Received: July 4, 2024. Revised: September 23, 2024.

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Zero-knowledge Proof Based Federated Learning with Blockchain for COVID-19 Classification

Parikshith Nayaka Sheetakallu Krishnaiah[1]* **Dayanand Lal Narayan[1]**

[1]*Department of Computer Science and Engineering,*
GITAM School of Technology, GITAM University, Bengaluru, India
* Corresponding author's Email: Pari2sn@gmail.com

Abstract: The diversity and scarcity of the medical information makes it difficult to create precise global classification approach for the healthcare applications. The main motive is the privacy issue that restricts the data exchanging scope between healthcare institutions. On the contrary, an information from single source is not adequate for developing the worldwide diagnosis approach. The Federated Learning (FL) is a promising solution for privacy and data multiplicity issues, an appropriate aggregation model for multi class and dissimilar medical information is still challenging task in the recognition. Moreover, the FL approaches does not effectively analyzes the each participant execution in the local model and secures the user data. In order to overcome this issue, the Zero-Knowledge Proof (ZKP) based FL approach is developed over blockchain (BC) for performing the COVID-19 classification. The global model of FL uses the two layer Long Short Term Memory (2LLSTM) with federated proximal term (FedProx) namely 2LLSTMFP while the Convolutional Neural Network (CNN) is used in the local model. The integration ZKP and BS is used to improve the data confidentiality while the immutability of BC helps to prevent unauthorized variations for the ledger. The developed FLBC-ZKP is analyzed with two datasets such as COVID-19 Radiography, and CXR images pneumonia and COVID-19. The FLBC-ZKP is evaluated using accuracy, recall, precision, specificity, F1-score, False Negative Rate (FNR) and False Positive Rate (FPR). The existing researches such as WMT, MCCF, 3SFDL and TOTL are used to compare the FLBC-ZKP method. The FLBC-ZKP achieves improved accuracy of 98.34 % for COVID-19 Radiography dataset that is better than the MCCF and 3SFDL.

Keywords: Blockchain, Convolutional neural network, COVID-19 classification, Federated learning, Federated proximal term, Two layer long short term memory, Zero-knowledge proof.

1. Introduction

Lung diseases or respiratory diseases are the important reasons of mortality and disability throughout the world. The most general lung disease comprises of pneumonia, tuberculosis and recent Coronavirus Disease 2019 (COVID-19) [1, 2]. In that, COVID-19 is initially appeared in Wuhan, China at early 2020 and it is spread to over 200 other countries and regions of the world [3, 4]. The typical symptoms which are linked with COVID-19 are fever, chest discomfort, myalgia, sore throat, headache and dry cough that may be characterized as a respiratory disease [5, 6]. The CXR images i.e., sensitive approach for detecting the COVID-19 and also other chest related issues [7]. CXR is a noninvasive, painless and powerful investigatory approach which carries an effectual respiratory disease data [8]. The manual diagnosis of COVID-19 needs a skilled radiologists and it is time consuming, error-prone, extensive and tedious process, because radiologists are estimated to analyze huge amount of COVID-19 patients [9]. The radiology includes decision-making under the situations of uncertainty, hence it always cannot generate dependable clarifications or reports [10].

A Computer-aided diagnosis (CAD) is considered as substantial approach because it creates the diagnoses as easier and minimizes an amount of misdiagnoses [11]. The Artificial Intelligence (AI) based identifying tool is mandatory for helping the

radiologist in discovering COVID-19 existences in a quicker, sudden and precise way. Else, the diseased persons cannot be discovered and quarantined as soon as practicable and doesn't receive an adequate treatments [12]. But, the medical organizations and possessors of CXR images are preferred training over their own medical data has strict privacy necessities that creates it challenging for medical organizations with lesser samples for training the model with predictable performance. These issue is overcome by using the transfer learning and FL where medical organizations uses the local data to train without performing the data centralization [13]. FL gathers the locally trained models from various sources and cooperatively train the global model through decentralized network. But, broadcasting the personal data is impossible, because of the lack of a privacy-preserving in health care centers. Therefore, the blockchain technology is used to overcome the security issues on the decentralized network [14]. The blockchain technology is used to enhance the reliability, transparency and system security [15].

The contributions are summarized as follows:

- The global and local learning are balanced by incorporating the FedProx (FP) with 2LLSTM while the BC is used security the privacy of the data. The ZKP used in the FLBC improves the data confidentiality while accomplishing local training process.
- In FLBS, the 2LLSTMFP is used as the global model and CNN is used as the local model. The developed FLBC-ZKP performs the cooperative and secure manner to accomplish the COVID-19 classification while preserving the trust and privacy of participants.

The remaining paper is categorized as: The related works are given in the section 2. Section 3 details classification of FLBC-ZKP based COVID-19 classification. Simulation results are given in section 4 while section 5 discusses the conclusion.

2. Related works

The related works about the COVID-19 classification using CXR images are given as follows:

Noman [15] presented the FL based classification for repository disease classification, whereas the BC was incorporated for confirming the privacy. The Weight Manipulation Technique (WMT) was used during the local model aggregation which equipped the model with less and inappropriate parameters contributing to lesser than the model with huge and appropriate parameter. Further, an authentic and diverse medical information were provided by using the adaptive incentive approach in BC. The balancing among global and local model was essential for further improving the classification.

Jangam, E [16] presented the Multi-Class Classification Framework (MCCF) for reducing either false negative or false positive in the computer aided detection or CAD respectively. Here, the stacked ensemble from pre-trained models and fully connected layers based on selection metric and systematic approach was created to decrease false negatives or false positives. The base classifier's diversity was mainly depends on the diverse set of false negatives or false positives. The security of participant data was not considered while doing the disease diagnosis.

