Federated Learning aided Breast Cancer Detection with Intelligent Heuristic-based Deep Learning Framework

atul b kathole (atul.kathole@pccoepune.org)
PCCOE

Research Article

Keywords: Federated Learning aided Breast Cancer Detection, Densenet architecture, Enhanced Recurrent Neural Networks, Hybrid Dragon-Rider Optimization

Posted Date: June 16th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1764898/v1

Federated Learning aided Breast Cancer Detection with Intelligent Heuristic-based Deep Learning Framework

*Atul B.Kathole
*Assistant Professor,PCCOE,Pune
*atul.kathole@pccoepune.org

Savita Kumbhare
Assistant Professor,PCCOE,Pune
savita.kumbhare@pccoepune.org

Swati Shinde
Professor,PCCOE,Pune
swati.shinde@pccoepune.org

***Abstract*-Breast cancer is the second largest cause of female cancer death and one of the most hazardous diseases that leads to a higher mortality rate. Breast cancer is initialized with the malignant stage, where the abnormal growth of cancerous lumps is initiated from the breast cells. Periodic clinical checks and self-tests assist in early identification and thus progress the survival rates considerably. One of the eminent medical approaches is breast cancer recognition, which offers scientists and researchers huge complications. Breast cancer detection at an early stage permits the patients to receive suitable treatment, which increases the chances of survival. Thus, this paper utilizes a new form of artificial intelligence training called Federated Learning (FL), especially for breast cancer detection, the most eminent technique in the last few years. FL permits individual hospitals to benefit from the rich datasets of multiple non-affiliated hospitals without centralizing the data in one place. Hence, FL utilizes numerous collaborators for building a strong deep-learning model using a large dataset. In this paper, a hybridization of this type of training with a meta-heuristic and deep learning is aimed to be proposed for breast cancer diagnosis. This model encloses diverse steps that include (a) image collection, (b) feature extraction, and (c) classification phase. Initially, the mammogram images related to breast cancer are collected with the concept of FL from the affected individuals. The federated learning helps in reducing the processing time and ensures better performance of the proposed model. The obtained images are considered for the feature extraction phase. The Densenet architecture is used to extract the features used in the classification phase with the help of Enhanced Recurrent Neural Networks (E-RNN) for detecting breast cancer. Here, the performance is enhanced by tuning the certain parameter in the RNN network using a hybrid optimization algorithm called Hybrid Dragon-Rider Optimization (HDRO) with Dragonfly Algorithm (DA) and Red deer algorithm (RDA) to achieve accurate classification results. The experimental results demonstrate the effectiveness of the suggested breast cancer diagnosis model compared with conventional approaches using diverse quantitative measures.**

Keywords— Federated Learning aided Breast Cancer Detection; Densenet architecture; Enhanced Recurrent Neural Networks; Hybrid Dragon-Rider Optimization;

I. INTRODUCTION

Breast cancer is a type of nasty growth that begins from breast tissue, generally in the interior coating of the milk ducts or "breast lobules" and spreads throughout the body parts [9][10]. Breast cancer is the second most spreading disease in the world. Hence it is essential to find the number of death due to "breast cancer" in advance for the drop and thriving treatments. After the analysis process, it was found that there has been a considerable rise in breast cancer in the last ten years [11]. For the diagnosis and treatment process, the breast cancer medical imaging method will be helpful for doctors to look into a person's body [12]. With the existing imaging process, breast cancer could be diagnosed timely; this image could be malevolent. There are chances of massive threat while placing the cancer cell in interstitial tissue vein or fluid till the microscope examination of tissues was conducted to ensure the spiteful begging [13]. Rising cancer chances are higher when the cell moves through a needle route or operates intelligently [14]. To remove pectoral muscles in breast cancer diagnosis, the specialized medical image is pre-processed to enhance detecting process. Thus, the search for abnormalities can be confined to the breast profile region by removing the background of the image and pectoral muscle [15].

However, the classic diagnosis method is generally used to enhance the correctness. The handcrafted elements may cause deviation in the diagnosis correctness; an experienced doctor plays a vital role in manual feature extraction [16]. Based on a doctor's experiences, usually prejudiced features such as texture, morphology, density, and other characteristics are physically used [17]. The deep learning technique is used in the last few years, which could extract hierarchical features manually with the help of image data. It has shown immense progress in applications, namely natural language, image, and speech recognition [19]. The object features overlook artificial experiences, and the subjectivist overlooks vital image elements [20]. Hence, from the above, we conclude to replicate the essential property of image and artificial experience by merging the objectives and subjectivist features [18].

It is crucial to observe the chances of hybridizations with the help of a heuristic algorithm and the existing state of machine learning [21]. In the previous year, excellent progressions have been made in artificial intelligence training methods. Multi-instance has been suggested that there are no similarities to classic training using a single instant. The final method is aggregating

and resenting for the next round of training, this concept is called federated learning. The private database is one the vital benefit of this learning [22]. To re-train the model, every instant has its own data set. Likewise, every instance is trained in the small set of data, so that the calculation can be done in a short period [23]. Unluckily, less protection of exchanged data is still a drawback [24]. While training the instant model, the aggregations are delivered. Sometimes, it's not safe to deliver. The same problem is faced during the training process; it is firm to control data. There is both advantage and disadvantage to federate learning. Still, there is more hope in finding the resolution by implementing the advanced innovation to train the classifier [25].

The "objectives of the designed model" are given here.

- To suggest new FL-based "breast cancer detection" with an intelligent heuristic-based deep learning framework with the aim of maximizing the diagnosis results in terms of maximizing the precision and accuracy.
- To gather the mammogram images regarding breast cancer utilizing the FL from the affected individuals for minimizing the processing time and superior efficiency.
- To classify the Densenet-based features through E-RNN to detect breast cancer using HDRO to get precise experimental results to assist medical professionals in offering timely treatment.
- To propose a HDRO algorithm for implementing a new enhanced RNN to optimize the number of epoch and hidden neurons in RNN to maximize the precision and accuracy.
- To evaluate the performance of the designed model over other traditional "approaches in terms of various performance measures" for demonstrating the superiority.

The rest of the paper is evaluated here. Section II explains the existing works. Section III implements a FL concept for enhanced "breast cancer diagnosis using mammogram images". Section IV specifies the dataset preparation, architecture description, and deep learning-based feature extraction. Section V designs the enhanced RNN with hybrid dragon-rider optimization for breast cancer diagnosis. Section VI "evaluates the performance. Section VII finishes this paper".

II. LITERATURE SURVEY

A. Related Works

In 2021, Houby and Yassin [1] implemented a "system using deep learning to categorize breast lesions" in ultrasonography images into malignant and non-malignant built on two outlooks: the whole image and the other by using a patch of the region of interest. This work has two phases: the classification phase using the "Convolution Neural Network (CNN) and pre-processing" step. The pre-processing stage was included image enhancement, region of interest extract, formatted unified noise removal, image resizing, and augmentation. The CNN enhances the current model from scratches to learn features to categorize the breast lesions in images. However, the recent work was classified in a reasonable rate, which showed a huge hopeful solution.

In 2021, Feki *et al.* [2] found a technique using deep learning method to collaborate federate framework, which allowed multi-medical institutes to take X-rays of a human chest for covid screening without data leak. The proposed system studied various vital points and specifics of federated learning, including the non-identity distribution and unbalanced data distributions. This proposed paper has observed and showed the demo of the federated framework. Also, it produced an aggressive solution for the trained model using two various architecture models and sharing data. Therefore, this could be helpful for the medical field to embrace the collaboration process, which attained benefits for private data that would lead to building an enhanced model for covid screening.

In 2021, Hirra *et al.* [3] developed a deep belief network that a "novel patch-based deep learning technique named Pa-DBN-BC" classified breast cancer on histopathology images. The supervised fine-tune phase and unsupervised pre-trained phase extracted features through this phase. From image patches, the features are extracted anonymously from the network. The histopathology images were classified from logistic regressions. The model has taken the extracted feature as an input, and the solution produced by the probability matrix was the positive or negative sample. With the help of four data sections, the current model was trained and tested using a histopathology image dataset, which resulted in colossal accuracy. As a result, the current system is more advanced than the existing one, as this model automatically produces the best solution.

In 2019, Wang *et al.* [4] found a method focusing on CNN to detect breast cancer. This work consists of three-phase; firstly, using CNN in-depth feature and unsupervised Extreme Learning Machine (ELM) clusters were based on a mass detection method. Secondly, they assembled and merged some feature sets such as texture, density, and morphological features. Thirdly, a combined set of ELM classifiers was created so that it's affectionate and spiteful breast masses. Hence, this proposed system produced an enhanced and effective process for detecting breast cancer.

In 2020, Zheng *et al.* [5] developed a " deep learning technique for detecting" breast cancer, using mathematical propose with improved computerized technique. In addition to previous computer vision, the tumor classified technique was actively transferred and improvised using deep CNN. Authors experimented with deep learning for various diagnosis predictions, transferred learning to characterize breast cancer for various diagnoses, prognosis, predicted tasks, and several image modals such as Ultrasound, mammography, and Magnetic Resonance Image (MRI) digital breast tomosynthesis. Max-pooling and LSTM

layers were the convolutions of deep learning. The softmax and "fully connected layers" were used to perform the error removal and classification. The proposed system aimed on merging the machine learning with process of selecting features and gathered them "through evaluating their results" using classified and segmented techniques to proceed towards the suitable output. The proposed model showed a huge accuracy level, specificity, and sensitivity compared with other breast cancer detection models.

In 2021, Połap and Woźniak [6] have implemented a meta-heuristic with hybrid type of training. It was helpful to reduce the attack on best model and managed the entire process using meta-heuristic algorithm. It has mainly utilized on lowering the "error of the model", while controlling the mechanisms for incoming model. Using CNN and meta-heuristic algorithm, the image classification result was observed with the effective solution. It was observed, that the safety against poison attack and the correctness of this model was better. The correctness of detecting poison sample in database was higher than optimized algorithm.

.

In 2021, Priya and Peter [7] founded a technique named DenseNets to detect different chest diseases. This model used the image dataset for chest X-rays from the Kaggle repository. In terms of different evaluation metrics, the new models were tested with complete datasets and both samples of chest X-ray. To enhance the performance, adoption of transfer learning was done with the pertained network named DenseNet121 and also it helped to boost up the feature propagation, vanish gradient issue, and also lesser the number of arguments. Hence, this system helped doctors and specialists identify various diseases from a single chest X-ray.

In 2021, Saber *et al.* [8] had developed a new technique named Transfer Learning (TL), which helped to automatically detect and diagnose suspected regions that were cross-validation and 80-20 techniques. This technique got the idea for solving the other relevant problem. The features were extracted from the Mammographic Image Analysis-Society (MIAS)data set using a pre-trained deep learning network in the existing system. Hence, it is noticed that VGG16 model was more powerful for diagnosing breast cancer through the classification of mammogram breast image.

B. Problem statement

Breast cancer detection is a serious healthcare problem that makes the affected individuals fall to death. Thus, many deep learning and federated learning are used in the literature for detecting breast cancer, and some of them are listed in Table 1. CNN [1] supports radiologists and physicians for breast cancer diagnosis and reduces the possibility of the biopsies process. However, it is necessary to investigate more validations in terms of accuracy before applying them to practical applications. VGG16 and ResNet50 [2] contain high practical values and can detect the early stage of breast cancer at the right time. But, it does not solve the problem of data imbalance and requires more time for the training process. DBN [3] efficiently solves the class "efficiency of optimizing the networks on the imbalanced dataset".

On the other hand, it cannot be applied to standard imbalanced datasets and various medical image analyses. US-ELM [4] simplifies the multi-optimization problem and enhances performance in terms of high accuracy. Yet, it is complex to interpret, and it won't work well when the data is strongly non-linear. DLA-EABA [5] recognizes the ROI mapping in the embedding process without any help of additional information. But, it is difficult to train the data, and the computation of the network is slow. Federated learning+CNN [6] is better with accuracy in the classification performance. Yet, it is not implemented in an embedded system to minimize costs. DenseNet [7] attains low complexity and secures more robust to numerical error.

On the other hand, it takes longer as it contains a more convolutional layer. "Inception V3, Inception-V2 ResNet, VGG16, VGG19, and ResNet50" [8] ensures enhanced performance regarding the accuracy, overhead, and breast cancer prediction time. Yet, it is costly to train due to complex data models. Hence, it is revealed that an efficient breast cancer diagnosis model is necessary to develop by considering these existing challenges.

TABLE I. "FEATURES AND CHALLENGES OF EXISTING BREAST CANCER DETECTION WITH DEEP LEARNING MODEL"

Author [citation]	Methodology	Features	Challenges
Houby and Yassin [1]	CNN	• It supports the radiologists and physicians for the breast cancer diagnosis and reduces the possibility of the biopsies process.	• It is necessary to investigate more validations in terms of accuracy before applying them to practical applications.
Feki *et al.* [2]	VGG16 and ResNet50	• It contains high practical values and can "detect the early stage of breast cancer" at the right time.	• It does not solve the problem of data imbalance and requires more time for the training process.
Hirra *et al.* [3]	DBN	• It effectively handles the class efficiency on the imbalanced dataset using the optimization of the networks.	• It is not suitable for various medical image analyses and standard imbalanced datasets.

