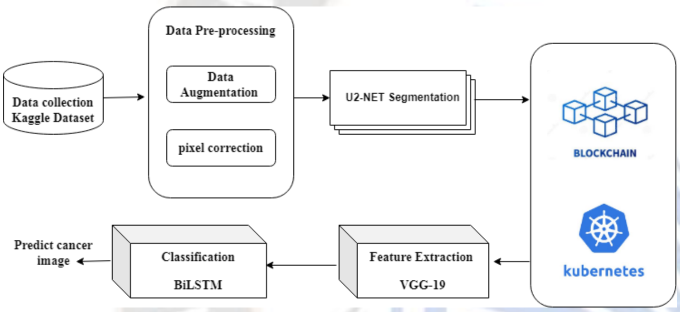


Figure 2. Block representation of the suggested framework.

A. Data collection: Kaggle Dataset

Since its founding in 2010, Kaggle has emerged as the top community for data scientists to work together to explore and create data-based models, compete, and connect with one another primarily through dynamic notebooks and forum discussions [23]. The community is expanding, and there are currently 2.9M users from 194 different countries. Data scientists and machine learning aficionados can engage in challenges on Kaggle, a well-known platform for data science competitions, and demonstrate their abilities. Also, Kaggle provides access to a sizable collection of datasets that may be used to practice data analysis and machine learning. In Kaggle, thousands of datasets are available, covering a variety of subjects like business, medicine, sports, politics, and more. These datasets are frequently added by members of the Kaggle community or by businesses that want to share their data with others.

B. Data Pre-processing

Data preparation involves pixel correction and data augmentation. To solve the shortcomings of unbalanced data, the suggested approach employed data augmentation methods to balancing the data samples. One of the key picture preparation methods that can be used offline or online is data augmentation. While online augmentation techniques are mostly employed with huge datasets, offline augmentation methods are employed to enlarge tiny datasets [24]. Rotation, Flipping, Translation, Scaling and Shearing, Zooming, and Adding Noise are examples of data augmentation procedures for image data. In data pre-processing for image data, pixel correction is a technique that involves adjusting the pixel values of images that may have been warped by the lighting, the camera, or other elements. Brightness and contrast adjustment, color correction, picture normalization, image filtering, and image cropping are some of the pixel correction methods used for image data. Pixel correction is a crucial step in the preparation of image data since it has a big impact on how well machine learning model's function. High-quality photos can help the model learn and generalize to new data, whereas inconsistent pixel values might cause mistakes and unreliable predictions. Images are segmented utilizing after image preprocessing.

C. U2-NETSegmentation

The sophisticated deep learning model U2-Net builds a mask for background removal, which is then utilized to segment the image using Pillow libraries and OpenCV image processing features, as detailed in. The U2-Net model receives the background-containing image and creates a mask for it. By removing the backdrop from the original image, and to extract the area of interest, the mask is utilized [25]. Figure 6 illustrates the process of segmenting leaves from images utilizing the U2-Net model, with the object of interest represented by the white portion of the mask and the backdrop by the black portion.

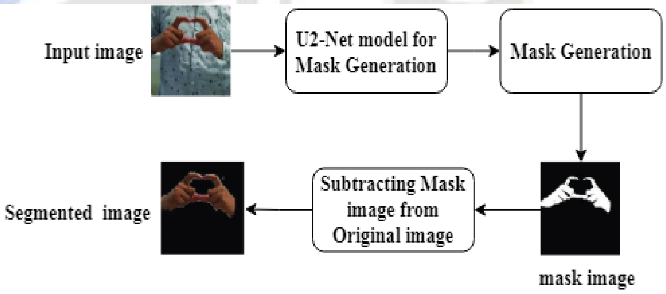


Figure 3. U2-Net segmentation.

D. Blockchain

Blockchain is capable of running on billions of devices and aids in protecting the anonymity of numerous reports. One of the top sectors in today's technology for communication and information market is the healthcare sector. The IoT in the healthcare sector has now made it possible to keep electronic medical records and monitor patients from a distance. Because of the quantity and variety of the healthcare data produced by multiple sources, questions regarding the data's accuracy are also raised. Medical data can also be used for a variety of purposes, such as prediction and diagnosis. As a result, it is difficult to verify data quality when combining data from different devices. As healthcare data are exchanged over the network, data confidentiality issues occur, and if the data is kept in a prominent central location, there is a risk of single-point failure. Moreover, centralized approved storage makes a denial-

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

of-service attack more likely to occur [26]. The blockchain makes it possible to find a solution to the problems described previously. The distributed storage system known as blockchain promises to synchronize data among healthcare providers. Each block contains private health information that is only accessible to those with permission.