Bayram, F. and Eleyan, A [17] developed the 3-Stream Fusion-based Deep Learning (3SFDL) approach to detect the Viral Pneumonia and COVID-19 in CXR images. The 3-stream Convolutional Neural Network (CNN) was used in the fusion model. The incorporated CNN was extracted and combined the features from Local Binary Pattern (LBP), Histograms of Oriented Gradients (HOG) and grayscale images, before performing the final decision. Further, the classification result was achieved in dense layer along with the Softmax activation function. However, the performance of 3SFDL was mainly depends on the database size whereas 3SFDL was offered better performance when the database was large.

Kumar, S. and Mallik, A [18] presented the Trained Output-based Transfer Learning (TOTL) to detect the COVID-19. A denoising, contrasting and segmentation were done to preprocess the CXR of the patients. Next, these preprocessed images were given to the various pre-trained transfer learning architectures such as VGG16, VGG19, ResNet50, ResNet50V2, Xception, MobileNet, InceptionV3 and InceptionResNetV2. Further, the outputs of these models were trained by deep neural network for achieving the improved performance. The TOTL was required to analyze with the large dataset for further enhancing the classification.

Paul, A [19] developed the inverted bell curve based ensemble of deep learning approach to discover the COVID-19. The principle of transfer learning was utilized for transferring and fine tuning the pretrained weights for enhancing the classification. This ensemble approach was provided worse classification for lower and higher learning rate. The artefacts exist in the CXR images were affected the classification. The ensemble approach was sensitive to weight factor, an incorrect

classification was occurred when there was less weight.

Aktas, K [20] analyzed the efficiency of the deep CNN in CHR classification over huge amount of images. Here, the InceptionV3 network was used to extract the features from the processed images. Moreover, this deep CNN was used to address the issues developed small dataset such as low generalization, bias between classes, low sample size and so on. The developed deep CNN was analyzed only less amount of classes during the recognition.

Li, Q [21] developed the multi-level residual feature fusion network (MLRFNet) for performing the multi-label classification. This MLRFNet was obtained the receptive field data over various lesion sizes and enhanced disease specific features on spatial positions for minimizing the intervention of inappropriate regions. The classification vectors were generated by multi-level residual feature classifier that completely utilizes the spatial position of disease at the feature map which used to enhance the classification. This work was failed to consider the dependencies among the labels while performing the recognition.

Priya, K.V. and Peter, J.D [22] presented the federated approach to identify the chest diseases. The performance was enhanced by using the adopted transfer learning with pre-trained network. Consequently, DenseNet121 was integrated for avoiding the issue of vanishing gradient and enhanced the feature propagation. However, the developed federated approach was failed to consider the data leakage of participants during the transaction.

3. FLBC-ZKP method

In this research, the classification of COVID-19 is performed using the FL with blockchain architecture. The developed FL considers the CNN as the local models and 2LLSTMFP as the global model for the classification. The FP incorporated in the 2LLSTMFP balances the global and local learning whereas blockchain incorporated for transparency and ZKP for privacy. Therefore, the developed FLBC-ZKP offered the cooperative and secure way to perform the COVID-19 classification while preserving the trust and privacy of participants. The architecture of FLBC-ZKP is shown in Fig. 1.

The steps processed in this FLBC-ZKP method are given as follows:

- At first, the global model of 2LLSTM is initialized with FedProx parameter. A decentralized network with local nodes (i.e., hospitals) are deployed with the input data.
- Every local model accomplishes classification using CNN for the given input. Here, the local model is trained for recognizing the pattern and features related to the classification.
- The training over the datasets is performed by local models where each model generated the model update according to the training process.
- The update of model i.e., CNN is given to the 2LLSTMFP via FL. The FP incorporate in the global model is ensured the CNN is aligned with the 2LLSTMFP that balanced the reliability and local relevance.
- The updates of model and its relevant information is recorded over the blockchain. Here, the ZKP is incorporated for ensuring the privacy where only the valid information is verified without disclosing the sensitive information.