Wang *et al.* [4]	US-ELM	● It simplifies the multi-optimization problem. ● It increases the performance in terms of high accuracy.	● It is complicated to interpret. ● It won't work well when the data is strongly non-linear.
Zheng *et al.* [5]	DLA-EABA	● It recognizes the ROI mapping in the embedding process without any help of additional information.	● It is challenging to train the data. ● The computation of the network is slow.
Połap and Woźniak [6]	Federated learning+CNN	● The classification efficiency is improved with accuracy. ● Higher performance in terms of feature selection is attained.	● It gets higher costs during implementation.
Priya and Peter [7]	DenseNet	● It Attains low complexity. ● It Secures more robust to numerical error.	● It takes a longer time as it contains a more convolutional layer.
Saber *et al.* [8]	"Inception V3, Inception-V2 ResNet, VGG16, VGG19, and ResNet50"	● It ensures enhanced performance regarding the accuracy, overhead, and breast cancer prediction time.	● It is enormously costly for training owing to intrinsic data models.

III. IMPLEMENTATION OF FEDERATED LEARNING CONCEPT FOR ENHANCED BREAST CANCER DIAGNOSIS USING MAMMOGRAM IMAGES

C. Importance of Federated Learning

FL is an approach utilized to train any artificial intelligence approaches with data from various sources to manage data anonymity. Thus, it helps remove multiple barriers to sharing data without the data being exposed or transported outside from their original location. FL differs from the standard collaborative learning with characteristics which are (i) data distribution is non-independent and identically distributed while considering the real-world scheme scenario, (ii) the central aggregate server does not have control over devices or individual nodes, and also there is no participation of entire nodes due to passive devices, which do not react to the server, and (iii) a massive range of client nodes participates in training and observes slower communication speed among the aggregate server and client nodes. FL has a recent surge in popularity, and it is a paradigm that holds promising results in learning with divided sensitive data. FL "enables training of a shared global model" instead of relying on the conventional "discovery then replication design or aggregating data" from various places altogether, where they originate while "keeping the data in local institutions" [30]. FL also achieves outstanding accomplishments in healthcare data analytics. During the FL process, the sensitive patient data is located either in individual consumers or with local institutions for both consumer and provider. FL is a viable approach for connecting data from electronic health records in various institutions that have permitted to offer privacy. It is also known that FL is an approach to train the predictive models without distributing patient-level data for managing data security when enabling inter-institutional collaboration. FL has attained more interest in recent years as it allows model training without distributing confidential patient data.

Moreover, FL avoids the requirement of collecting and storing private data at a centralized location through permitting institutions while utilizing the ML model. Overall, federated learning has the ability of several actors to build a robust, standard machine learning model without distributing data. So, it permits addressing complications like access to heterogeneous data, data access rights, data security, and data privacy.

D. Federated Learning Framework Preliminaries

A new FL framework is implemented by considering the client-server architecture, where the input images with normal and abnormal classes are classified. While considering the configuration, a global model is maintained by a centralized parameter sever, which focuses on sharing the details with the clients and performs coordination among their updates. A robust model is built by clients coordinates by taking their private datasets. This paper has suggested a new DenseNet for extracting the features and proposed an HDRO-based E-RNN architecture for classifying the mammogram images for breast cancer diagnosis. This model has taken the input from the mammogram images and classified the breast cancer classes regarding normal and abnormal results. The learning process of this E-RNN includes various communication rounds in which the clients focus on interacting with the central server synchronously. During the initial process, the E-RNN model initializes the number of hidden neurons and the

number of epochs ξ. Let us assume that there are C-number of clients, where each client consists of a_k private mammogram images saved locally. Here, four steps are there in every communication round r, explained here.

Step 1: Firstly, a global central model is maintained by the central server along with an initial number of parameters that is distributed with a subset of clients Z_r, which are chosen randomly and specified as a fraction Fr, where $Fr \in [0,1]$.

Step 2: Every client $c \in Z_r$ receives the initial parameters that carry out training procedures on a mini-batch of their local own private data using "mini-batch stochastic gradient descent" based on the reduction of local objective O_c along with a number of epochs Ep and local learning rate $\Re_{local}$.The classification is carried out by reducing the "categorical cross-entropy loss", where the clients focus on optimizing the model.

Step 3: After finishing the local training, they run stochastic gradient descent for epochs on local data points. Then, the model updates are forwarded from the user Z_r to the server.

Step 4: At last, the updates are received to the server from all the clients, and then, an average model is estimated for updating the constraints of the global model using Eq. (1).

$$\xi^r \leftarrow \sum_{c=1}^{C} \frac{a_k}{a} \xi_c^r \tag{1}$$

In Eq. (1), the total count of data points utilized in collaborated training is termed as a, the number of data points saved on client" c is mentioned as a_k, the "parameters sent by client" $c$at round r is noted as $\xi_c^r$and the parameters updated at round $r$is specified as ξ^r.

These four steps demonstrate the one round of FL-based E-RNN model for breast cancer diagnosis. Several rounds are conducted for performing this operation. Sever resends the new parameters of the global model by the server. When various clients are available, the subset of clients is modified from one round to another.

E. Proposed Federated Learning

The "breast cancer diagnosis model" is implemented with a collaborative and decentralized scheme using mammogram images. The primary goal of FL-based E-RNN focuses on sharing rich private data when conserving privacy with the help of maximizing the accuracy rate of a breast cancer diagnosis. To achieve this, the E-RNN is selected in this paper for maximizing the overall performance. For performing the optimization for E-RNN, the HRDO algorithm is used, which maximizes the diagnosis performance. Fig. 1 illustrates the FL-based breast cancer diagnosis model, which includes two parities like a central server and federated clients to exchange information.

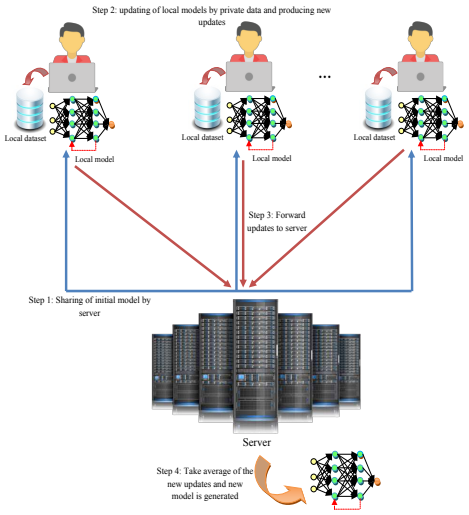


Fig. 1. A framework of federated learning for the breast cancer diagnosis model

The server owns the global model used to manage the complete progression of the "model training" and distributes the original model to all the client participants. The server gets synchronized updates from each client at every federated round and then aggregates them for building a novel model by considering the updated constraints. Further, the training is carried out on the client-side, where every federated client includes a specific dataset with computational capabilities for running a mini-batch stochastic gradient descent algorithm. Every local model is initialized at every round using a global model coming from the server. The client generates the new updated model through computing a gradient update that is further shared with the aggregation server. The local data remains private to every client by following this training protocol and is never shared.

IV. Dataset Preparation, Architecture Description, and Deep Learning-Based Feature Extraction

F. Data Preparation for Federated Learning

The proposed model training and testing images are from "http://www.eng.usf.edu/cvprg/Mammography/Database.html: Access Date: 2022-04-01". This dataset is named "Curated Breast Imaging Subset of DDSM (CBIS-DDSM)," which is a standardized and updated version of "Digital Database for Screening Mammography (DDSM)". This dataset consists of 2,620 mammogram images, including malignant, benign, and normal cases with verified pathology information. The dataset is split into training and validation sets, where 30% of the data is used for testing and 70% is used for training. The data distribution is carried out where each client has an individual set of images, including three cases. The entire client has a similar amount of data based on distribution. A skewed class distribution is used for simulating the proposed model. They have further divided the learning data, where every client attains a various count of images from every class. Lastly, the training dataset is generated for testing the unbalanced data distribution over clients. Here, the overlap among the client sub-sets is avoided for making a federated setup a realistic one.

As the gathered dataset is small, the data augmentation is carried out for artificially expanding the "size of the training and testing" the sub-sets through building the versions of the images. Moreover, the zoom or rotation operations are carried out for modifying the images. Here, zoom augmentation is carried out by zooming in or zooming out the image based on the small range. Then, rotation augmentation focuses on rotating the image left or right on an axis through a small or random degree. Finally, the FL-based data preparation is carried out for classifying the breast cancer diagnosis. The prepared data for breast cancer classification is termed as D_n, $n = 1,2,3,\cdots, N$ and the "total number of gathered images" is specified as N. The sample images gathered from the dataset are given in Fig. 2.

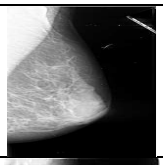
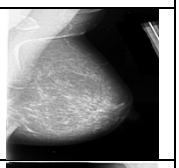
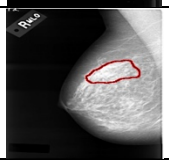
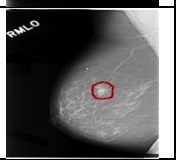
Image description	Image 1	Image 2	Image 3
Normal			
Cancer			

Fig. 2. Gathered sample images from the dataset

G. Overall Architecture

A growing and urgent universal complication is cancer, leading to mortality and morbidity that is more dependable for one in six global deaths. There is a need to analyze the consequences of cancer by taking the noteworthy inequities noticed in medical care worldwide. Undoubtedly, early diagnosis is the most significant prognostic constraint for breast cancer. Moreover, timely detection helps to make appropriate decisions to save lives. In terms of breast cancer diagnosis, the prevention measures are reducing obesity, minimization of alcohol consumption, and encouraging breastfeeding. Breast cancer detection at an early stage must consider two clear medical cases. The diagnosis is carried out by either physical exam finding and the other is the earlier examination of a patient's complaint. Various screening approaches are presented in recent years: Magnetic Resonance Imaging (MRI), Mammography, Ultra-Sound, etc. Among all these approaches, mammography screening minimizes the mortality that is restricted to resource-limited medical care systems.

Along with that, Artificial Intelligent (AI) developments are adopted for the early detection of breast cancer that helps in accurate and necessary treatment. In early detection, the ability to detect a tumor is a painful process, relatively long waiting time, and is expensive for a patient. Although various studies focus on detecting a new early detection of breast cancer, it suffers from complications like lower accuracy, high error rate, etc. Existing studies have tried to train the neural networks through entire mammograms without depending on any annotations. Though, it is complicated to know these networks that have the ability to locate the medically significant lesions and base predictions on the equivalent portions of the mammograms. Thus, it is eminent to use deep learning approaches for breast cancer diagnosis, but it also needs a vast range of training datasets that must be effective. Hence, it is required to leverage both larger and fully annotated datasets to understand the cancer status of every image for "enhancing the accuracy of breast cancer classification" approaches. Along with the traditional breast cancer diagnosis models, this scheme adopts a federated learning scheme to give superior efficiency. A new FL-assisted breast cancer diagnosis model is shown in Fig. 3.

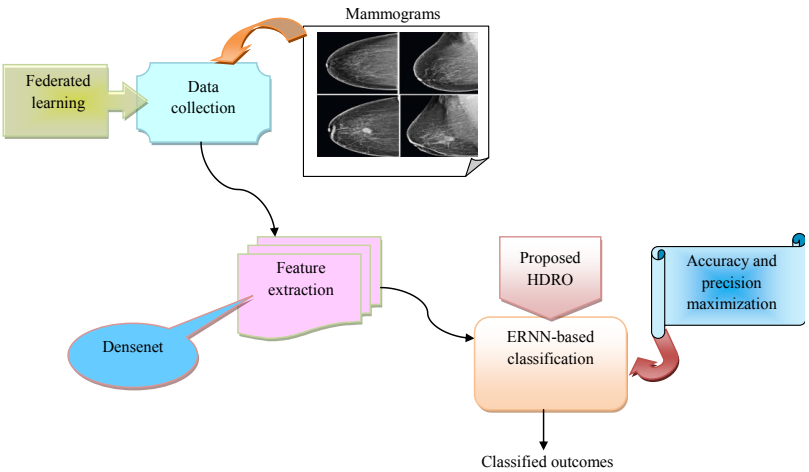


Fig. 3. An overall framework for breast cancer diagnosis using federated learning with an enhanced deep learning strategy

This paper has implemented a new FL-assisted breast cancer diagnosis model with standard procedures for helping medical professionals. This model encloses various steps that include "(a) image collection, (b) feature extraction, and (c) classification phase". The image collection is carried out at the initial stage, where the FL is used to prepare the data to make the breast cancer diagnosis easy. FL is a "distributed machine learning" scheme that guarantees "multi-institutional collaboration" on various deep learning approaches without utilizing client data. FL enables us to use the training data on user nodes or local devices while comparing with logging it to a data center. These devices carry out computations from their data for updating a global model. The prepared data is further fed to the feature extraction phase, where the deep features are extracted using Densenet architecture. The extracted features are then forwarded to the classification stage to get the normal and abnormal classes regarding breast cancer. It is carried out by E-RNN, where the heuristic improvement regarding the optimization of hidden neuron counts and the number of epochs of RNN are optimized using the HDRO algorithm. Moreover, the primary aim of this HDRO-based E-RNN is to

maximization of accuracy and precision among classified outcomes. Finally, superior performance is observed by deep feature extraction with HDRO-based E-RNN-assisted breast cancer classification.

H. Densenet-based Feature Extraction

DenseNet is used in this proposed breast cancer classification model for extracting the features from the gathered images D_n.