The benefits of blockchain include immutability, decentralized storage, permission, and enhanced capabilities. As blockchain data is immutable, nobody else can alter it, and machine learning algorithms may be used to create any illness prediction model. Blockchain can provide a decentralized solution for security and integration. By trying to make the stored data and digitally signing each block to ensure a high level of legitimacy, the blockchain offers potential benefits. As the healthcare sector contains numerous participants and demands a high degree of trust between them, blockchain may be an appropriate option. Generally speaking, blockchains are perfect for highly dispersed systems where there is a large potential for activity tracking and where the accuracy of the data is crucial. It distributes the data throughout the medical field while preserving confidentiality and privacy [27]. Sharing data between hospitals that are the source and the requester must be secure. Hospitals are only able to disclose learned models to requestors, not the entire data set. The hospitals can connect, and federated data is used to learn using a consensus mechanism. Data from requesters and providers are kept on blockchain nodes. Only learning models, not the original data itself, are provided to protect data privacy. Every hospital uploads its image datasets for group learning during the first phase. In the next phase, hospitals communicate the adequately trained network weights with the blockchain and combine all of the local models into global models via federated learning.

E. Kubernetes cluster

A system called Kubernetes is used to manage diverse containerized applications across several hosts. It offers the fundamental tools and features needed for the deployment, upkeep, and scaling of numerous and various types of applications [28]. The Kubernetes network, which is entirely open source, will offer a more scalable service that can be expanded as needed by users. The proposed system will have a master node connected to numerous client nodes, each of which will have specialized functions. These features will be available to users via the master node.

TABLE I. DIFFERENCES BETWEEN KUBERNETES AND CONVENTIONAL DISTRIBUTED SYSTEMS.

Attribute	Traditional Period	Kubernetes Period
DISTRIBUTION OF WORKLOAD	Prior to this, there was no structure in place for dividing up the work services among several machines, therefore scheduling, allocating, and managing resources became essential.	The master and worker nodes are connected using the SSH protocol, which makes it much simpler to divide work across the nodes and manage resources.
SCALABILITY	Scaling the resources to meet user needs and requirements is a problem. Scaling the resources demanded a significant labour and financial investment. Moreover, the cluster deployment would not function if one node went down.	For the distribution of work, Kubernetes provides containers, and we can increase the resources as much as we need. Scaling the resources will enable efficient workload distribution and resource allocation. Kubernetes automatically restarts the container even if one fails, protecting the cluster in the process.
REQUIREMENTS OF THE SYSTEM	There has been much discussion over how many resources are needed and how much CPU and memory (RAM) are used by the resources. Hence, in order for the work to be done effectively, we would need to work on incredibly powerful computers every time.	Although there are also concerns with RAM and fundamental system requirements in this case, the user manages the CPU and RAM requirements for a specific container. The resources are automatically adapted and managed by Kubernetes so that the nodes are fully utilized.
MANAGEMENT OF STORAGE	Earlier, it cost a lot of money to store or install a storage system in our machine. Also, data centres were responsible for handling the data, raising concerns about its security.	Using Kubernetes, we can mount storage based on user preferences and requirements. Depending on the data's legitimacy, the user can save the information in a public, private, or hybrid mode.

Such criteria seek to identify the node that transfers artifacts a nodes from a distant repository particular dataset in the shortest amount of time for each new container request. With each new container request, we must select the node that has the required items in its central dataset in order to satisfy this need. Imagine that there are two nodes, node1, and node2 and that node1 has images A and B in its remote directory. Nevertheless, the local database of node2 only has image C. Any time a consumer provides a new container (container x) that requires image C, Node2 is by far the best node to use for container x execution. This is because node2 can expressly run container x even though it must delete image c from the server.

Platform-as-a-Service (PaaS) cloud services may deploy and maintain containerized modules thanks to Kubernetes [29]. This is made possible by Kubernetes' capacity to spread out several pods across different machines, enabling the scalability of a constantly changing workload. Each pod can support many dockers, and these dockers are in charge of hosting the services related to the pod. Kubernetes offers a much-needed method that enables the deployed application's scalability and resilience, increasing the platform's elasticity.

The goal of Kubernetes is to evaluate how well the Kubernetes mechanism works for building and running Kubernetes clusters. This cluster is created with the intention of making bioinformatics workflow deployment and scaling simple. The user may easily distribute resources and offer some front end functionalities (like logging). The network is designed to expand horizontally, and if the resources provided are insufficient, it will create slaves.

F. Feature Extraction

By gradually extracting significant information, a distinct representation of the input image is created in each layer of the CNNs. VGG-19 is used in the suggested technique to extract important data from the images obtained. The visualization reveals the type of categorized image used in the depiction.