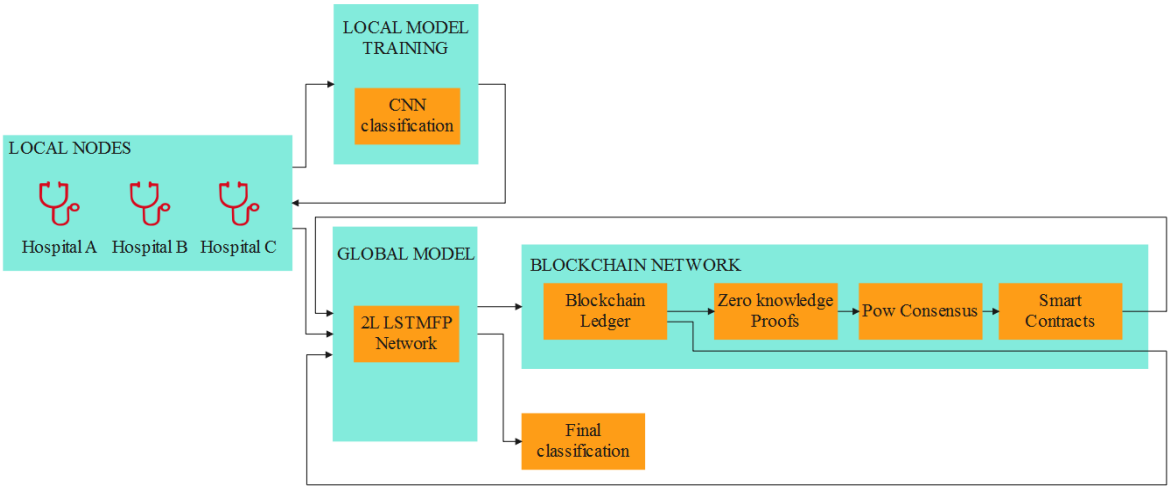

Figure. 1 Architecture of FLBC-ZKP

- The Proof of Work (PoW) calculations is performed by users for including the new blocks to the blockchain. Here, the PoW incentivizes participants and secured the blockchain against malicious behaviors.
- A model update confirmation and PoW solutions are automated by using Smart Contract (SC) on the blockchain. Here, an authenticate user gets the incentives via the SC operation.
- In this FL, the blockchain is decentralized and transparent ledger which performed as zero-trust approach. A model update from CNN is combined using 2LLSTMFP over the blockchain.
- The updated 2LLSTMFP is combined the predictions from CNN and this information is preserved blockchain. This 2LLSTMFP is considered as a reference for upcoming local training iterations.
- The FL iterates over the time and supports uninterrupted enhancement of 2LLSTMFP. Subsequently, the participants are uninterruptedly incentivized via ZKP and SC operations.
- The combination of ZKP with BC in FL is used to enhance the data confidentiality during local training process. Here, the blockchain's immutability is used to improve the security and prevent the unauthorized variations for the ledger. The ZKP secures the blockchain and includes an extra layer of trust.

3.1 Dataset acquisition

This research considered the two different CXR datasets for analysis of multi class classification. The two datasets are COVID-19 Radiography [23], and CXR images pneumonia and COVID-19 [24]. Here, the COVID-19 Radiography is huge dataset which has three different classes such as positive COVID-19, viral pneumonia and healthy images. Similarly, CXR images pneumonia and COVID-19 is small dataset that has less amount of images. The detailed information about datasets is shown in Table 1.

Table 1. Datasets information

Parameters	COVID-19 Radiography	CXR images pneumonia and COVID-19
Total images	15153	6118
Positive COVID-19	3616	262
Viral pneumonia	1345	1583
Healthy	10192	4273

3.2 ZKP based FLBC

The images from the datasets are taken as input to perform the classification.

The architecture of FL includes the local model (CNN) and global model (2LLSTMFP). The updates of model and its data is persevered in the blockchain. The ZKP is incorporated for ensuring the privacy where only the valid information is verified without disclosing the information. In FL, the sensitive information utilized for classifying the COVID-19 is kept private while contributing in training process. The incorporation of BC offers an immutable record of the proofs which helps in increasing the trust of FL process. BC preserves the outcomes of ZKP verifications offering a transparent and traceable log of all behaviors and proofs which used to maintain the integrity of FL. The ZKP has three different stages such as model distribution, training, and aggregation.

3.2.1. Model distribution

In this stage, the training tasks and models are processed and disseminated by publisher.

- The training approach F is separated as q identical pieces P by publisher.
- The piece P is converted into arithmetic circuit constraints R for $Groth16$ by publisher. Then it operates the $Groth16.Setup()$ for creating the Common Reference String $(\sigma;\tau)$ i.e., CRS which has proving key (pk) and verification key (vk) for both the ZKP creation and verification, where CRS has verification key is denoted by σ, trapdoor is denoted as τ
- The piece P and related Common Reference String (CRS) $(\sigma;\tau)$ and constraints R to n number of trainers are transmitted by publisher.

3.2.2. Model training

In this training stage, every trainer operates the training approach for training the local model and creating the ZKP for training process's correctness.

- At first, the local data d_i is used to train the model by i th trainer via operating the algorithm piece P for many rounds. Moreover, the inputs and outputs are kept as statements ϕ_i for each round.
- The $Groth16: Prove()$ algorithm is used for producing the ZKPs π_i of every round of training by ith trainer. Trainer operated the proof generation process for eliminating the

unmasking of training process 's intermediary inputs and outputs to the publisher while doing the proof verification. The $Groth16: Prove()$ algorithm modified the statement and verification key as ϕ_i' and vk_i' for hiding the information. The changes in the ϕ_i and vk is valid to produce the ZKP s_i^1 and s_i^2 as well as the outcome of earlier piece is the input to successive price, where s_i^1 and s_i^2 are the sigma proof. The statement ϕ_i is expressed in Eq. (1).

$$\phi_i = \{a_{i,j}\}_{j=1}^{l} \quad (1)$$

Where, number of elements in the statement is l and piece output is denoted as a.

- A ith trainer broadcasts all the proof and modified information to the publisher, once the training and proof generation are done. A last reformed statement is kept as the final outcome in the local model i.e., CNN.

3.2.3. Model aggregation

The proofs of publisher and m trainer are checked to operate the secure sum protocol for computing the global model (i.e., 2LLSTMFP) without exposing the CNN.

- The publisher initialized the $Groth16.Verify()$, when π_i, s_i^1, s_i^2, ϕ_i' and vk_i' are received. The $Groth16.Verify()$ is verified the proof π_i and initialized the proof verification for checking the s_i^1 and s_i^2 which used to confirm the perfect training process of ith Trainer.
- For performing the homomorphic encryption and randomly selecting the generator $g_{pub} \in G$, the public key pk_h and secret key sk_h are acquired by operating the $KenGen()$ approach using publisher. Subsequently, the pk and g_{pub} are transmitted to trainers who has checked and confirmed its proof. Moreover, the smart contract (SC) is generated and revealed to perform the final calculation of parameter aggregation.
- The random value $s_{i,i+1} \in Z_q$ is generated and broadcasted by ith trainer to trainer $i+1$. Further, the trainer m transmits to the 1st trainer.
- A encrypted value (c) is computed using Eq. (2) by i th trainer, where Enc denotes encryption. The proof generation of addition is operated for obtaining s_i^3 by ith trainer which argues that c_i has the ϕ_i'.

$$c_i = Enc_{pk}(\phi_i + s_i - s_{i-1}) \quad (2)$$

- A smart contract SC received the c_i and s_i^3 by ith trainer.
- The publisher measures the global model for decrypted value $(\bar{c})$ via SC as shown in Eq. (3), when the c_i and s^3 are received from ith trainer. A proof verification by summation is operated by publisher for checking the precision of models used in the aggregation of the 2LLSTMFP.

$$\bar{c} = \frac{Dec_{sk}(\sum_{i=1}^{m} c_i)}{n} \quad (3)$$

Where, Dec denotes decryption.