"Dense Convolutional Network (DenseNet)" [29] connects every layer layer in a feed-forward manner. DenseNet layers are narrower that adds only a "small set of feature maps" for making the final decisions. DenseNets have various advantages in terms of offering superior parameter efficiency, gradients throughout the network, and enhanced flow of information, which makes the DenseNet simpler for training. DenseNet always focuses on reducing the "overfitting on tasks" by considering the "smaller training set sizes" owing to the regularizing effect. The gathered images D_n are passed through the convolutional network of DenseNet, in which the network consists of $\ell-$ layers, where every layer designs a "non-linear transformation" $\Im_l(\cdot)$. Here, the indexes of the layer are specified as l. This $\Im_l(\cdot)$ function is taken as a "composite function of operations like convolution, pooling, Rectified Linear Units (ReLU), and Batch Normalization (BN)". The result of the lth layer is denoted as Y_l. In ResNets, the output of the lth layer is taken as the input to the layer connected with existing convolutional feed-forward networks that help formulate the following derivation.

$$Y_l = \Im_l(D_{l-1}) \tag{2}$$

ResNets include a skip-connection with an identity function for bypassing the non-linear transformations as equated in Eq. (3).

$$D_l = \Im_l(D_{l-1}) + D_{l-1} \tag{3}$$

ResNet has the benefit of forwarding the gradient flow directly via the identify function from the previous layers to the former layers. Though, the output of $\Im_l$ and the identify function are integrated with a summation that can obstruct the information flow in the network.

Secondly, a dense connectivity pattern is introduced in DenseNet to further enhance the information flow among layers. It has adopted "direct connections from any layer to entire consequent layers". Traditionally, the feature maps of the entire preceding layers are received by l^{th} layer that is taken as input $d_{0,} \cdots, d_{l-1}$ and derived in Eq. (4).

$$D_l = \Im_l(\lfloor d_{0,} \cdots, d_{l-1} \rfloor) \tag{4}$$

The integration of the feature maps generated in layers $0, \cdots, l-1$ is denoted as $\lfloor d_{0,} \cdots, d_{l-1} \rfloor$. Thus, with the utilization of dense connectivity, this network is named as DenseNet. To simplify the process, the various inputs of $\Im_l(\cdot)$ concatenated into a single tensor.

Further, a composite function is described as $\Im_l(\cdot)$ that consists of three equivalent operations like a 3×3 convolution, ReLU, and BN. Eq. (4) is derived based on the concatenation operation; however, that is not feasible while changing the size of feature maps. However, an essential section of convolutional networks is down-sampling layers, which modify the size of feature maps. This network is divided into several densely linked dense blocks to facilitate downsampling in DenseNets. The transition layers are considered as the layers among blocks that carry out pooling and convolution operations. Further, the transition layers consist of 1×1 convolution layer, BN layer, and a 2×2 average pooling layer.

Then, the growth rate is formulated, where while generating every function $\Im_l$ produces $S-$ feature maps, which pursues that the lth layer consists of $S_0 + S \times (l-1)$ feature-maps, in which the amount of channels in the input layer is noted as S_0. Here, the growth rate of the network is referred to as hyperparameter S. The feature maps are taken as the "global state of the network", where every layer includes their S own feature maps to this state.

In addition, the growth rate helps regulate "how much new information" in every layer contributes to the global state. After writing the global state once, there is no requirement of replicating it from layer to layer. Even though every layer has produced S-output feature maps, it generally includes various inputs. Thus, introducing 1×1 convolution as the bottleneck layer is needed before performing every 3×3 convolution operation considerably improves the computational efficiency and number of "input feature maps". The model compactness has to be improved, and minimized the number of feature maps at transition layers. The

transition layer produces the output feature maps φ_j when a dense block includes h feature maps, in which the compression factor is given as $0 < \varphi \leq 1$. The number of feature maps over transition layers remains unmodified while $\varphi = 1$.

Finally, the DenseNet extracts the features from the gathered images, which is denoted as $fs_o^{DenseNet}$, where $o = 1,2,3,\cdots O$, and the "total number of gathered features" is expressed as O. These gathered features are further forwarded to the E-RNN phase, where the breast cancer classification is performed to get the classes regarding normal and abnormal categories.

V. ENHANCED RNN WITH HYBRID DRAGON-RIDER OPTIMIZATION FOR BREAST CANCER DIAGNOSIS

I. Basic RNN Model

RNN [28] is one category of neural network framework that processes both parallel and sequential information. NN has similar functionalities to the human brain that processes memory cells. The primary feature of RNN is its capacity to combine dynamic and static information from patients. This benefit is more necessary for clinical applications as most medical datasets exist in some background information about patients like main disease, blood type, gender, and so on that is integrated with dynamic information by taking the patient information taken from the hospitals at multiple visits. The standard RNN frames take the input as sequential information while the static information of the patients is not combined into the neural network. RNN is also useful in characterizing the relationship between a sequence's previous output with the current output to model the sequential progression. While building an RNN, there is a need to distinguish the asymmetry of the bilateral hippocampus and high-level correction for disease detection. RNN is chosen here due to its generalization ability and superior prediction accuracy. The hidden state of a one-time step in RNNs is determined by integrating the recent inputs with the hidden state of the earlier time step. RNNs have the efficiency of learning and remembering the events from the past that are suitable for predicting future results.

RNN also has the constraint of unlimited memory, and thus, the suitable time window is utilized in feed-forward neural networks for prediction tasks that must be specified. RNNs have the ability to learn to remember any combination of events or any particular process for predicting future events. It is more helpful for learning the long-term dependencies by taking the most valuable clinical dataset feature. The resultant of the RNN is determined in Eq. (5).

$$op_t = \eta\left(hs_t Ws_p\right) \tag{5}$$

In Eq. (5), a vector including the predicted result is specified as op_t, the hyperbolic function or logistic sigmoid function is known as η, a vector containing the hidden state of the Neural Network is denoted as hs_t, and the matrix containing the parameters of the model is indicated as Ws_p.

The sequence of input is given as $fs_o^{DenseNet} = \left(fs_1, fs_2, \cdots, fs_T\right)$, and then the hidden state of RNN is determined in Eq. (6).

$$hs_t = \Re\left(hs_{t-1}, fs_o^{DenseNet}\right) \tag{6}$$

Here, the function is specified as $\Re$. Moreover, RNN focuses on capturing long-term dependencies. RNN is more useful as it specifies the ability to remember events that have occurred a long time ago. It is more helpful for some clinical applications, where the events are noticed at a similar point in the patient's life.

J. Enhanced RNN for Breast cancer Diagnosis

Although RNN poses various benefits in the medical field for recognizing medical events, it suffers from vanishing gradient issues. Also, it propagates the error via the network at earlier time steps. The error also affects the performance of the designed model. Moreover, RNN requires a vast range of data for classification. While handling complex data structures, RNN suffers from complications. Thus, this paper adopts a new E-RNN for suggesting a new breast cancer diagnosis model. The framework of E-RNN for breast cancer diagnosis using FL approach is given in Fig. 4.

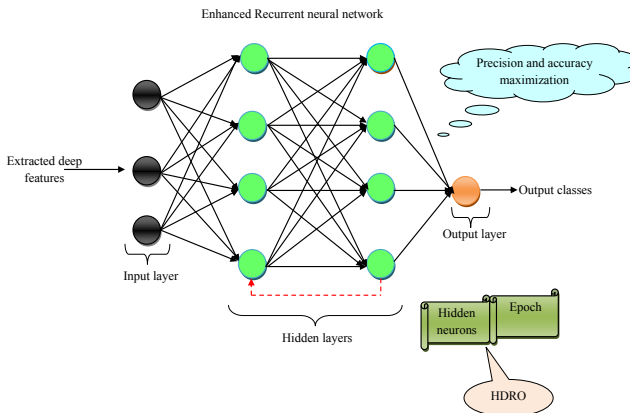


Fig. 4. The architecture of Enhanced RNN for Breast cancer Diagnosis

The E-RNN is implemented by "optimizing the number of hidden neurons and number of the epoch" of RNN architecture using HDRO technique, which focuses on adding advantages to the breast cancer detection model and improves the convergence behavior. This new FL-assisted E-RNN for breast cancer diagnosis using the HDRO algorithm helps maximize "accuracy and precision". It is derived in Eq. (7).

$$Obf = \underset{\{EH, HE\}}{\operatorname{argmin}}\left(\frac{1}{Acs + Pn}\right) \tag{7}$$

In Eq. (7), the number of hidden neurons is noted as/that "lies in the range" of [5, 255] and the number of epoch of RNN is indicated as EH and assigned in the "bounding limit" of [0.01, 0.99]. Precision is denoted as Pn and described as "the ratio of positive observations that are predicted exactly to the total number of observations that are positively predicted".

$$Pn = \frac{po^{tru}}{po^{tru} + po^{fal}} \tag{8}$$

Accuracy is specified as Acs and defined as a "ratio of the observation of exactly predicted to the whole observations".

$$Acs = \frac{\left(po^{tru} + ng^{tru}\right)}{\left(po^{tru} + ng^{tru} + po^{fal} + ng^{fal}\right)} \tag{9}$$

Here, terms po^{tru}, ng^{tru}, po^{fal}, ng^{fal} refer to the "true positives, true negatives, false positives, and false negatives," respectively. Finally, the designed E-RNN model using the HDRO algorithm improves accuracy and precision and ensures their detection rate.

K. Developed HDRO for Improving Breast cancer Diagnosis

A hybrid meta-heuristic technique termed HDRO is proposed with the help of both RDA and DA techniques. This algorithm HDRO implements an E-RNN model by optimizing the number of epochs and "number of hidden neurons" of RNN based on FL. It assists in maximization of performance due to accuracy and precision and ensures their detection rate for breast cancer diagnosis.

A new HDRO algorithm is proposed by utilizing the features of RDA and DA approaches. RDA is chosen in this paper because it has superior balancing among the exploration and exploitation stages, gets enhanced convergence rate, and easily reaches the global optima solutions. On the other hand, RDA also has some complexities in tuning diverse "controlling parameters, fewer execution speeds" and suffers from finding global solutions. Thus, a new HDRO algorithm is proposed by adopting the features of the DA technique to provide a higher balance among various phases, improve the diversity of the solutions, and enhance performance. It is simpler owing to less control constraints. DA with RDA tries to offer optimal solutions at a reasonable convergence rate.

HDRO algorithm is suggested based on the sensing area, where the random vector uniformly distributed in the range [0, 1] used in DA is used for updating the solutions; where this $rnd > 0.5$ is satisfied, then the solutions are updated using DA technique or else the solutions are updated using RDA algorithm. Here, the random number rnd is formulated by determining the ratio among the best fitness and the mean of best fitness solutions formulated in Eq. (10).

$$rnd = \frac{M^{best}(j)}{mean\left(M^{best}(j)\right)} \tag{10}$$

Here, the mean of best fitness solutions is known as $M^{best}(j)$, and the best fitness solutions are given as $M^{best}(j)$. If $rnd > 0.5$ is satisfied, then the solutions are updated using DA technique.

DA [27] is motivated by the food searching behavior of dragonflies. It also considers both "dynamic and static behaviors of dragonflies". There are five constraints such as "control cohesion, alignment, separation, attraction (towards food sources), and

distraction (towards outward enemies) of individuals in the swarm" are taken for updating the position of individuals in swarms. Eq. (11) computes the population of dragonflies.

$$M_i^{j+1} = M_i^j + \Delta M_i^{j+1} \tag{11}$$

$$\Delta M_i^{j+1} = \left(soSo_i + avAV_i + csCS_i + hsHs_i + enEn_i\right) + \zeta \cdot \Delta M_i^j \tag{12}$$

The aforementioned equations, the cohesion weight is derived cs, iteration counter is referred to as j, the position of dragonflies is given as M_i^j and inertia weight is known as ζ, i^{th} dragonfly food source is indicated as Hs_i, velocity or step vector is termed as ΔM_i^{j+1}, the cohesion and alignment of the ith dragonfly is represented as CS_i and AV_i, respectively, and the food factor is suggested as hs and i^{th} dragonfly enemy position is specified as En_i, the alignment weight is noted as av and the ith dragonfly's separation and the separation weight are specified as So_i and so, respectively, and the enemy factor is given as en. These constraints are formulated as follows.

$$So_i = -\sum_{t=1}^{Nj}\left(M - M_t\right) \tag{13}$$

$$AV_i = \frac{\sum_{t=1}^{Nj} Vo_t}{Nj} \tag{14}$$

$$CS_i = \frac{\sum_{t=1}^{Nj} M_j}{Nj} - M \tag{15}$$

$$Hs_i = Fo_m - M \tag{16}$$

$$En_i = Ey_m + M \tag{17}$$

Here, the position of the food source is noted as Fo_m, Ey_t denotes the enemy's position, t^{th} neighboring individual's position is given as M_j, the current individual's position is indicated as M, the velocity is denoted as Vo_t at t^{th} neighboring individual, and the count of neighboring individuals is considered as Nj.

If $rnd \leq 0.5$ is satisfied, then the solutions are updated using the RDA technique. RDA is implemented owing to the abnormal mating behavior of Scottish red deer in a breeding season. The population of Red Deers (RDs) is initialized first, and then it is categorized into two RDs like "hinds and male RDs". As well, the harem is known as a set of female RDs. The fighting and roaring behaviors of RDs are taken for establishing the hunting behavior of RDs.

"Roar male RDs": The superior solutions are attained by performing roaring by male RDs. The position of every male RD is changed, and then it is equated in Eq. (18).

$$M_{new} = \begin{cases} M_{old} + bs_1 \times \left(\left(\left(ub - lw\right) * bs_2\right) + lw\right), & if\ bs_3 \geq 0.5 \\ M_{old} - bs_1 \times \left(\left(\left(ub - lw\right) * bs_2\right) + lw\right), & if\ bs_3 < 0.5 \end{cases} \tag{18}$$

Here, the current position and updated position of male RD are correspondingly specified as M_{old} and M_{new}, and constraints like bs_1, bs_2, and bs_3 are generated among 0 and 1 through a uniform distribution in a random way.