1) VGG-19

The sole aspect of VGG-19 architecture that differs from VGG-16 is layer depth. It has three more layers, including three fully linked layers, five max-pooling layers, and 16 convolutional layers. There are 144 million parameters in the VGG-19 architecture [30]. These layers comprehend intricate visual patterns for efficient training. The VGG-19 is practical due to its simplicity, which includes three convolutional layers stacked on top to expand with depth level. In VGG-19, the volume size is controlled by using max pooling layers. The two FC layers employ 4096 neurons. Figure 4 shows how the vessel segmentation photographs were used as input data by the VGGNet DNN. During the training phase, convolutional layers are used to extract features, and some of these layers also include max pooling layers attached to them to reduce the dimensionality of the features. To extract features from the input photos, the first convolutional layer has 64 kernels (3 × 3 filter size). Fully linked layers are used to build the feature vector. The gathered feature vector is subsequently put through SVD and PCA to reduce dimension and extract features from picture data for better classification results. It takes a lot of work to minimize the higher dimensional data using PCA and SVD. PCA and SVD are superior to other reduction techniques because they are faster and quantitatively more stable. Last but not least, the SoftMax activation technique is applied during testing to classify the DR pictures using 10-fold cross-validation.

G. Classification

a) Bidirectional LSTM

In contrast to traditional LSTMs, bi-directional long term memory models, or BiLSTM, retain data for both the past and the future [31]. This kind of network is useful for applications where prediction depends on the complete input sequence. The input sequence is passed through one LSTM from beginning to end and another LSTM from end to the first patch in the sequence. A BiLSTM is created by combining these two LSTMs. The BiLSTM model is shown in Figure 4 with g(t) representing its state as it progresses backward through the sequence h(t) representing the state of the sub-BiLSTM as it moves forward through the ordered sequence and where t = 1, 2, 3,..., m. V(t) is a representation that depends on both the past and the future of the sequence but is most sensitive to the current inputs. The output unit V(t) resulting from superimposing h(t), and g(t). An output vector V(t) is calculated as

$$V_{(t)} = f\left(h_{(t)}, g_{(t)}\right) \tag{1}$$

When the two output sequences are combined using the function f. Concatenating functions, summing functions, average functions, and multiplication functions are some examples. The end result of a BiLSTM layer can be represented by the following vector,

$$V_{(m)} = f\left(h_{(m)}, g_{(m)}\right) \tag{2}$$

in which V_m, is the predicted sequence. Such a network is helpful for predicting the label of the patch sequence because it only needs the final output vector to summarize a sequence. To train this network, information is back-propagated first through forward h states and then through backward g states using the cross-entropy loss function L. The weights are updated following forward and backward passes. Each set of sequences is sent through the BiLSTM algorithm once, and the expected output indicates the class of the sequence. For our prediction, we employed a SoftMax classifier.

IV. RESULTS AND DISCUSSIONS

The model developed using a centralized approach, in which all the data are held in a single database, and the training is carried out without incorporating blockchain technology, is utilized as the standard reference.

A. Dataset used

The data set that is utilized to test and train the suggested framework. Some of the key characteristics that are analyzed to predict breast, kidney, and oral cancer. The dataset's goal is to determine if a patient has had disease-related symptoms or not. Images are segmented using U2-Net segmentation after the raw data has been preprocessed with data augmentation and pixel correction.

B. U2-NET

A segmentation method called U2-Net uses deep learning to create a mask a mask made from the primary image, which is then utilized to split up the image as illustrated in Figure 3. The technique outperformed the Watershed as well as GrabCut techniques, as evidenced by the minimal number of erroneously segmented photos. 99%, 91%, and 96% of the photos in Table 2 are accurately categorized as Susceptible, Healthy, and Resistant respectively. Figure 3 displays the segmentation outcomes on several U2-net image classes. Table 2 displays the outcomes of U2-Net on various classes. A tiny fraction of inaccurate segmentations is primarily due to photos with an inadequate focus on the leaf. However, few photos contain information other than leaves, such as a human hand, which makes it challenging to segment using the segmentation model.

TABLE II. SEGMENTED RESULTS AFTER USING THE U2-NET METHOD.