3.3 Prediction of classification probability using local model

In this phase, the data from the hospitals are processed using the local model i.e., CNN for obtaining the classification probability. CNN is developed to analyze the certain image I of size $w \times h \times d$ (width-height-deep) whereas the d value is based on the selected color model. For example, the Red-Green-Blue (RGB) model is defined on 3 components which denotes the colors, hence the dimension is equals to 3. This model has some variation in the layers when compared to the conventional structure of neural networks. CNN uses two different layers: The first layer is convolutional layer and it has the capacity of discovering essential features which is done by varying the image structure by using the filter k with a $p \times p$ matrix size as shown in Eq. (4).

$$k \times I_{x,y} = \sum_{i=1}^{p} \sum_{j=1}^{p} k_{i,j} . I_{x+i-1,y+j-1} + b_1 \quad (4)$$

Where, pixel's location over image I is denoted as $I_{x,y}$ and bias value is denoted as b. The image is transformed as feature maps via filter with the smaller size than the original of size $(p-1) \times (p-1)$. Additionally, if some computed values are negative i.e., outside of the color range, then the values are transformed by ReLu function according to Eq. (5).

$$R(x) = \max(0, x) \quad (5)$$

The outcome is processed by the successive layer i.e., pooling layer once the feature map is obtained in

CNN. This layer is required to minimize the size of image to reduce amount of computations and reduce the amount of inappropriate features which is accomplished by choosing a certain function $\theta(.)$ like maximum, minimum and so on. The function is analyzed the given pixel and its adjacent size $k \times k$ namely kernel. The investigation is performed by choosing only one point that highly satisfies this specific function.

The convolutional and pooling layers are incorporated in the CNN followed by fully connected layer is included with multiple hidden layers and one output. A neurons in the layers are linked with the adjacent neurons by synapse that denoted by weights using $w \in \langle 0,1 \rangle$. The function of such neuron $x_{i,j}$ that is denoted as shown in Eq. (6).

$$x_{i+1,k} = \tanh\left(\sum_{j=0}^{r-1} w_{i,j} x_{i,j}\right) + b_2 \quad (6)$$

Where, i is the layer; location is j; r is an amount of neurons in the preceding layer and bias is denoted as b_2. This CNN provides the vector $[y_0, y_1, \dots, y_{c-1}]$ where c is amount of classes.

3.4 Global model classification using 2LLSTMFP

The classification probabilities from the local model CNNs are given as input to the 2LLSTMFP to perform the final prediction of diseases. The FedProx loss function is included in the 2LLSTM to improve the classification. The developed 2LLSTM has two LSTM layers where FedProx is used to solve the issue of statistical heterogeneity. LSTM is used to learn the dependency relations and it has the ability to solve the complexity and synthetic long time lag operations. Each LSTM cell comprises the gates which has the ability to perform various operations in identifying the data from an earlier process for forgetting and remembering the parameters.

An overall 2LLSTM's stability is enhanced by using the proximal term i.e., FedProx. To contribute in each round update, a subgroup of devices are selected by FedProx. This update is used to ensure the local update and calculates an average for local update for generating the global updates. Eq. (7) expresses the FedProx objective function.

$$\underset{w}{\operatorname{argmin}}\, h_k(w, w^t) = F_k(w) + \frac{\mu}{2}\|w - w^t\|^2 \quad (7)$$

Where, objective function is denoted as h_k, device is denoted as k; weight value is w; local loss is F_k; time step is t; and hyperparameter is denoted as μ.

The pseudo code for FLBC-ZKP is given in Algorithm 1.

Algorithm 1: Pseudo code for FLBC-ZKP

Initialize and develop a global model i.e., 2LLSTMFP.
Decentralized network is initialized with local CNN model.
Set the FLBC parameters such as number of epochs, criterion.
For each epoch in number of epochs:
For each local node in FLBC.
Local model (CNN) is trained in CXR data.
Transmit local model to local node.
Local model is updated and appended to CNN update.
Retrieve the CNN from local node
End for
Combine the CNN using 2LLSTMFP.
Transmit 2LLSTMFP to global worker.
Update 2LLSTMFP using FP.
Transmit 2LLSTMFP to local node.
Update CNN using 2LLSTMFP.
Generate the ZKP for 2LLSTMFP update.
Submit the ZKP to blockchain.
End for

4. Simulation results and discussion

In this section, a various evaluation is performed to evaluate the efficiency of FLBC-ZKP for COVID-19 classification. Two different datasets such as COVID-19 Radiography, and CXR images pneumonia and COVID-19 are considered to analyze the FLBC-ZKP. The ratio of 20:80 is used to perform the testing and training process.