"Choose η percent of the best male RD as male commanders": RD is divided into two cases such as "stags and commanders". The number of stags Nj_{stags} is calculated in Eq. (19).

$$Nj_{stags} = Nj_{male} - Nj_{commander} \tag{19}$$

$$Nj_{commander} = round\{\eta, Nj_{male}\} \tag{20}$$

In Eq. (12), the commanders of harems $Nj_{commande}$ are computed from the number of males Nj_{male}, and the initial value is noted as η with the bounding range of 0 and 1.

"Fight among male commanders and stags": Let us assume that every commander will fight with stages randomly in the solution space. Here, after obtaining the better solutions, it is replaced with commander by performing the below-mentioned two new solutions. The fighting process is given here.

$$X_1 = \frac{(commander + stags)}{2} + bs_1 \times (((ub - lw) * bs_2) + lw) \quad (21)$$

$$X_2 = \frac{(commander + stags)}{2} - bs_1 \times (((ub - lw) * bs_2) + lw) \quad (22)$$

Here, there are four solutions such as X_1 and X_2, $stag$ and $commande$, and the best solutions are chosen via objective function. Two solutions produced in the fighting process are given as X_1 and X_2, the symbol of commander and stages are specified as $stag$ and $commande$.

"Form harems": It is carried out where the power of male commanders determines the Nj_{Hind} in harem through their objective function, and the hinds among commanders are proportionally separated in Eq. (23).

$$Y_v = y_v - \max_i \{y_i\} \quad (23)$$

In Eq. (23), the normalized value of power is noted as Y_v, the power of the vth commander and their objective function is expressed as y_v, and in Eq. (24), the normalized power of commanders is formulated.

$$Po_v = \left| \frac{Y_v}{\sum_{j=1}^{Nj_{commander}} Nj_i} \right| \quad (24)$$

The number of hinds of a harem is given in Eq. (25).

$$No.harem_v = round\{Po_v \cdot Nj_{Hind}\} \quad (25)$$

In Eq. (25), the number of hinds in v^{th} harem is referred to as $harem_v$ and the number of all hinds is known as Nj_{Hind}.

"Mate commander of a harem with γ percent of hinds in his harem": Deer mate with each other mate through taking the γ percent of hinds in a commander and his harem as derived in Eq. (26).

$$Nj.harem_v^{mate} = round\{\gamma \cdot Nj.harem_v\} \quad (26)$$

In Eq. (26), the number of hinds of the vth harem mate with its commander is formulated as $Nj.harem_v^{mate}$, and an initial parameter value is considered as γ in the range of 0 to 1. Eq. (27) formulates the mating process.

$$X_3 = \frac{(commander + Hind)}{2} + (ub - lw) \times ad \quad (27)$$

Here, a new solution is represented as X_3 and the randomly generated using a uniform distribution function among 0 and 1 is termed as ad.

"Mate commander of a harem with α percent of hinds in another harem": The Nj_{Hind} in the harem that "mate with the commander" by initializing the parameter α is formulated in Eq. (28)

$$Nj.harem_k^{mate} = round\{\alpha \cdot Nj.harem_k\} \quad (28)$$

"Mate stag with the nearest hind": Every stag "mates with its nearest hind in the breeding season". There is a need to determine the distance among an entire hind and stag are equated in Eq. (29) for finding the closest hind.

$$di_i = \left(\sum_{ks \in KS} \left(stags_{ks} - Hind_{ks}^{i} \right)^2 \right)^{\frac{1}{2}} \tag{29}$$

Here, the minimum value in this matrix is specified as the chosen hind and the distance amid the ith hind and a stag is depicted as di_i.

"Chose the next generation": The next generation is selected by taking two schemes. Roulette wheel mechanism is used for determining the new solutions with the mating procedure.

Finally, the solutions are updated until reaching the final solutions. The "pseudo-code" of the developed HDRO is given in Algorithm 1.

Algorithm 1: Proposed HDRO algorithm

```
Initialization of the population
Computation of the fitness among individuals
While j < j_max
    if rnd > 0.5
    Solutions are updated based on DA
        Update the solution vector by Eq. (11).
        Update the velocity vector by Eq. (12).
    Else
    Solutions are updated based on RDA
        for every male RD
            Utilize Eq. (18) for roaring
            Update the better positions
        end for
        Update stage and commanders using Eq. (20) and Eq. (19)
        for every male commander
            Use Eq. (21) and Eq. (22) for fighting among stages and
            commanders
            Update position of commanders and stages
        end for
        Carry out Eq. (23), Eq. (24), and Eq. (25) to form harems
        for every male commander
            Apply Eq. (27) for mating with randomly chosen hinds
            Use Eq. (28) to select the harem
            Update Eq. (27) for mating with selected hinds of the harem
        end for
        for every stage
            Determine the distance through Eq. (29)
            Update Eq. (27) for the mating stage with selected hind
        end for
            Use a roulette wheel strategy for choosing the next generation
    end if
Update the best solutions
end while
Return optimal solutions
```

Thus, the hybrid algorithms are focused on getting superior solutions. The flowchart of the suggested HDRO algorithm is specified in Fig. 5.

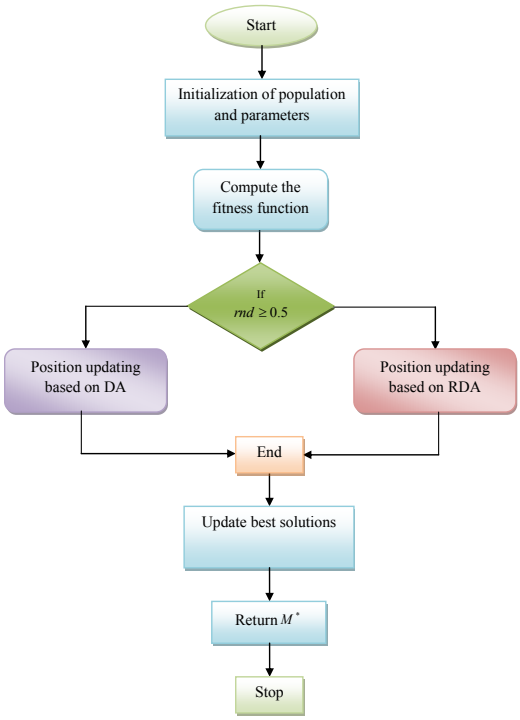


Fig. 5. The flowchart of the designed HDRO algorithm for FL-assisted breast cancer diagnosis

VI. RESULTS AND DISCUSSIONS

L. Experimental setup

The evaluation of the designed FL-assisted breast cancer detection model was implemented in python. Some standard performance measures were used to execute the suggested model's efficiency. The performance of the proposed model was estimated over state-of-the-art schemes like FL+CNN [6], DenseNet [7], Ensemble [8], and RNN [28]. Other heuristic approaches like Particle Swarm Optimization (PSO) [31], Deer Hunting Optimization Algorithm (DHOA) [32], DA [26] and RDA [27]. The execution was conducted by taking the "number of population and the maximum number of iterations" as 10 and 25, respectively.

M. Performance measures

The execution was done by using various standard measures as given here.

(a) F1 score: "harmonic mean between precision and recall. It is used as a statistical measure to rate performance".

$$F1score = \frac{2po^{tru}}{2po^{tru} + po^{fal} + ng^{fal}} \tag{30}$$

(b) MCC: "correlation coefficient computed by four values".

$$MCC = \frac{po^{tru} \times ng^{tru} - po^{fal} \times ng^{fal}}{\sqrt{\left(po^{tru} + po^{fal}\right)\left(po^{tru} + ng^{fal}\right)\left(ng^{tru} + po^{fal}\right)\left(ng^{tru} + ng^{fal}\right)}} \tag{31}$$

(c) NPV: "probability that subjects with a negative screening test truly don't have the disease".

$$NPV = \frac{ng^{tru}}{ng^{fal} + ng^{tru}} \tag{32}$$

(d) FDR: "the number of false positives in all rejected hypotheses".

$$FDR = \frac{po^{fal}}{po^{fal} + po^{tru}} \tag{33}$$

(e) FPR: "the ratio of false positive predictions to the entire count of negative predictions".

$$FPR = \frac{po^{fal}}{po^{fal} + ng^{tru}} \tag{34}$$

(f) FNR: "the proportion of positives which yield negative test outcomes with the test".

$$FNR = \frac{ng^{fal}}{ng^{tru} + po^{tru}} \tag{35}$$

(g) Sensitivity: "the number of true positives recognized exactly".

$$Se = \frac{po^{tru}}{po^{tru} + ng^{fal}} \tag{36}$$

(h) Specificity: "the number of true negatives determined precisely".

$$Sp = \frac{ng^{tru}}{ng^{tru} + po^{fal}} \tag{37}$$

N. Convergence analysis

The effectiveness of the suggested FL-assisted "breast cancer diagnosis model" is estimated in Fig. 6. This analysis has shown that the performance using HDRO-FL+E-RNN depicts the maximum performance rate in terms of higher accuracy and precision over all the iterations by estimating with other heuristic algorithms.

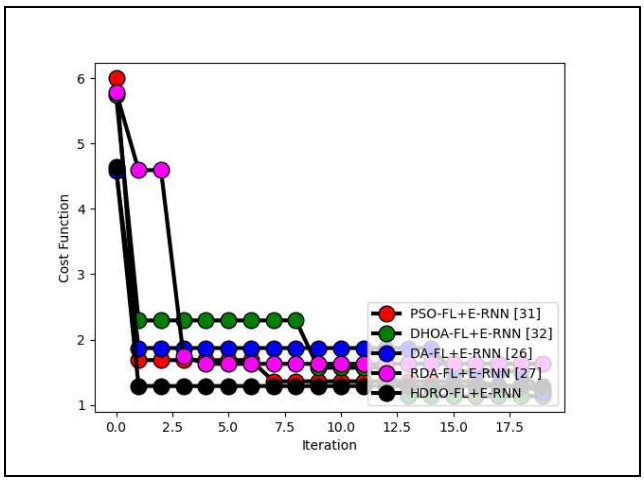
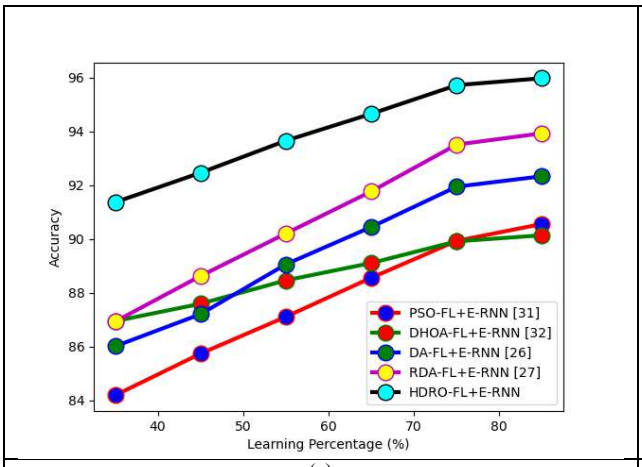


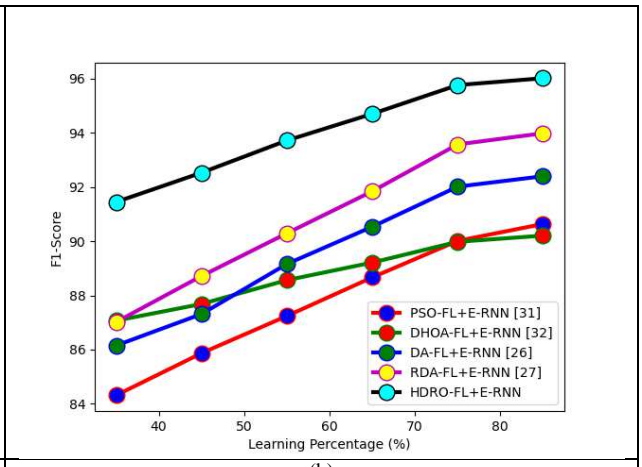
Fig. 6. Convergence analysis of the FL-assisted "breast cancer diagnosis model."

O. Investigation on algorithms

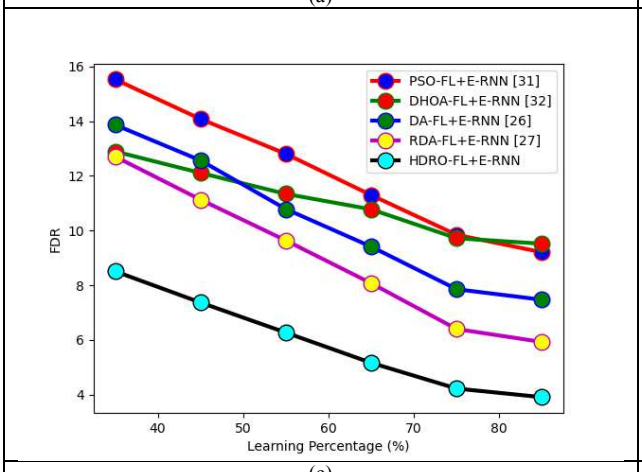
The designed HDRO-FL+E-RNN-based breast cancer diagnosis model is evaluated by varying the learning percentages as given in Fig. 7. The performance analysis has specified that superior efficiency is observed while evaluating with other heuristic algorithms. For example, the F1-score of the designed HDRO-FL+E-RNN gets 6%, 6.6%, 4.3%, and 2%, correspondingly progressed than PSO-FL+E-RNN, DHOA-FL+E-RNN, DA-FL+E-RNN and RDA-FL+E-RNN at 85%. Thus, the maximum efficiency is observed by the HDRO-FL+E-RNN-based breast cancer diagnosis model.