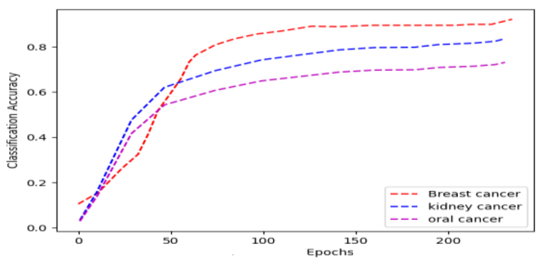
Class	Result (Total)	Accurate Segmentation	Inaccurate Segmentation
Susceptible	735	707	28
Healthy	673	667	6
Resistant	516	468	48

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

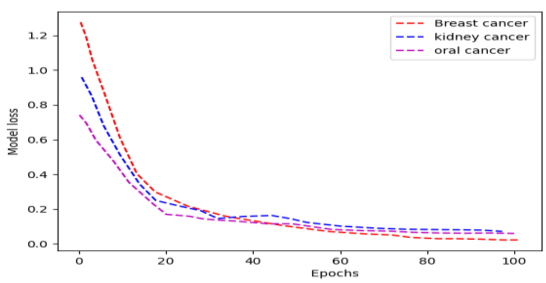
C. Analysis of Security & Privacy

The information is gathered from a wide range of sources, including hospitals that are outfitted with numerous different types of gadgets. The datasets have been divided across three separate cancer images to carry out accurate performance analysis. Many hospitals will have access to data-sharing benefits from DL as a result of this cooperation. Figure 1 illustrates how the performance of our suggested Kubernetes model altered as the hospitals or providers changed. Results are better when a range of information sources are used. As seen in Figure 4(a), there are no gradual fluctuations in accuracy. The degree of accuracy attained directly correlates with the number of patient slices. The model loss can also be used in a similar way. Figure 4(b) illustrates how the model loss has converged. Figure 4(c) displays the F1-score obtained by the suggested model after dividing the data into three distinct cancer images. Figure 4(d) displays the time needed to execute models that utilize the DL framework. While the DL model is guided by the local models, the local model has been prepared to utilize the entire data. Performance increases significantly as the variety of data sources expands, as seen in Figure 4(a, b). Yet, DL does not sacrifice high accuracy and manages data transfer while upholding anonymity.

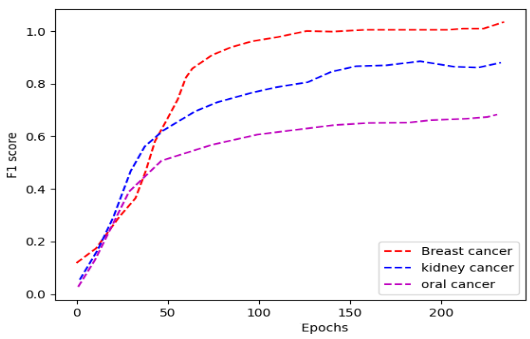
A



B



C



D

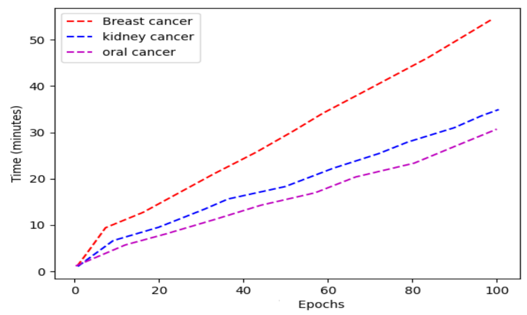
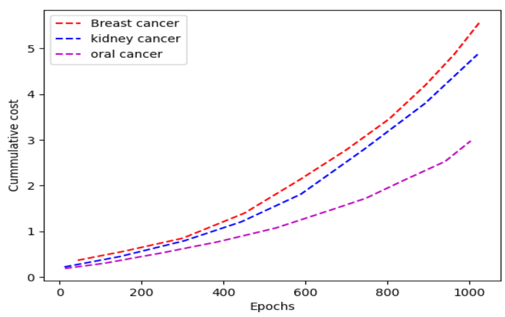


Figure 4.A result obtained by the suggested blockchain-based approach. (a) Accuracy of the suggested model on dataset for three different cancers. (b)Model loss on dataset for three different cancers. (c) The proposed model's F1 score(d) Time of dataset for three different cancers.

D. Computational Cost

The multiple local DL learning techniques are compared with the suggested blockchain-based DL. Despite this, blockchain-based federated architectures offer a very high level of privacy and security. Figure 5(A) demonstrates how the communication load increases along with the number of hospitals or transactions, increasing operational costs. Also, Figure 5(A) illustrates the computational expense of our suggested method for the various cancers with conventional DL models.

A



B

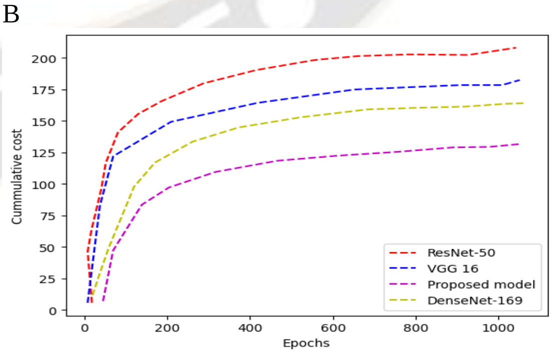


Figure 5. (A), (B)comparison of the models' computing costs.