4.1 Experimental setup and evaluation metrics

The implementation is done in the system with intel core i7 processor, 16GB RAM and Windows 10 operating system. Here, the analysis is done by using the Python 3.7 language over Anaconda Navigator. The parameters of accuracy, recall, precision, specificity, F1-score, FNR and FPR expressed in Eqs. (8) to (14) are utilized to evaluate the FLBC-ZKP.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (8)$$

$$Recall = \frac{TP}{TP+FN} \quad (9)$$

$$Precision = \frac{TP}{TP+FP} \quad (10)$$

$$Specificity = \frac{TN}{TN+FP} \quad (11)$$

$$F1 - score = \frac{2TP}{2TP+FP+FN} \quad (12)$$

$$FNR = \frac{FN}{FN+TP} \quad (13)$$

$$FPR = \frac{FP}{TN+FP} \quad (14)$$

Where, TP and TN expresses the true positive and true negative; FP and FN expresses the false positive and false negative.

4.2 Ablation study

The ablation study is performed for 2LLSTMFP utilized in the FLBC-ZKP. Hence, a different optimizer along with different learning rate are considered in the ablation study for analyzing the performances with different layers. The optimizers of AdaGrad, RMSprop, SGD and NADAM are used for the evaluation. The performances are evaluated using validation accuracy by (VAcc), validation loss (VLoss), test accuracy by (TAcc) and test loss (TLoss). The evaluation shown in Table 2 shows the performances of 2LLSTMFP for COVID-19 Radiography dataset. The NADAM provides better performance in most of the cases than the other optimizers because of its improved stability achieved during the classification.

Additionally, the evaluation shown in Table 3 shows the performances of 2LLSTMFP for CXR images pneumonia and COVID-19 dataset. This table confirmed that the NADAM optimizer mostly gives enhanced performances than the other optimizers.

Table 2. Analysis of 2LLSTMFP for COVID-19 Radiography dataset

Optimizer	Learning Rate	VAcc (%)	VLoss (%)	TAcc (%)	TLoss (%)
AdaGrad	0.00000001	96.442	0.052	95.254	0.978
	0.000001	96.687	0.122	96.131	0.968
	0.0000001	96.266	0.131	96.139	0.753
	0.00001	96.217	0.177	96.220	0.724
RMSprop	0.00000001	97.611	0.197	96.382	0.516
	0.000001	96.674	0.313	96.355	0.479
	0.0000001	97.851	0.343	96.891	0.440
	0.00001	97.905	0.623	96.414	0.432
SGD	0.00000001	97.970	0.647	96.916	0.428
	0.000001	97.926	0.801	96.463	0.427
	0.0000001	97.118	0.810	97.822	0.351
	0.00001	97.463	0.830	97.213	0.337
NADAM	0.00000001	97.552	0.926	97.895	0.219
	0.000001	97.226	0.939	97.083	0.180
	0.0000001	97.848	0.943	97.646	0.031
	0.00001	97.727	0.956	97.750	0.024

Table 3. Analysis of 2LLSTMFP for CXR images pneumonia and COVID-19 dataset

Optimizer	Learning Rate	VAcc (%)	VLoss (%)	TAcc (%)	TLoss (%)
AdaGrad	0.00000001	94.173	0.157	95.477	0.137
	0.000001	94.375	0.156	95.663	0.136
	0.0000001	94.433	0.152	95.693	0.134
	0.00001	94.495	0.144	95.838	0.130
RMSprop	0.00000001	94.625	0.137	95.936	0.128
	0.000001	94.660	0.135	95.957	0.128
	0.0000001	94.785	0.131	96.009	0.098
	0.00001	94.799	0.121	96.118	0.092
SGD	0.00000001	94.810	0.120	96.123	0.079
	0.000001	94.873	0.115	96.191	0.078
	0.0000001	94.973	0.089	96.365	0.076
	0.00001	95.042	0.088	96.386	0.061
NADAM	0.00000001	95.058	0.088	96.446	0.059
	0.000001	95.328	0.086	96.574	0.058
	0.0000001	95.339	0.083	96.685	0.051
	0.00001	95.504	0.072	96.910	0.043

Table 4. Performances of 2LLSTMFP for COVID-19 Radiography dataset

Class	Precision (%)	Recall (%)	Specificity (%)	F1-Score (%)	FPR	FNR
Positive COVID-19	**97.45**	**98.31**	**95.36**	**97.48**	**0.37**	**0.09**
Viral pneumonia	95.12	94.67	96.48	94.89	0.99	0.77
Healthy	96.78	95.48	97.45	95.78	0.51	0.63
Average	93.42	91.56	98.31	92.48	0.62	0.50

Table 5. Analysis of classifiers with different configuration for COVID-19 Radiography dataset

Configuration	Classifiers	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Individual classifier	CNN	67.45	57.9	53.34	55.52
	LSTM	78.57	72.83	68.09	70.38
	2LLSTMFP-CNN	80.45	74.89	71.48	73.14
Classifier with FL	CNN	91.67	87.98	85.05	86.49
	LSTM	94.28	80.58	80.53	80.55
	2LLSTMFP-CNN	98.34	93.42	91.56	92.48

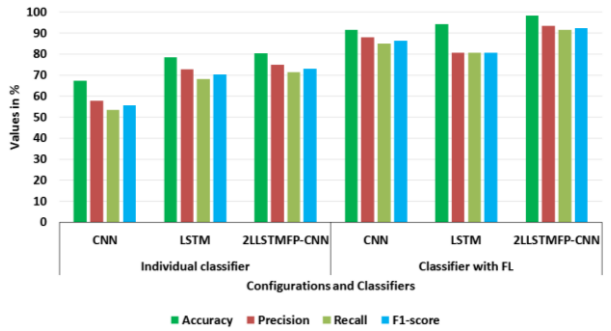

Figure. 2 Graph of classifiers with different configuration for COVID-19 Radiography dataset

4.3 Analysis of FLBC-ZKP for COVID-19 radiography dataset

The evaluation carried out for each class in the COVID-19 Radiography dataset is shown in Table 4 for 2LLSTMFP. This evaluation shows that how well the classifiers in the FLBC performs classification. The 2LLSTMFP of FLBC has the average specificity of 98.31%, it shows that overall FLBC reduces false positives during the classification.