Accuracy
Learning Percentage (%)
PSO-FL+E-RNN [31]
DHOA-FL+E-RNN [32]
DA-FL+E-RNN [26]
RDA-FL+E-RNN [27]
HDRO-FL+E-RNN

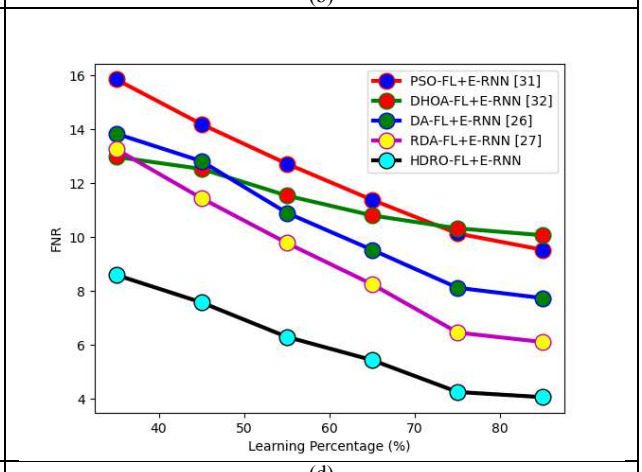
(a)


F1-Score
Learning Percentage (%)
PSO-FL+E-RNN [31]
DHOA-FL+E-RNN [32]
DA-FL+E-RNN [26]
RDA-FL+E-RNN [27]
HDRO-FL+E-RNN

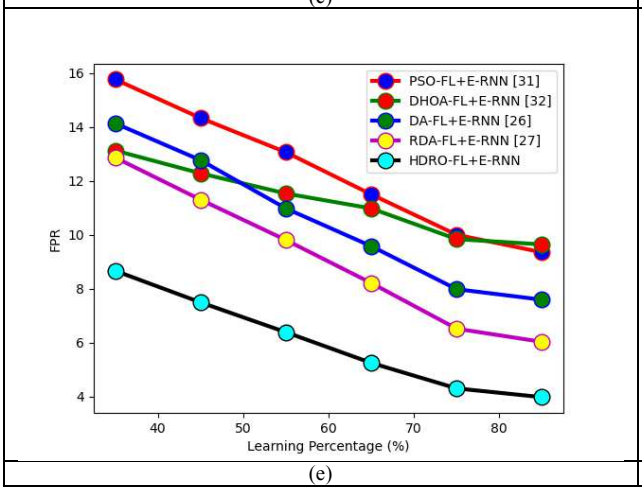
(b)


FDR
Learning Percentage (%)
PSO-FL+E-RNN [31]
DHOA-FL+E-RNN [32]
DA-FL+E-RNN [26]
RDA-FL+E-RNN [27]
HDRO-FL+E-RNN

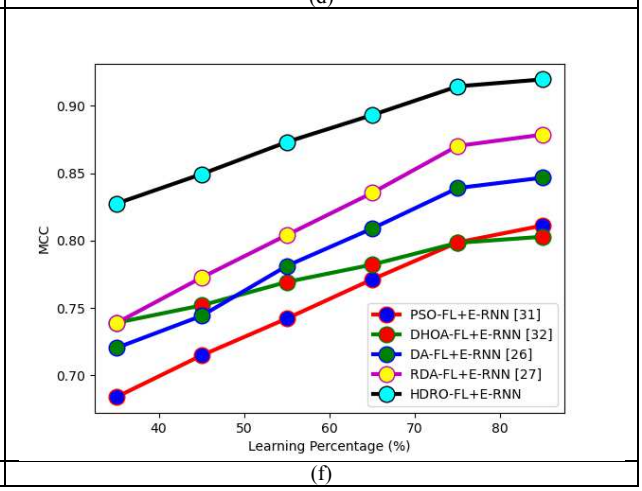
(c)


FNR
Learning Percentage (%)
PSO-FL+E-RNN [31]
DHOA-FL+E-RNN [32]
DA-FL+E-RNN [26]
RDA-FL+E-RNN [27]
HDRO-FL+E-RNN

(d)


FPR
Learning Percentage (%)
PSO-FL+E-RNN [31]
DHOA-FL+E-RNN [32]
DA-FL+E-RNN [26]
RDA-FL+E-RNN [27]
HDRO-FL+E-RNN

(e)


MCC
Learning Percentage (%)
PSO-FL+E-RNN [31]
DHOA-FL+E-RNN [32]
DA-FL+E-RNN [26]
RDA-FL+E-RNN [27]
HDRO-FL+E-RNN

(f)

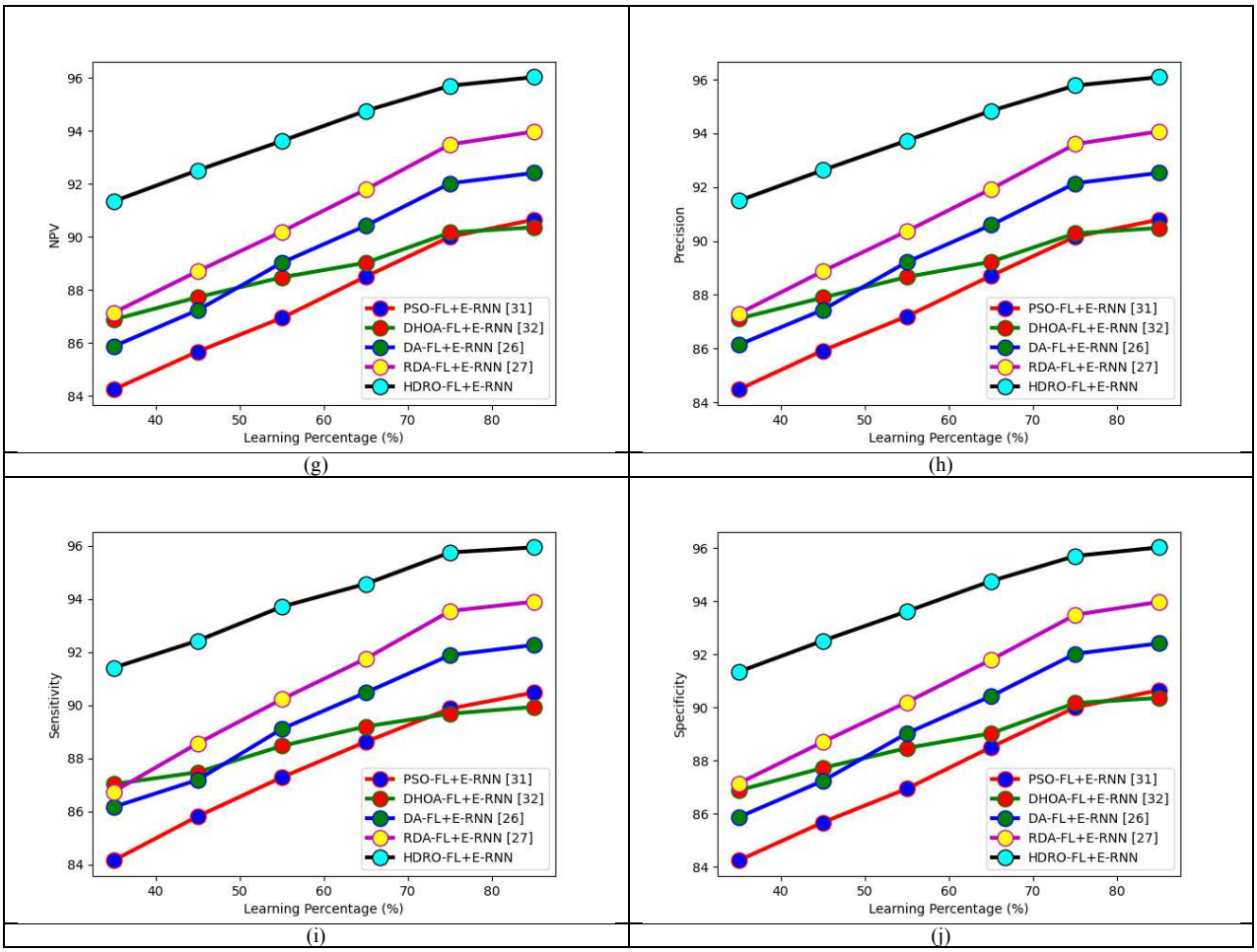
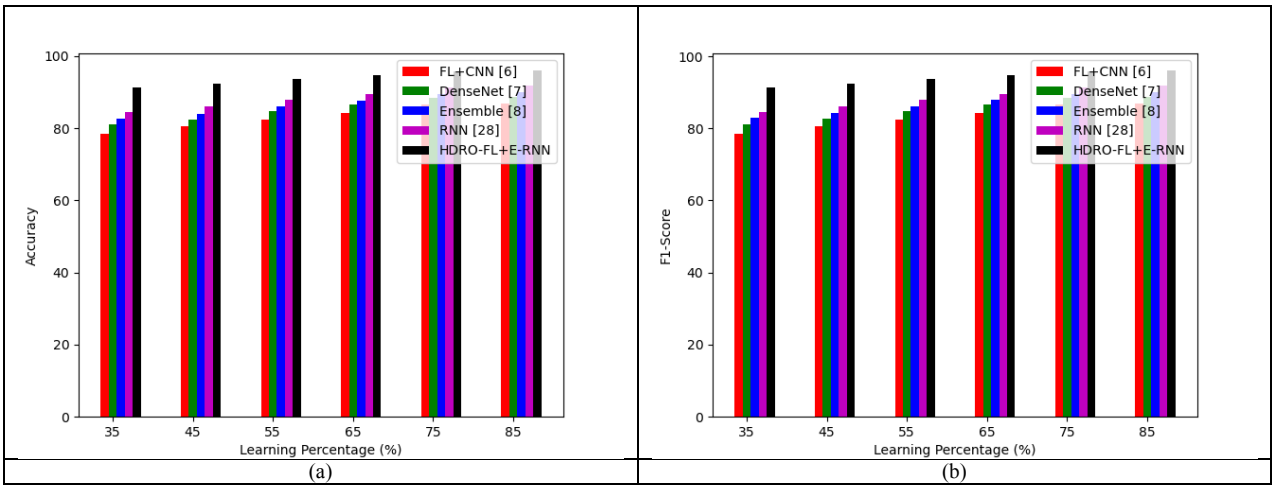
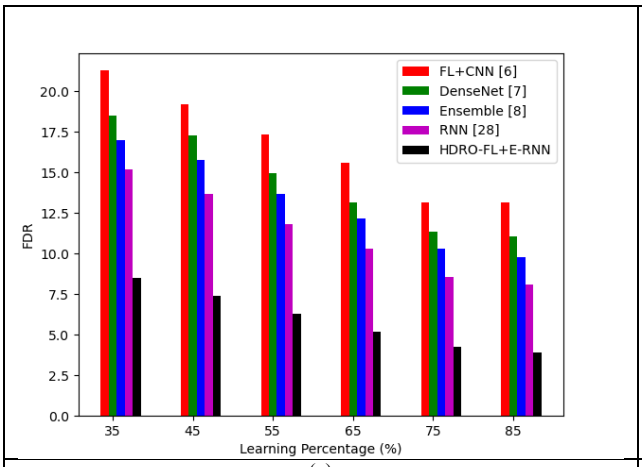


Fig. 7. Analysis of the FL-assisted "breast cancer diagnosis model" over heuristic-based algorithms regarding "(a) Accuracy, (b) F1 Score, (c) FDR, (d) FNR, (e) FPR, (f) MCC, (g) NPV, (h) Precision, (i) Sensitivity, and (j) Specificity"

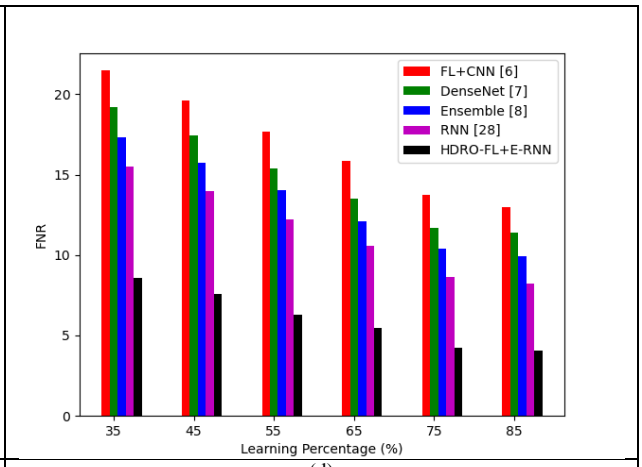
P. Investigation on existing approaches

The efficacy of the implemented HDRO-FL+E-RNN-based breast cancer diagnosis model is analyzed over other existing diagnosis models, as given in Fig. 8. The FPR of the designed HDRO-FL+E-RNN is 68.75%, 64%, 61.5%, and 58.3% more advanced than FL+CNN, DenseNet, Ensemble, and RNN, respectively at 65%. From this analysis, it is verified that the suggested breast cancer diagnosis model using FL with deep feature extraction helps in maximization of accuracy along with various performance measures.