E. CPU utilization

The CPU utilization percentage is calculated after testing both architectures ten times. The Ubuntu operating system's TOP

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

command is used to calculate CPU utilization. To calculate the proportions of both servers in the two systems (Master and Worker Nodes in Kubernetes and Server 1 and Host 2 in Docker), the resource utilization is first recorded. Figure 6 displays the CPU use for each pod. The machine's effective pod utilization rises as the number of CPUs grows. Consequently, if we wish to execute on local Virtual Machines, a high-performance system is recommended for completing the cluster functions.

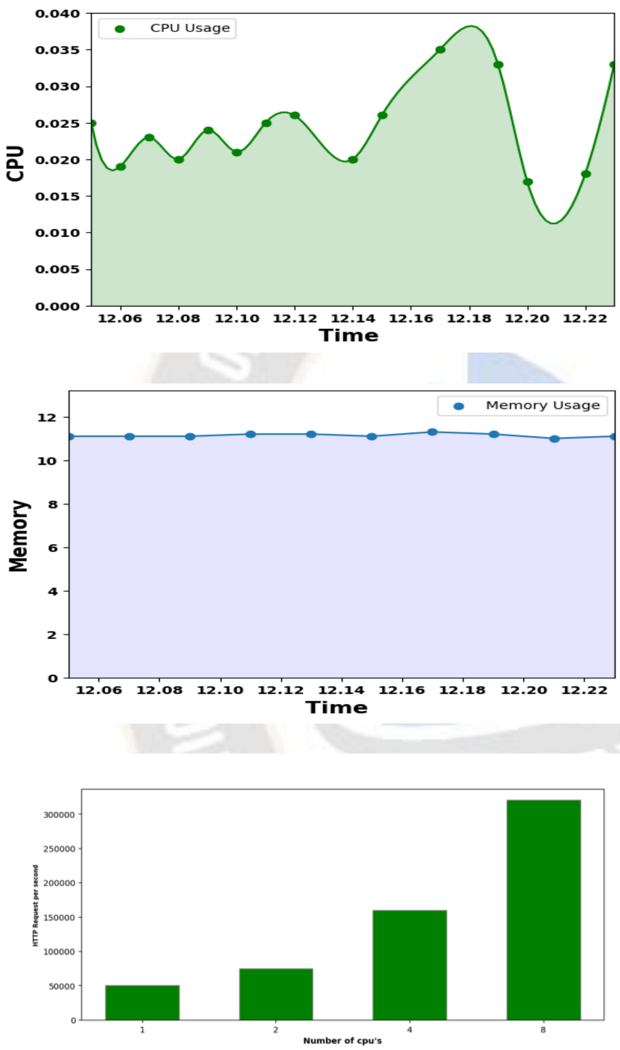
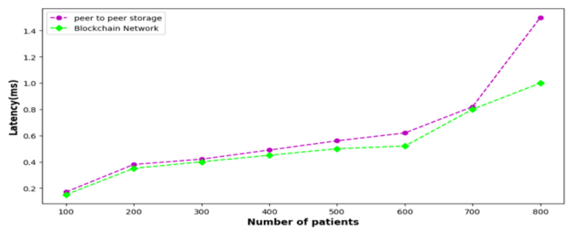


Figure 6. Number of CPU vs requests.

Instead of using centralized storage, the next round of research uses blockchain technology to store patient information. It is clear from Fig. 7(A) that as the quantity of patients in the network grows, so does the latency. Yet as opposed to peer-to-peer storage, the latency is always kept as low as possible for blockchain-based decentralized storage. Blockchain-based storage outperforms centralized storage with a minimum accuracy of 34.63% once there are 200 patients. When there are 800 patients, blockchain-based storage has a minimum latency that is 19.45% low compared to peer-to-peer storage. From the fewest number of users to the most, blockchain-based decentralized storage has the lowest latency compared to traditional storage solutions.

A



B

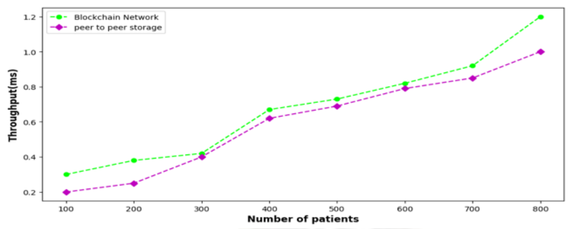


Figure. 7(A). Comparison of latency. 7(B). Comparison of throughput.