The evaluation shown in the Table 5 and Fig. 2 shows the evaluation of different configuration of classifier. Here, the performance is analyzed for individual classifier and classifier with FL. The developed 2LLSTMFP-CNN outperforms better even when it is processed with as individual classifier or as a FL. The statistical heterogeneity issue solved by the FP integrated in 2LLSTM helps to improve diagnosis. Moreover, the dependency relations learning using 2LLSTMFP additionally enhances the diagnosis. The incorporation of FL makes the model to be trained with data from diverse sources that possibly obtaining a wider range of patterns and representations that is beneficial in enhancing the diagnosis in next level. Moreover, the performances of ZKP used with FLBC is shown in Table 6.

Table 6. Analysis for ZKP for COVID-19 Radiography dataset

Parameters	Number of epochs						
	1	**2**	**3**	**5**	**10**	**15**	**20**
SetupTime	403.81	404.72	408.32	410.41	413.83	416.28	418.04
Proofgeneration time	53.23	55.59	57.72	59.82	62.24	64.89	67.25
Proofverification time	1.25	1.5	1.7	1.99	2.24	2.46	2.73
Proof size (kb)	6	6.24	6.48	6.76	7.02	7.22	7.45

Table 7. Performances of 2LLSTMFP for CXR images pneumonia and COVID-19 dataset

Class	Precision (%)	Recall (%)	Specificity (%)	F1-Score (%)	FPR	FNR
Positive COVID-19	97.10	95.00	98.10	96.81	0.05	0.09
Viral pneumonia	95.73	93.94	96.96	95.67	0.05	0.09
Healthy	96.74	94.69	97.95	96.62	0.07	0.07

Table 8. Analysis of classifiers with different configuration for CXR images pneumonia and COVID-19 dataset

Configuration	Classifiers	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Individual classifier	CNN	78.54	71.42	65.32	68.23
	LSTM	85.31	81.03	76.15	78.51
	2LLSTMFP-CNN	90.05	87.67	81.32	84.37
Classifier with FL	CNN	89.45	76.94	98.45	92.00
	LSTM	90.56	86.42	97.34	94.45
	2LLSTMFP-CNN	99.07	95.02	99.67	97.25

Table 9. Analysis for ZKP for CXR images pneumonia and COVID-19 dataset

Parameters	Number of epochs						
	1	2	3	5	10	15	20
SetupTime	402.53	405.17	408.02	410.15	412.31	414.77	417.11
Proofgeneration time	53.23	55.47	58.05	60.56	63.08	65.42	67.77
Proofverification time	1.25	1.46	1.66	1.87	2.14	2.42	2.65
Proof size (kb)	6	6.28	6.55	6.85	7.09	7.34	7.61

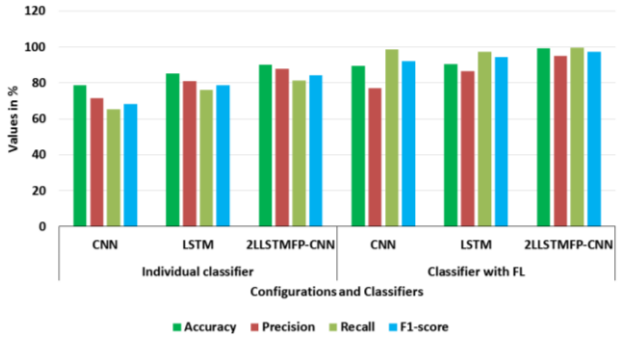

Figure. 3 Graph of classifiers with different configuration for CXR images pneumonia and COVID-19 dataset

4.4 Analysis of FLBC-ZKP for CXR images pneumonia and COVID-19 dataset

The evaluation carried out for each class in the CXR images pneumonia and COVID-19 dataset is shown in Table 7 for 2LLSTMFP. This evaluation shows that how well the 2LLSTMFP in the FLBC performs classification. The 2LLSTMFP of FLBC achieves an average specificity of 97.67%, it confirmed that FLBC reduces false positives while classifying the diseases.

The evaluation shown in the Table 8 and Fig. 3 shows the evaluation of different configuration of classifier. Here, the performance is analyzed for individual classifier and classifier with FL. The 2LLSTMFP-CNN has improved performance when it is processed with as individual classifier or as a FL. The disease diagnosis is improved by solving the statistical heterogeneity issue based on the FP term integrated in 2LLSTM. The FL makes the model to be trained with data from diverse sources that possibly obtaining a wider range of patterns and representations that is beneficial in enhancing the diagnosis in further level. Moreover, the performances of ZKP used with FLBC is shown in Table 9.

4.5 Analysis of FLBC-ZKP for CXR images pneumonia and COVID-19 dataset

The existing researches such as WMT [15], MCCF [16], 3SFDL [17] and TOTL [18] are used to compare the FLBC-ZKP. The comparison carried out for COVID-19 Radiography is shown in Table 10 and for CXR images pneumonia and COVID-19 dataset is shown in Table 11. Similarly, a Fig.4 and Fig.5 shows the graph comparison for accuracy. This comparison confirmed that the FLBC-ZKP has enhanced performance than the WMT [15], MCCF [16], 3SFDL [17] and TOTL [18]. The accuracy of FLBC-ZKP for COVID-19 Radiography dataset is 98.34 % that is better than the MCCF [16] and 3SFDL [17]. The FP integrated in the 2LLSTM supports an appropriate aggregation of the local CNN model, even when trained using heterogeneous data distributions. The global model i.e., 2LLSTMFP of FL combines the updates from local CNN model. The ZKP confirms the updates are authorized and trustworthy without sharing information in BC, and hence each user's privacy is maintained during the secure model aggregation using 2LLSTMFP. The performance of FLBC-ZKP is enhanced by balancing learning among the global and local model. The disease classification is improved by solving the statistical heterogeneity issue based on the FP term of 2LLSTMFP.