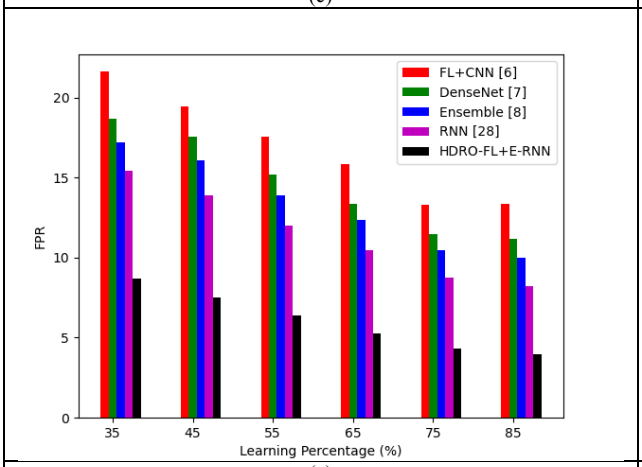



FL+CNN [6]
DenseNet [7]
Ensemble [8]
RNN [28]
HDRO-FL+E-RNN
FDR
Learning Percentage (%)

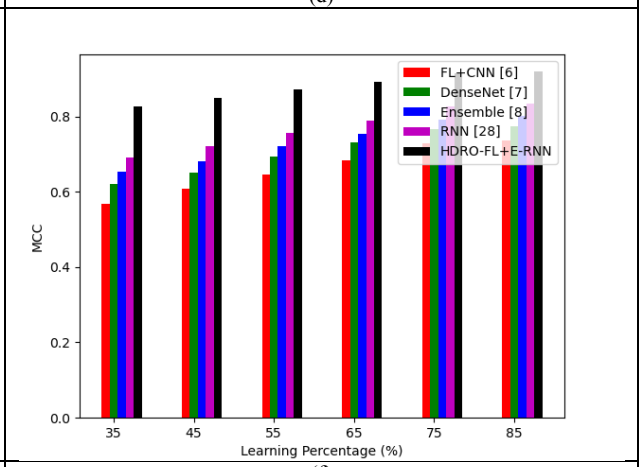
(c)


FL+CNN [6]
DenseNet [7]
Ensemble [8]
RNN [28]
HDRO-FL+E-RNN
FNR
Learning Percentage (%)

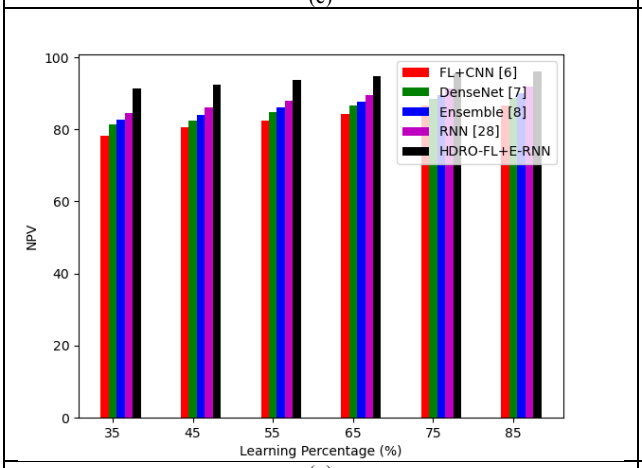
(d)


FL+CNN [6]
DenseNet [7]
Ensemble [8]
RNN [28]
HDRO-FL+E-RNN
FPR
Learning Percentage (%)

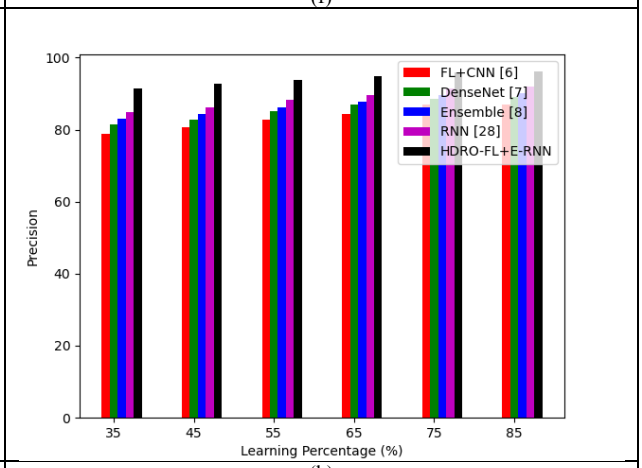
(e)


FL+CNN [6]
DenseNet [7]
Ensemble [8]
RNN [28]
HDRO-FL+E-RNN
MCC
Learning Percentage (%)

(f)


FL+CNN [6]
DenseNet [7]
Ensemble [8]
RNN [28]
HDRO-FL+E-RNN
NPV
Learning Percentage (%)

(g)


FL+CNN [6]
DenseNet [7]
Ensemble [8]
RNN [28]
HDRO-FL+E-RNN
Precision
Learning Percentage (%)

(h)

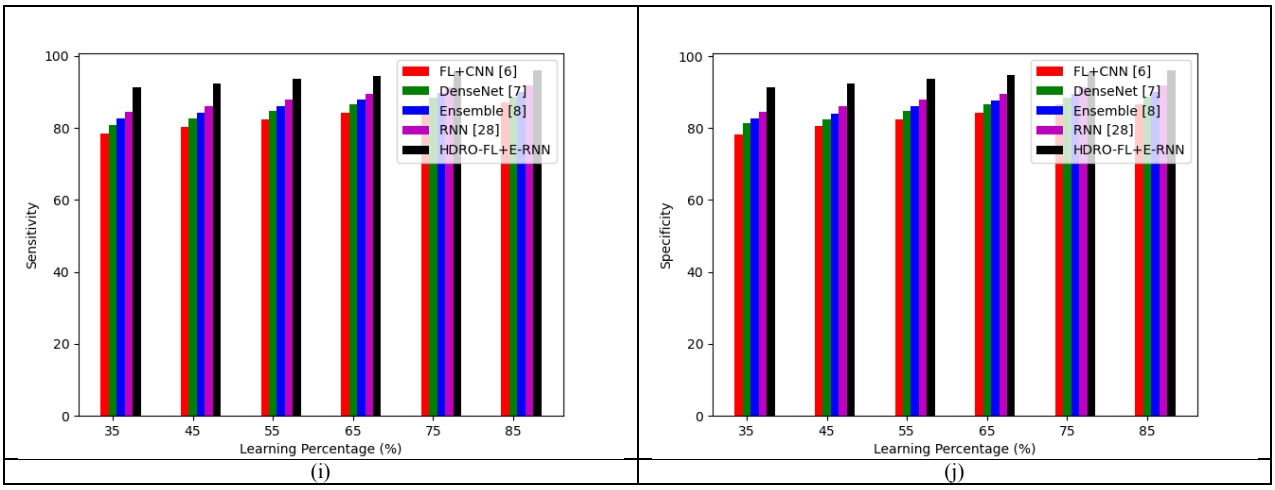
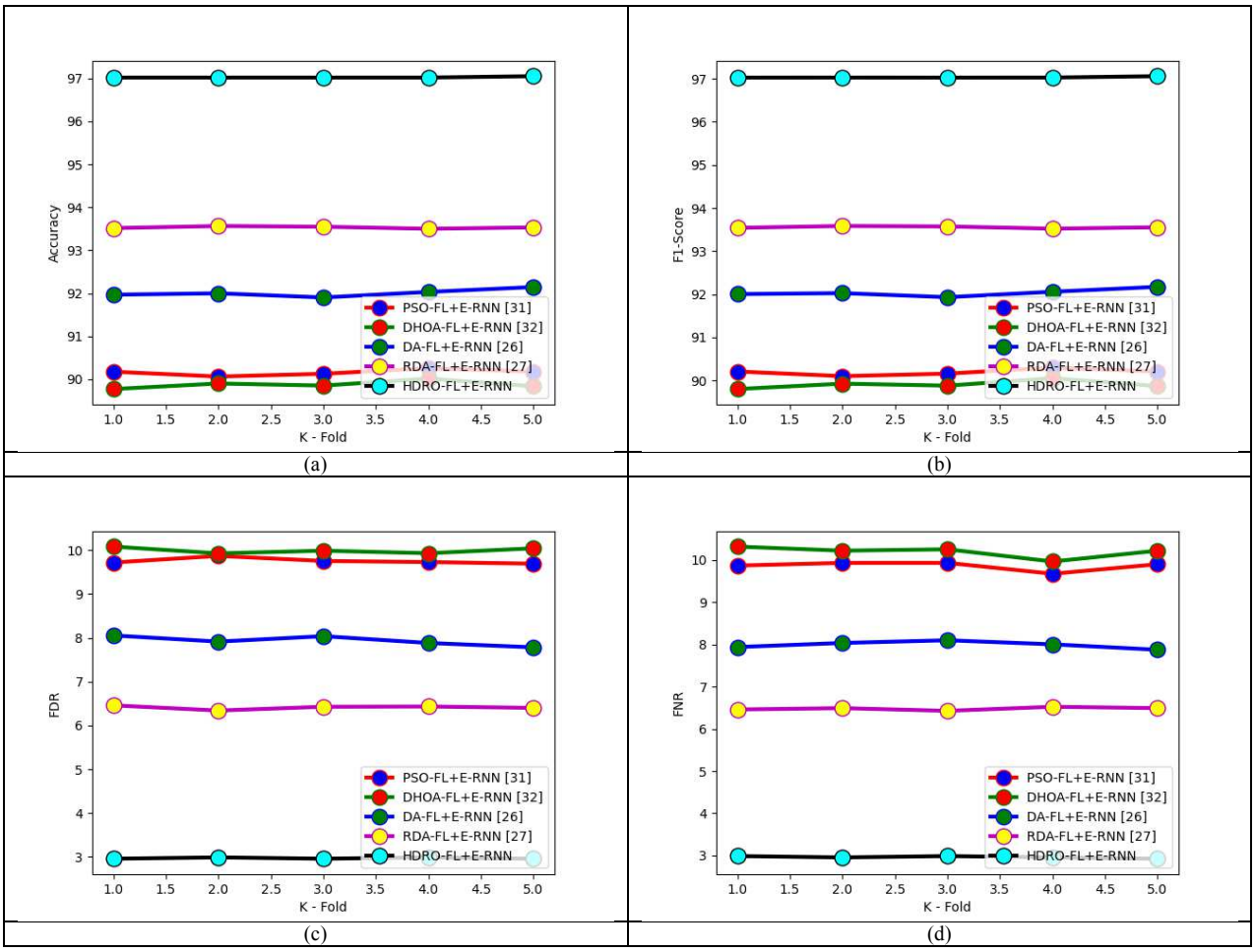


Fig. 8. Analysis of the FL-assisted "breast cancer diagnosis model" over existing models regarding "(a) Accuracy, (b) F1 Score, (c) FDR, (d) FNR, (e) FPR, (f) MCC, (g) NPV, (h) Precision, (i) Sensitivity, and (j) Specificity"

Q. K-fold analysis

The K-fold analysis estimates the efficiency of the designed FL-assisted "breast cancer diagnosis model" over other heuristic algorithms and existing diagnosis models, as given in Fig. 9 and Fig. 10. Consequently, superior efficiency is observed by the initial folds itself over various k-fold while analyzing with various approaches. Thus, it is shown that the designed model has exhibited its efficacy over other approaches.