The following experiment measures the throughput for decentralized and peer-to-peer storage using a blockchain. It is clear from Figure 7(B) shows that, in terms of throughput, blockchain-based storage outperforms peer-to-peer storage. Blockchain-based storage outperforms peer-to-peer storage with 800 patients, having a throughput that is 43.52% higher.

F. Training and Testing

The training data for adaptive moment estimate is enhanced with cancer patients. The suggested architecture included 1000 data points to train the three types of cancer. Figure 8 illustrates how smoothly and convergently training progress is. Also, security analysis distinctions refer to a systematic procedure to permits hospitals to draw conclusions from the vast majority of data while assuring that those findings prevent the separation or re-identification of the data. In other words, the approach enables hospitals to make inferences from the vast majority of data while safeguarding personal information. Data security is increased because everything can operate automatically in accordance with a predetermined programme thanks to the system of distributed trust offered by BC Technology. The decentralized BCT uses a demanding set of algorithms, enabling it to guarantee that the data is correct, transparent, true, unaltered, and traceable. Data vendors are accountable for their data and are free to make any necessary alterations. Real data and the owner's digital signature are then added to the database that the BC keeps. The blockchain employs a vast number of cryptographic methods to ensure that user data is kept secret. A comparative is performed between the existing local DL learning models and the proposed blockchain-based model. Despite this, blockchain-based federated architectures offer a significant level of privacy and anonymity. The training and validation accuracy for the suggested models are shown in Figure 8. With the binary classification model, we witnessed improvements in all evaluation metrics, which showed how well our model performed throughout training, validation, and testing.

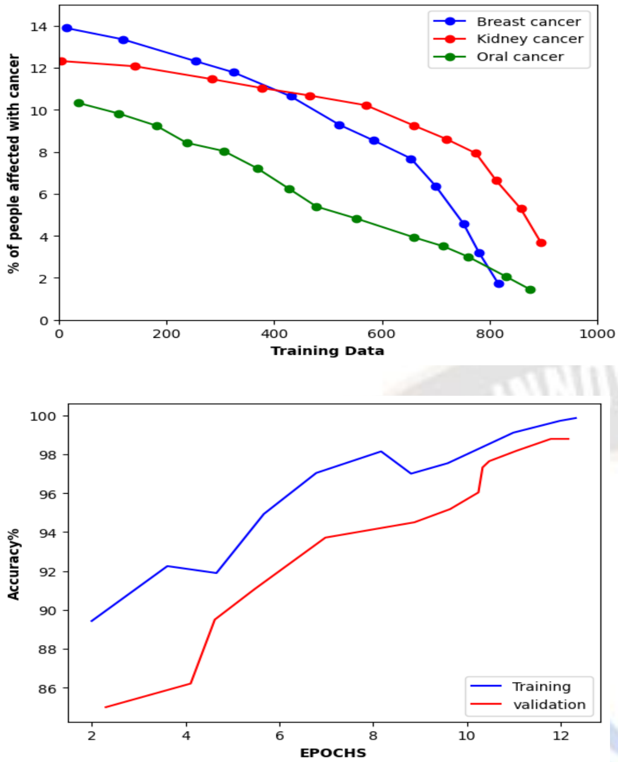


Figure 8. Training Dataset for Three cancer disease.

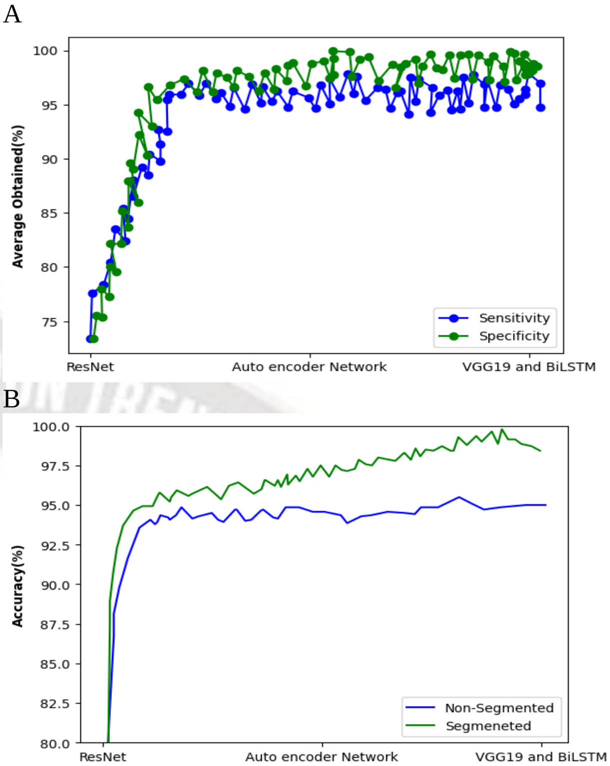


Figure 9. Comparison of Different Classification model (A)Analysis of Sensitivity and Specificity. (B) Evaluation of Segmented and Non-Segmented Image Accuracy.