Table 10. Comparison for COVID-19 Radiography dataset

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
MCCF [16]	97.67	98.36	97	97.68
3SFDL [17]	97.76	97.78	97.76	97.76
FLBC-ZKP	98.34	93.42	91.56	92.48

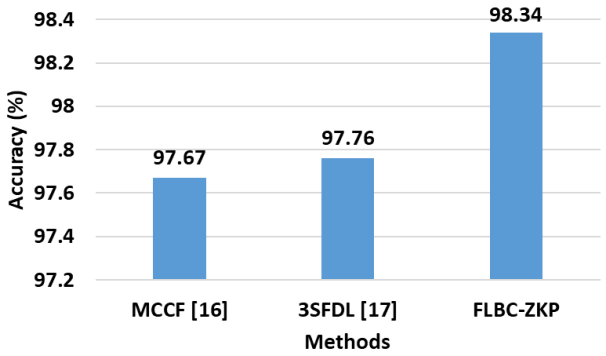

Figure. 4 Comparison graph of accuracy for COVID-19 Radiography dataset

Table 11. Comparison for CXR images pneumonia and COVID-19 dataset

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
WMT [15]	88.10	NA	NA	NA
MCCF [16]	97.33	97.37	97.33	97.36
TOTL [18]	96.47	97.81	95.71	96.75
FLBC-ZKP	99.07	95.02	99.67	97.25

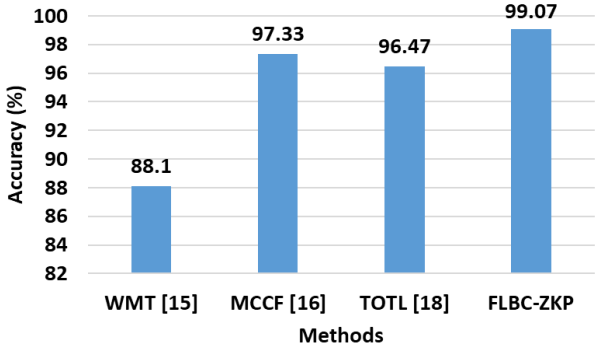

Figure. 5 Comparison graph of accuracy for CXR images pneumonia and COVID-19 dataset

5. Conclusion

This paper uses the ZKP in the FLBS for performing the cooperative and secure way of COVID-19 classification while maintaining the trust and privacy of participants. The classification probabilities from the CXR images are found initially by using the CNN models in the local models. Next, the 2LLSTMFP is used as the global model for performing the final classification. The FP incorporated in the 2LLSTM balances the global and local learning whereas blockchain incorporated for transparency and ZKP for privacy. The ZKP used in the FLBS enhanced the data confidentiality during local training process. Further, the security is improved and an unauthorized variation for the ledger are prevented using ledger. The simulated extensive experiments confirmed that the FLBC-ZKP provides better performance than the WMT, MCCF, 3SFDL and TOTL. The FLBC-ZKP achieves improved accuracy of 98.34 % for COVID-19 Radiography dataset that is better than the MCCF and 3SFDL.

Conflicts of Interest

The authors declare no conflict of interest.

Author Contributions

Conceptualization, PNSK and DLN; methodology, PNSK; software, DLN; validation, DLN; formal analysis, PNSK; investigation, DLN; resources, DLN; data curation, DLN; writing—original draft preparation, PNSK; writing—review and editing, DLN; visualization, PNSK; supervision, DLN; project administration, DLN.

Notation:

Parameter	Description
F	Training approach
P	Piece
q	Number of pieces
vk	Verification key
pk	Proving key
R	Constraints
n	Number of trainers
d_i	Data
ϕ_i	Statement
vk_i'	Modified verification key
ϕ_i'	Modified statement
s_i^1 and s_i^2	Sigma proof
l	Number of elements in the statement
a	Piece output
$g_{pub} \in G$	Generator
$s_{i,i+1} \in Z_q$	Random value
c	Encrypted value
Enc	Encryption
$\bar{c}$	Decrypted value
Dec	Decryption
$w \times h \times d$	Width-height-deep
I	Image

d	Value based on the selected color model
k	Filter
$p \times p$	Matrix size
$I_{x,y}$	Pixel's location over image I
b and b_2	Bias value
$R(x)$	ReLu function
$\theta(.)$	Function
r	Number of neurons in the preceding layer
i	Layer
j	Location
w	Weight
h_k	Objective function
k	Device
F_k	Local loss
t	Time step
μ	Hyperparameter
FNR	False negative rate
FPR	False positive rate
TP	True positive
TN	True negative
FP	False positive
FN	False negative

References

[1] P. Yadav, N. Menon, V. Ravi, and S. Vishvanathan, "Lung-GANs: Unsupervised Representation Learning for Lung Disease Classification Using Chest CT and X-Ray Images", *IEEE Transactions on Engineering Management*, Vol. 70, No. 8, pp. 2774-2786, 2023.

[2] M. K. Islam, M. M. Rahman, M. S. Ali, S. M. Mahim, and M.S. Miah, "Enhancing lung abnormalities detection and classification using a Deep Convolutional Neural Network and GRU with explainable AI: A promising approach for accurate diagnosis", *Machine Learning with Applications*, Vol. 14, p. 100492, 2023.