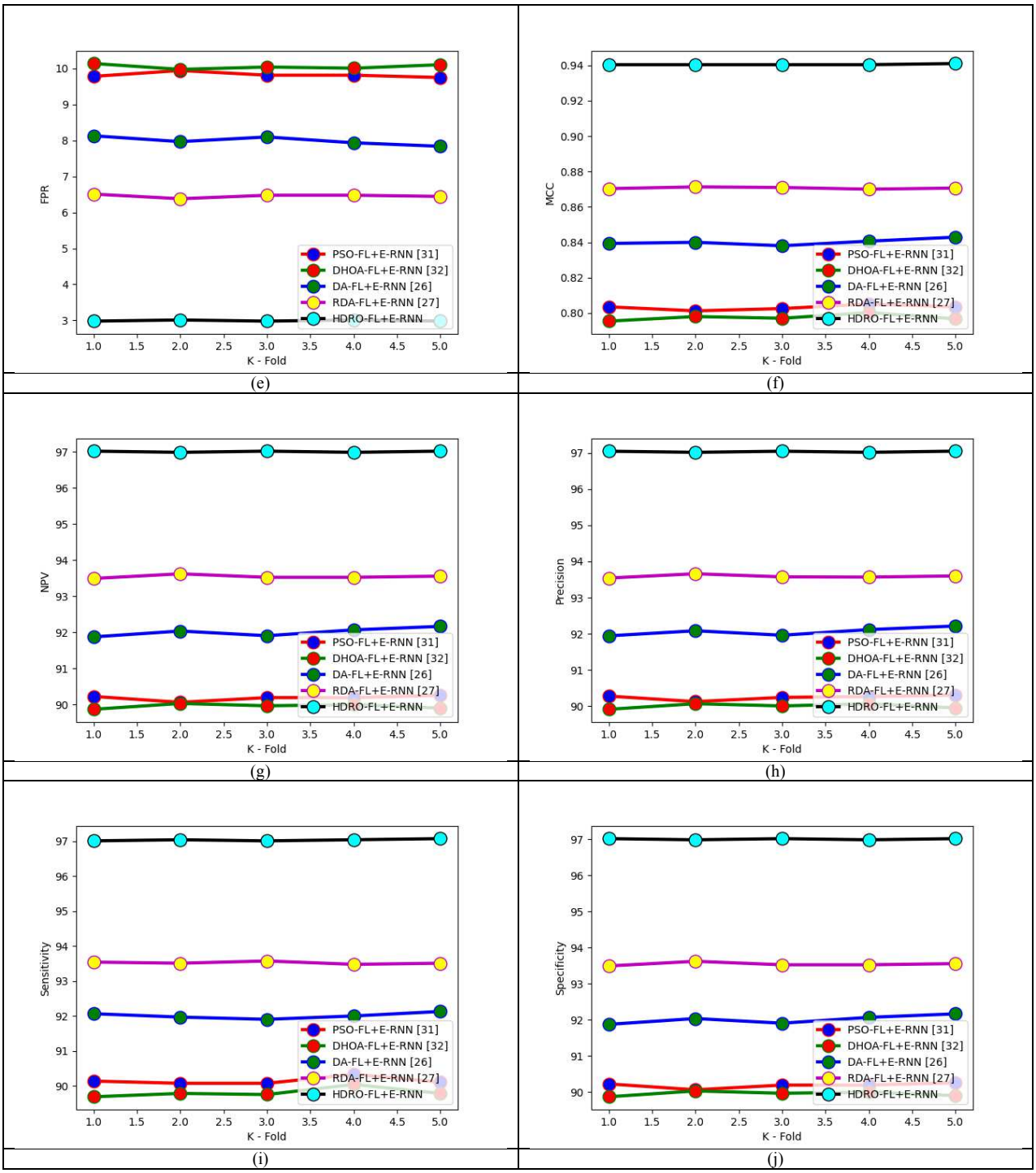
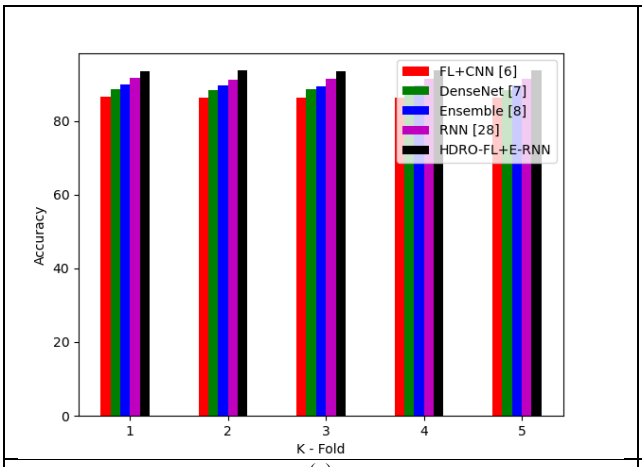
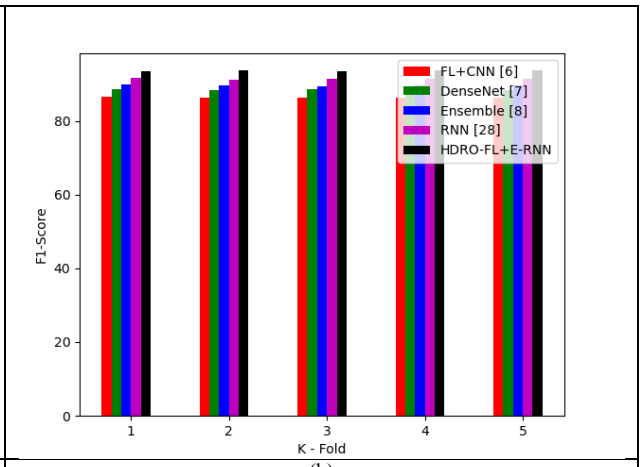


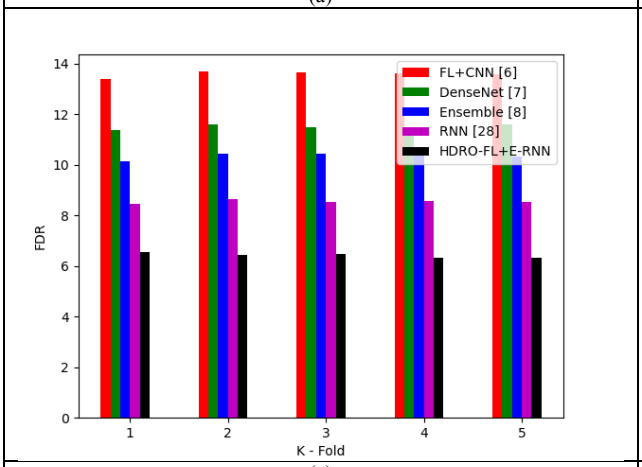
Fig. 9. K-fold validation Analysis of the FL-assisted "breast cancer diagnosis model" over heuristic-based algorithms regarding "(a) Accuracy, (b) F1 Score, (c) FDR, (d) FNR, (e) FPR, (f) MCC, (g) NPV, (h) Precision, (i) Sensitivity, and (j) Specificity"


FL+CNN [6]
DenseNet [7]
Ensemble [8]
RNN [28]
HDRO-FL+E-RNN
Accuracy
K - Fold

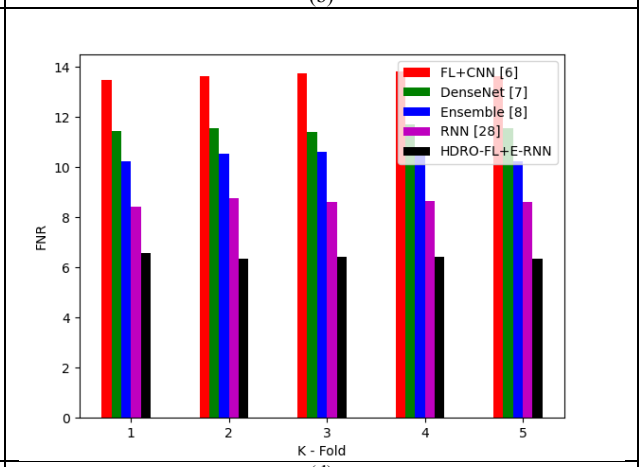
(a)


FL+CNN [6]
DenseNet [7]
Ensemble [8]
RNN [28]
HDRO-FL+E-RNN
F1-Score
K - Fold

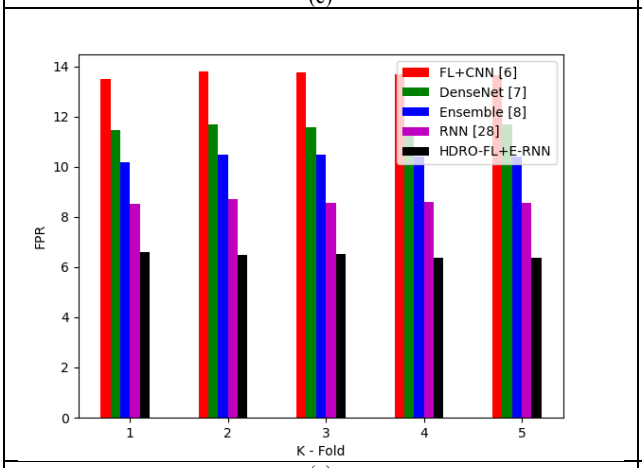
(b)


FL+CNN [6]
DenseNet [7]
Ensemble [8]
RNN [28]
HDRO-FL+E-RNN
FDR
K - Fold

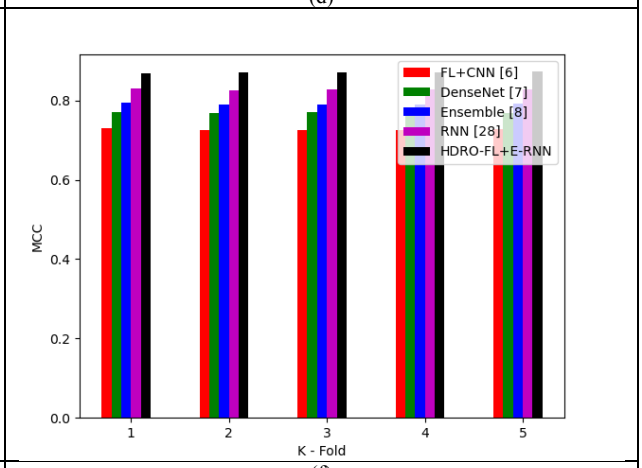
(c)


FL+CNN [6]
DenseNet [7]
Ensemble [8]
RNN [28]
HDRO-FL+E-RNN
FNR
K - Fold

(d)


FL+CNN [6]
DenseNet [7]
Ensemble [8]
RNN [28]
HDRO-FL+E-RNN
FPR
K - Fold

(e)


FL+CNN [6]
DenseNet [7]
Ensemble [8]
RNN [28]
HDRO-FL+E-RNN
MCC
K - Fold

(f)

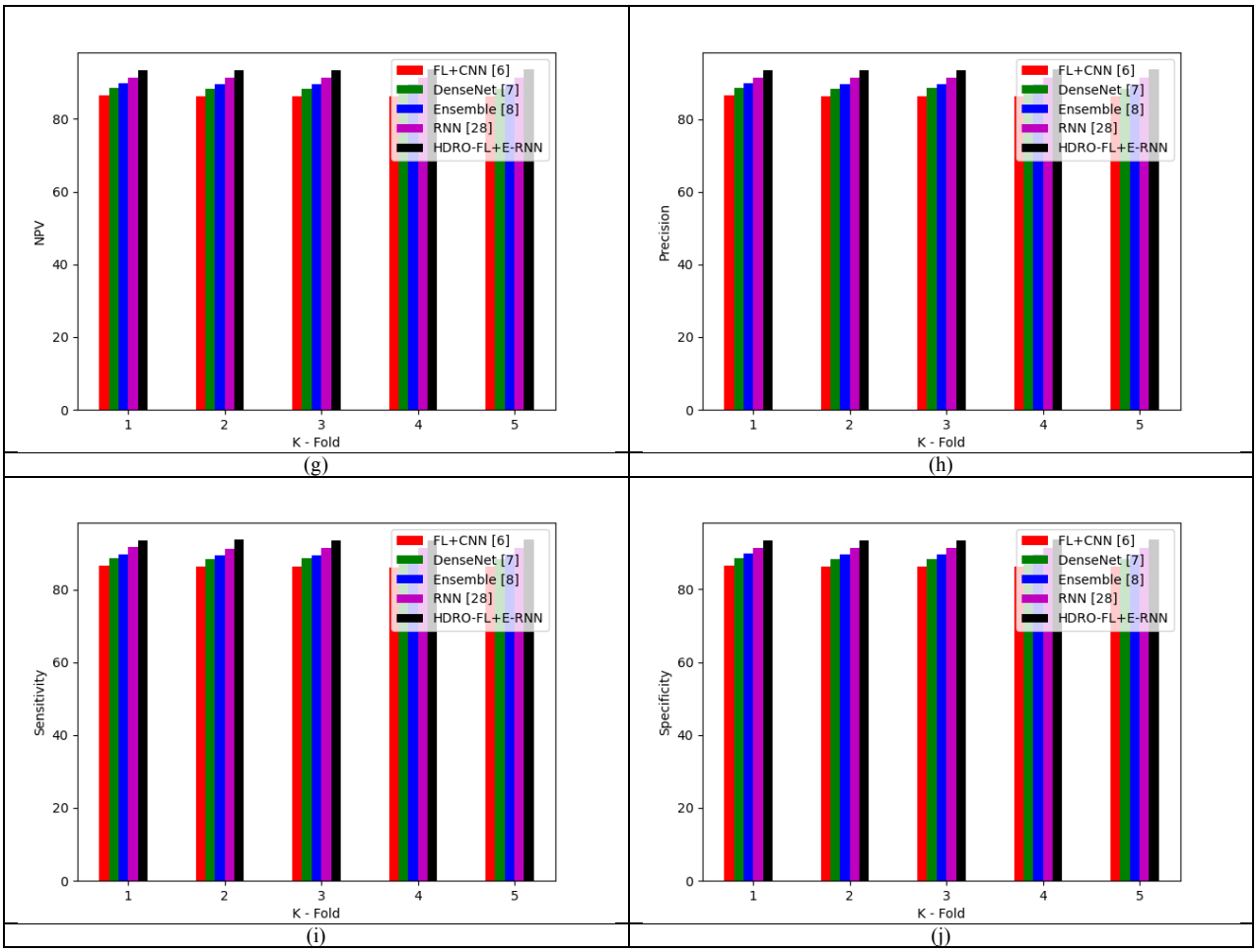


Fig. 10. K-fold validation of the FL-assisted "breast cancer diagnosis model" over existing appraoches regarding "(a) Accuracy, (b) F1 Score, (c) FDR, (d) FNR, (e) FPR, (f) MCC, (g) NPV, (h) Precision, (i) Sensitivity, and (j) Specificity"

R. Performance evaluation

The performance estimation of the designed FL-assisted "breast cancer diagnosis model" over existing heuristic algorithms is represented in Table II. From the evaluation, the accuracy of the designed HDRO-FL+E-RNN model has exhibited their efficiency, which is 3.7%, 3.8%, 1.5%, and 1.8% progressed than PSO-FL+E-RNN, DHOA-FL+E-RNN, DA-FL+E-RNN and RDA-FL+E-RNN, respectively. Therefore, the superior effectiveness of the designed model is ensured while evaluating existing algorithms.

TABLE II. INVESTIGATION ON THE SUGGESTED FL-ASSISTED "BREAST CANCER DIAGNOSIS MODEL" OVER HEURISTIC ALGORITHMS

TERMS	PSO-FL+E-RNN [31]	DHOA-FL+E-RNN [32]	DA-FL+E-RNN [26]	RDA-FL+E-RNN [27]	HDRO-FL+E-RNN
"Accuracy"	89.93548	89.91935	91.95161	93.31613	93.53871
"Sensitivity"	89.86897	89.67721	91.88239	93.14426	93.41643
"Specificity"	90.00326	90.16607	92.02214	91.48746	93.25952
"Precision"	90.1571	90.28314	92.14744	92.60409	93.38658
"FPR"	9.996744	9.83393	7.977857	6.512537	5.740475
"FNR"	10.13103	10.32279	8.11761	6.455737	5.583573
"NPV"	90.00326	90.16607	92.02214	93.48746	93.95952
"FDR"	9.842898	9.71686	7.852564	6.395907	5.613419
"F1-Score"	90.0128	89.97916	92.01472	93.57417	94.4015
"MCC"	79.87023	79.84006	83.90261	87.03122	88.67622

S. Comparative analysis

The performance analysis of the designed HDRO-FL+E-RNN-based breast cancer detection is given in Table III. In the course of estimation, HDRO-FL+E-RNN gets maximum performance over other existing diagnosis models in terms of FDR, which is 24%, 59%, 67.8%, and 62.7% more advanced than FL+CNN, DenseNet, Ensemble, and RNN, respectively. Consequently, the maximum performance is achieved while estimating with other diagnosis models.