G. Evaluation criteria

The data must next be classified and assigned to a particular class after the necessary feature has been extracted. Since our data is balanced and amongst the different categorization performance attributes, our analysis employs the following evaluation parameters: sensitivity, specificity, accuracy, F1 score and precision. Below are the definitions for some well-known parameters:

Evaluation criteria	Formula
Sensitivity	$\frac{TP}{FN+TP}$
Specificity	$\frac{TN}{FP+TN}$
Accuracy	$\frac{TP+TN}{FP+FN+TP+TN}$
F1	$2\times\frac{Recall\times Precision}{Recall+Precision}$
Precision	$\frac{TP}{FP+TP}$

H. Comparison of the Suggested Method with other Methods

Figure 9(B) displays the classification performance for various classification networks using segmented and non-segmented data. The x-axis displays the classification network, while the y-axis displays the proportion of correct classifications. For non-segmented images, the suggested VGG19-BiLSTM model achieved 95.77%, while for segmented images, 99.1%.

V. Conclusion

In this study, a system based on a deep learning model on Kubernetes that is enabled by blockchain can predict cancer. To preserve electronic healthcare data and uphold a high level of privacy, blockchain is crucial. The building of predictive models using the current classical AI techniques frequently calls for centralized data storage and training, which increases computing complexity and compromises privacy. This study suggests using a blockchain-enabled Kubernetes cluster to safeguard data to solve the issue. For the best prediction of breast, renal, and oral cancer diseases from the data recorded in a blockchain, a proposed algorithm is utilized. When compared to existing methods, our classification model's training, validation, and testing performance reflects higher accuracy. The future application of this research can be expanded to include the prognosis of many human cancer conditions.

References

[1] Haque, I. R. I., &Neubert, J. (2020). Deep learning approaches to biomedical image segmentation. Informatics in Medicine Unlocked, 18, 100297.