[3] S. Hallaci, B. Farou, Z. Kouahla, and H. Seridi, "New approach based on light enhancement and real-time dual CNN for classification of COVID-19 X-ray images", *Evolving Systems*, 2023.

[4] R. Kumar, R. Arora, V. Bansal, V.J. Sahayasheela, H. Buckchash, J. Imran, N. Narayanan, G.N. Pandian, and B. Raman, "Classification of COVID-19 from chest x-ray images using deep features and correlation coefficient", *Multimedia Tools and Applications*, Vol. 81, No. 19, pp. 27631-27655, 2022.

[5] H. Malik, T. Anees, M.U. Chaudhry, R. Gono, M. Jasiński, Z. Leonowicz, and P. Bernat, "A Novel Fusion Model of Hand-Crafted Features with Deep Convolutional Neural Networks for Classification of Several Chest Diseases using X-ray Images", *IEEE Access*, Vol. 11, pp. 39243-39268, 2023.

[6] N. Kumar, M. Gupta, D. Gupta, and S. Tiwari, "Novel deep transfer learning model for COVID-19 patient detection using X-ray chest images", *Journal of Ambient Intelligence and Humanized Computing*, Vol. 14, No. 1, pp. 469-478, 2023.

[7] L. Kong, and J. Cheng, "Classification and detection of COVID-19 X-Ray images based on DenseNet and VGG16 feature fusion", *Biomedical Signal Processing and Control*, Vol. 77, p. 103772, 2022.

[8] M. Gazda, J. Plavka, J. Gazda, and P. Drotár, "Self-Supervised Deep Convolutional Neural Network for Chest X-Ray Classification", *IEEE Access*, Vol. 9, pp. 151972-151982, 2021.

[9] A.M. Ayalew, A.O. Salau, B.T. Abeje, and B. Enyew, "Detection and classification of COVID-19 disease from X-ray images using convolutional neural networks and histogram of oriented gradients", *Biomedical Signal Processing and Control*, Vol. 74, p. 103530, 2022.

[10] I. Allaouzi, and M.B. Ahmed, "A Novel Approach for Multi-Label Chest X-Ray Classification of Common Thorax Diseases", *IEEE Access*, Vol. 7, pp. 64279-64288, 2019.

[11] A. Gopatoti, and P. Vijayalakshmi, "CXGNet: A tri-phase chest X-ray image classification for COVID-19 diagnosis using deep CNN with enhanced grey-wolf optimizer", *Biomedical Signal Processing and Control*, Vol. 77, p. 103860, 2022.

[12] Y.H. Bhosale, and K.S. Patnaik, "PulDi-COVID: Chronic obstructive pulmonary (lung) diseases with COVID-19 classification using ensemble deep convolutional neural network from chest X-ray images to minimize severity and mortality rates", *Biomedical Signal Processing and Control*, Vol. 81, p. 104445, 2023.

[13] Z. Li, X. Xu, X. Cao, W. Liu, Y. Zhang, D. Chen, and H. Dai, "Integrated CNN and Federated Learning for COVID-19 Detection on Chest X-Ray Images", *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 2022.

[14] R. Kumar, A.A. Khan, J. Kumar, Zakria, N.A. Golilarz, S. Zhang, Y. Ting, C. Zheng, and W. Wang, "Blockchain-Federated-Learning and Deep Learning Models for COVID-19

Detection Using CT Imaging", *IEEE Sensors Journal*, Vol. 2, No. 14, pp. 16301-16314, 2021.
[15] A.A. Noman, M. Rahaman, T.H. Pranto, and R.M. Rahman, "Blockchain for medical collaboration: A federated learning-based approach for multi-class respiratory disease classification", *Healthcare Analytics*, Vol. 3, p. 100135, 2023.
[16] E. Jangam, C.S.R. Annavarapu, and A.A.D. Barreto, "A multi-class classification framework for disease screening and disease diagnosis of COVID-19 from chest X-ray images", *Multimedia Tools and Applications*, Vol. 82, No. 10, pp. 14367-14401, 2023.
[17] F. Bayram, and A. Eleyan, "COVID-19 detection on chest radiographs using feature fusion based deep learning", *Signal, Image and Video Processing*, Vol. 16, No. 6, pp. 1455-1462, 2022.
[18] S. Kumar, and A. Mallik, "Covid-19 detection from chest x-rays using trained output based transfer learning approach", *Neural Processing Letters*, Vol. 55, pp. 2405-2428, 2023.
[19] A. Paul, A. Basu, M. Mahmud, M.S. Kaiser, and R. Sarkar, "Inverted bell-curve-based ensemble of deep learning models for detection of COVID-19 from chest X-rays", *Neural Computing and Applications*, Vol. 35, No. 22, pp. 16113-16127, 2023.
[20] K. Aktas, V. Ignjatovic, D. Ilic, M. Marjanovic, and G. Anbarjafari, "Deep convolutional neural networks for detection of abnormalities in chest X-rays trained on the very large dataset", *Signal, Image and Video Processing*, Vol. 17, No. 4, pp. 1035-1041, 2023.
[21] Q. Li, Y. Lai, M.J. Adamu, L. Qu, J. Nie, and W. Nie, "Multi-Level Residual Feature Fusion Network for Thoracic Disease Classification in Chest X-Ray Images", *IEEE Access*, Vol. 11, pp. 40988-41002, 2023.
[22] K.V. Priya, and J.D. Peter, "A federated approach for detecting the chest diseases using DenseNet for multi-label classification", *Complex & Intelligent Systems*, Vol. 8, No. 4, pp. 3121-3129, 2022.
[23] COVID-19 Radiography Dataset: https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database/data
[24] CXR images pneumonia and COVID-19 Dataset: https://www.kaggle.com/datasets/prashant268/chest-xray-covid19-pneumonia/data