TABLE III. EVALUATION ON THE SUGGESTED FL-ASSISTED "BREAST CANCER DIAGNOSIS MODEL" OVER CONVENTIONAL ALGORITHMS

TERMS	FL+CNN [6]	DenseNet [7]	Ensemble [8]	RNN [28]	HDRO-FL+E-RNN
"Accuracy"	94.22581	89.56452	86.48387	88.40323	95.72581
"Sensitivity"	94.11953	89.6133	86.25759	88.30297	95.74944
"Specificity"	94.33409	89.51482	86.71443	88.50537	95.70173
"Precision"	94.42129	89.6993	86.86836	88.67137	95.78005
"FPR"	5.665907	10.48518	13.28557	11.49463	4.298274
"FNR"	5.880473	10.38671	13.74241	11.69703	4.250559
"NPV"	94.33409	89.51482	86.71443	88.50537	95.70173
"FDR"	5.578711	10.3007	13.13164	11.32863	4.219949
"F1-Score"	94.27017	89.65628	86.5619	88.48679	95.76474
"MCC"	88.45141	79.12743	72.96901	76.80601	91.4509

T. Statistical analysis

The statistical evaluation of the designed HDRO-FL+E-RNN is given in Table IV over other heuristic algorithms. The "considered optimization algorithms are in stochastic nature and the experiment is executed five times. This analysis is carried out by considering the measures like best, worst, mean and standard deviation. The mean is the average value of the best and worst values and the median is referred to as the center point of the best and worst values whereas the standard deviation is represented as the degree of deviation between each execution". This analysis is demonstrated that the suggested model gets maximum performance over statistical measures and exhibited their efficiency.

TABLE IV. STATISTICAL EVALUATION OF THE FL-ASSISTED "BREAST CANCER DIAGNOSIS MODEL" OVER HEURISTIC ALGORITHMS

Measures	DHOA-FL+E-RNN [32]	DA-FL+E-RNN [26]	RDA-FL+E-RNN [27]	HDRO-FL+E-RNN
Worst	1.46299792	2.757118601	4.973125012	3.696535722
Best	1.191480828	1.177062843	1.396748346	1.243436415
Mean	1.28660055	1.425584989	1.880557774	2.030546813
Median	1.265692289	1.448225215	1.396748346	1.920687
Standard deviation	0.093410825	0.323868815	1.155700901	0.84622468

VII. CONCLUSION

This model has collected mammogram images related to breast cancer and adopted the concept of FL from the affected individuals. This concept has been used for reducing the processing time and ensuring better performance of the proposed model. Then, the FL-assisted gathered images were fed to the feature extraction phase, where the Densenet architecture was utilized for extracting the features. Then, the E-RNN has focused on classifying the extracted features for detecting breast cancer. Here, the performance enhancement was carried out by tuning certain parameters in the RNN by HDRO to achieve accurate classification results. With the evaluation, the accuracy of the designed HRDO-FL+E-RNN was 1.5%, 6.8%, 10.6%, and 8.2% more advanced than FL+CNN, DenseNet, Ensemble, and RNN, respectively. Therefore, the experimental results have demonstrated the effectiveness of the suggested breast cancer diagnosis model compared with conventional approaches using diverse quantitative measures.

Decleation:

- No funding consider from any external agiency. All work are self sponisered.
- This work is not consider and published anywhere.

Conflict of Interest: no conflict of interest.

References

[1] Enas M.F.El Houby and Nisreen I.R.Yassin,"Malignant and non-malignant classification of breast lesions in mammograms using convolutional neural networks,"Biomedical Signal Processing and Control, Vol 70, no. 102954, September 2021.

[2] Ines Feki, Sourour Ammar, Yousri Kessentini and Khan Muhammad,"Federated learning for COVID-19 screening from Chest X-ray images,"Applied Soft Computing, Vol. 106, no. 107330, July 2021.

[3] Irum Hirra; Mubashir Ahmad; Ayaz Hussain; M. Usman Ashraf; Iftikhar Ahmed Saeed; Syed Furqan Qadri; Ahmed M. Alghamdi; Ahmed S. Alfakeeh, "Breast Cancer Classification From Histopathological Images Using Patch-Based Deep Learning Modeling," IEEE Access, vol. 9, pp. 24273-24287, 2021.

[4] Zhiqiong Wang; Mo Li; Huaxia Wang; Hanyu Jiang; Yudong Yao; Hao Zhang; Junchang Xin, "Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion With CNN Deep Features," IEEE Access, vol. 7, pp. 105146-105158, 2019.

[5] J. Zheng, D. Lin, Z. Gao, S. Wang, M. He, and J. Fan, "Deep Learning Assisted Efficient AdaBoost Algorithm for Breast Cancer Detection and Early Diagnosis," IEEE Access, vol. 8, pp. 96946-96954, 2020.

[6] Dawid Połap and Marcin Woźniak,"Meta-heuristic as manager in federated learning approaches for image processing purposes,"Applied Soft Computing, Vol. 113, Part A, no. 107872, December 2021.

[7] K. V. Priya and J. Dinesh Peter, "A federated approach for detecting the chest diseases using DenseNet for multi-label classification," Complex & Intelligent Systems, 2021.

[8] A. Saber, M. Sakr, O. M. Abo-Seida, A. Keshk and H. Chen, "A Novel Deep-Learning Model for Automatic Detection and Classification of Breast Cancer Using the Transfer-Learning Technique," IEEE Access, vol. 9, pp. 71194-71209, 2021.

[9] C. Chen, Y. Wang, J. Niu, X. Liu, Q. Li and X. Gong, "Domain Knowledge Powered Deep Learning for Breast Cancer Diagnosis Based on Contrast-Enhanced Ultrasound Videos," IEEE Transactions on Medical Imaging, vol. 40, no. 9, pp. 2439-2451, Sept. 2021.

[10] P. Liu, B. Fu, S. X. Yang, L. Deng, X. Zhong and H. Zheng, "Optimizing Survival Analysis of XGBoost for Ties to Predict Disease Progression of Breast Cancer," IEEE Transactions on Biomedical Engineering, vol. 68, no. 1, pp. 148-160, Jan. 2021.

[11] B. Fu, P. Liu, J. Lin, L. Deng, K. Hu and H. Zheng, "Predicting Invasive Disease-Free Survival for Early Stage Breast Cancer Patients Using Follow-Up Clinical Data," IEEE Transactions on Biomedical Engineering, vol. 66, no. 7, pp. 2053-2064, July 2019.

[12] K. Park, W. Chen, M. A. Chekmareva, D. J. Foran and J. P. Desai, "Electromechanical Coupling Factor of Breast Tissue as a Biomarker for Breast Cancer," IEEE Transactions on Biomedical Engineering, vol. 65, no. 1, pp. 96-103, Jan. 2018.

[13] Anna N Karahaliou, Ioannis S Boniatis, Spyros G Skiadopoulos, Filippos N Sakellaropoulos, Nikolaos S Arikidis, Eleni A Likaki, George S Panayiotakis, Lena I Costaridou, "Breast Cancer Diagnosis: Analyzing Texture of Tissue Surrounding Microcalcifications," IEEE Transactions on Information Technology in Biomedicine, vol. 12, no. 6, pp. 731-738, Nov. 2008.

[14] L. Song, W. Hsu, J. Xu and M. van der Schaar, "Using Contextual Learning to Improve Diagnostic Accuracy: Application in Breast Cancer Screening," IEEE Journal of Biomedical and Health Informatics, vol. 20, no. 3, pp. 902-914, May 2016.

[15] K. Drukker, C. A. Sennett and M. L. Giger, "Automated Method for Improving System Performance of Computer-Aided Diagnosis in Breast Ultrasound," IEEE Transactions on Medical Imaging, vol. 28, no. 1, pp. 122-128, Jan. 2009.

[16] K. Drukker, C. A. Sennett and M. L. Giger, "Automated Method for Improving System Performance of Computer-Aided Diagnosis in Breast Ultrasound,"IEEE Transactions on Medical Imaging, vol. 28, no. 1, pp. 122-128, Jan. 2009.

[17] H. M. Whitney, H. Li, Y. Ji, P. Liu and M. L. Giger, "Comparison of Breast MRI Tumor Classification Using Human-Engineered Radiomics, Transfer Learning From Deep Convolutional Neural Networks, and Fusion Methods," Proceedings of the IEEE, vol. 108, no. 1, pp. 163-177, Jan. 2020.

[18] Y. A. Glickman, O. Filo, U. Nachaliel, S. Lenington, S. Amin-Spector and R. Ginor, "Novel EIS postprocessing algorithm for breast cancer diagnosis," IEEE Transactions on Medical Imaging, vol. 21, no. 6, pp. 710-712, June 2002.

[19] D. A. Woten and M. El-Shenawee, "Broadband Dual Linear Polarized Antenna for Statistical Detection of Breast Cancer," IEEE Transactions on Antennas and Propagation, vol. 56, no. 11, pp. 3576-3580, Nov. 2008.

[20] Q. Wuniri, W. Huangfu, Y. Liu, X. Lin, L. Liu and Z. Yu, "A Generic-Driven Wrapper Embedded With Feature-Type-Aware Hybrid Bayesian Classifier for Breast Cancer Classification," IEEE Access, vol. 7, pp. 119931-119942, 2019.

[21] Anuj J. Kapadia, Amy C. Sharma, Georgia D. Tourassi, Janelle E. Bender, Calvin R. Howell, Alexander S. Crowell, Matthew R. Kiser, Brian P. Harrawood, Ronald S. Pedroni and Carey E. Floyd, "Neutron Stimulated Emission Computed Tomography for Diagnosis of Breast Cancer," IEEE Transactions on Nuclear Science, vol. 55, no. 1, pp. 501-509, Feb. 2008.

[22] H. Lee, J. Park and J. Y. Hwang, "Channel Attention Module With Multiscale Grid Average Pooling for Breast Cancer Segmentation in an Ultrasound Image," IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, vol. 67, no. 7, pp. 1344-1353, July 2020.

[23] Uttam M. Pal, Anil Vishnu GK, Gayatri Gogoi, Saeed Rila, Saahil Shroff, Gokul AM, Pronami Borah, Manoj Varma, Vishnu Kurpad, Deb Baruah, Jayant S. Vaidya, Hardik J. Pandy, "Towards a Portable Platform Integrated With Multispectral Noncontact Probes for Delineating Normal and Breast Cancer Tissue Based on Near-Infrared Spectroscopy," IEEE Transactions on Biomedical Circuits and Systems, vol. 14, no. 4, pp. 879-888, Aug. 2020.

[24] R. K. Samala, H. -P. Chan, L. Hadjiiski, M. A. Helvie, C. D. Richter and K. H. Cha, "Breast Cancer Diagnosis in Digital Breast Tomosynthesis: Effects of Training Sample Size on Multi-Stage Transfer Learning Using Deep Neural Nets," IEEE Transactions on Medical Imaging, vol. 38, no. 3, pp. 686-696, March 2019.

[25] M. Sameti, R. K. Ward, J. Morgan-Parkes and B. Palcic, "Image Feature Extraction in the Last Screening Mammograms Prior to Detection of Breast Cancer," IEEE Journal of Selected Topics in Signal Processing, vol. 3, no. 1, pp. 46-52, Feb. 2009.

[26] Mohammad Jafari and Mohammad Hossein Bayati Chaleshtari, " Using dragonfly algorithm for optimization of orthotropic infinite plates with a quasi-triangular cut-out", European Journal of Mechanics A/Solids, vol. 66, pp.1-14, 2017.

[27] Amir Mohammad Fathollahi-Fard, Mostafa Hajiaghaei-Keshteli, and Reza Tavakkoli-Moghaddam, "Red deer algorithm (RDA): a new nature-inspired meta-heuristic", Soft Computing, October 2020.

[28] Hager Saleh, Sara F. Abd-el ghany, Hashem Alyami, and Wael Alosaimi, "Predicting Breast Cancer Based on Optimized Deep Learning Approach," Artificial Intelligence and Machine Learning-Driven Decision-Making, 2022.

[29] Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger, "Densely Connected Convolutional Networks," Computer Vision and Pattern Recognition, 2016.

[30] Constance Beguier, Jean Ogier du Terrail, Iqraa Meah, Mathieu Andreux, Eric W. Tramel, "Differentially Private Federated Learning for Cancer Prediction," Machine Learning, 2021.

[31] Vijayalakshmi S, John A, Sunder R, Senthilkumar Mohan, Sweta Bhattacharya, Rajesh Kaluri, Guang Feng, Usman Tariq, "Multi-modal prediction of breast cancer using particle swarm optimization with non-dominating sorting," International Journal of Distributed Sensor Networks, 2020.

[32] Weitao Ha and Zahra Vahedi, "Automatic Breast Tumor Diagnosis in MRI Based on a Hybrid CNN and Feature-Based Method Using Improved Deer Hunting Optimization Algorithm," Swarm Intelligence and Neural Network Schemes for Biomedical Data Evaluation, 2021.