[2] Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., Al-Nemrat, A., &Venkatraman, S. (2019). Deep learning approach for intelligent intrusion detection system. Ieee Access, 7, 41525-41550.
[3] Benhamou, P. Y. (2011). Improving diabetes management with electronic health records and patients' health records. Diabetes & metabolism, 37, S53-S56.
[4] Srivastava, G., Parizi, R. M., &Dehghantanha, A. (2020). The future of blockchain technology in healthcare internet of things security. Blockchain cybersecurity, trust and privacy, 161-184.
[5] Anjum, H. F., Rasid, S. Z. A., Khalid, H., Alam, M. M., Daud, S. M., Abas, H., ...&Yusof, M. F. (2020). Mapping research trends of blockchain technology in healthcare. ieee Access, 8, 174244-174254.
[6] Zerka, F., Urovi, V., Vaidyanathan, A., Barakat, S., Leijenaar, R. T., Walsh, S., ...&Lambin, P. (2020). Blockchain for privacy preserving and trustworthy distributed machine learning in multicentric medical imaging (C-DistriM). Ieee Access, 8, 183939-183951.
[7] Yewale, A. (2018). Study of blockchain-as-a-service systems with a case study of hyperledger fabric implementation on Kubernetes (Doctoral dissertation, University of Nevada, Las Vegas).
[8] Poniszewska-Marańda, A., &Czechowska, E. (2021). Kubernetes cluster for automating software production environment. Sensors, 21(5), 1910.
[9] Kumar, R., Wang, W., Kumar, J., Yang, T., Khan, A., Ali, W., & Ali, I. (2021). An integration of blockchain and AI for secure data sharing and detection of CT images for the hospitals. Computerized Medical Imaging and Graphics, 87, 101812.
[10] Chattu, V. K. (2021). A review of artificial intelligence, big data, and blockchain technology applications in medicine and global health. Big Data and Cognitive Computing, 5(3), 41.
[11] Bannon, D., Moen, E., Schwartz, M., Borba, E., Cui, S., Huang, K., ...& Van Valen, D. (2018). Dynamic allocation of computational resources for deep learning-enabled cellular image analysis with Kubernetes. BioRxiv, 505032.
[12] Medel, V., Tolosana-Calasanz, R., Bañares, J. Á., Arronategui, U., &Rana, O. F. (2018). Characterising resource management performance in Kubernetes. Computers & Electrical Engineering, 68, 286-297.
[13] Goethals, T., De Turck, F., &Volckaert, B. (2020). Extending kubernetes clusters to low-resource edge devices using virtual kubelets. IEEE Transactions on Cloud Computing, 10(4), 2623-2636.
[14] Burns, B., & Tracey, C. (2018). Managing Kubernetes: operating Kubernetes clusters in the real world. O'Reilly Media.
[15] Goethals, T., De Turck, F., &Volckaert, B. (2020). Extending kubernetes clusters to low-resource edge devices using virtual kubelets. IEEE Transactions on Cloud Computing, 10(4), 2623-2636.
[16] Ramanan, M., Singh, L., Kumar, A. S., Suresh, A., Sampathkumar, A., Jain, V., &Bacanin, N. (2022). Secure blockchain enabled Cyber-Physical health systems using ensemble convolution neural network classification. Computers and Electrical Engineering, 101, 108058.
[17] Noorbakhsh, J., Farahmand, S., Foroughi Pour, A., Namburi, S., Caruana, D., Rimm, D., ...& Chuang, J. H. (2020). Deep learning-based cross-classifications reveal conserved spatial behaviors within tumor histological images. Nature communications, 11(1), 6367.
[18] Toka, L., Dobreff, G., Fodor, B., &Sonkoly, B. (2021). Machine learning-based scaling management for kubernetes edge clusters. IEEE Transactions on Network and Service Management, 18(1), 958-972.
[19] Zerka, F., Urovi, V., Vaidyanathan, A., Barakat, S., Leijenaar, R. T., Walsh, S., ...&Lambin, P. (2020). Blockchain for privacy preserving and trustworthy distributed machine learning in multicentric medical imaging (C-DistriM). Ieee Access, 8, 183939-183951.
[20] Nasir, M. U., Zubair, M., Ghazal, T. M., Khan, M. F., Ahmad, M., Rahman, A. U., ...&Mansoor, W. (2022). Kidney cancer prediction empowered with blockchain security using transfer learning. Sensors, 22(19), 7483.
[21] Nguyen, N. D., & Kim, T. (2021). Balanced Leader Distribution Algorithm in Kubernetes Clusters. Sensors, 21(3), 869.
[22] Durga, R., & Poovammal, E. (2022). FLED-Block: Federated Learning Ensembled Deep Learning Blockchain Model for COVID-19 Prediction. Frontiers in Public Health, 10.
[23] Kulkarni, P., Kanhere, A., Yi, P. H., & Parekh, V. S. (2022). From Competition to Collaboration: Making Toy Datasets on Kaggle Clinically Useful for Chest X-Ray Diagnosis Using Federated Learning. arXiv preprint arXiv:2211.06212.
[24] Charalambous, C. C., &Bharath, A. A. (2016). A data augmentation methodology for training machine/deep learning gait recognition algorithms. arXiv preprint arXiv:1610.07570.
[25] Qin, X., Zhang, Z., Huang, C., Dehghan, M., Zaiane, O. R., &Jagersand, M. (2020). U2-Net: Going deeper with nested U-structure for salient object detection. Pattern recognition, 106, 107404.
[26] Jorge-Martinez, D., Butt, S. A., Onyema, E. M., Chakraborty, C., Shaheen, Q., De-La-Hoz-Franco, E., &Ariza-Colpas, P. (2021). Artificial intelligence-based Kubernetes container for scheduling nodes of energy composition. International Journal of System Assurance Engineering and Management, 1-9.
[27] Sarwinda, D., Paradisa, R. H., Bustamam, A., &Anggia, P. (2021). Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer. Procedia Computer Science, 179, 423-431.
[28] Bernstein, D. (2014). Containers and cloud: From lxc to docker to kubernetes. IEEE cloud computing, 1(3), 81-84.
[29] Mateen, M., Wen, J., Song, S., & Huang, Z. (2018). Fundus image classification using VGG-19 architecture with PCA and SVD. Symmetry, 11(1), 1.
[30] Nihar Ranjan, Midhun C. (2020). A Brief Survey of Machine Learning Algorithms for Text Document Classification on Incremental Database. TEST Engineering and Management, ISSN: 0193-4120, Volume 83, 25246 – 25251.
[31] Nihar Ranjan, Midhun C. (2021). Evolutionary and Incremental Text Document Classifier using Deep Learning" International Journal of Grid and Distributed Computing Vol. 14, No. 1, 587-595.
[32] Nihar Ranjan, Zubair Ghouse. (2017). A Multi-function Robot for Military Application. Imperial Journal of Interdisciplinary Research (IJIR) Vol-3, Issue-3, ISSN: 2454-1362, 1785-1788.