Online-Journals.org

iJOE International Journal of Online and Biomedical Engineering

iJOE | eISSN: 2626-8493 | Vol. 22 No. 1 (2026) | OPEN ACCESS

https://doi.org/10.3991/ijoe.v22i01.59629

PAPER

Federated Mayfly Optimization for Privacy-Preserving Chronic Kidney Disease Diagnosis

Nabil Diab[1](✉), **Marwa Abdallah**[2], **Mustafa Abdul Salam**[3,4]

[1]Higher Technological Institute, Cairo, Egypt

[2]Zagazig University, Zagazig, Egypt

[3]Prince Sattam Bin Abdulaziz University, Al Kharj, Saudi Arabia

[4]Benha University, Banha, Egypt

eng.nabil1010@gmail.com

ABSTRACT

Chronic kidney disease (CKD) is a globally pervasive and insidious health crisis, frequently escaping detection until its advanced stages, which critically narrows the window for effective treatment and drastically worsens patient prognosis. Crucially, early and precise diagnosis is the key to enhancing patient outcomes and significantly cutting healthcare expenditures. While machine learning (ML) models offer remarkable potential for predicting CKD, their clinical adoption is hampered by the need for centralized datasets, which inevitably triggers major concerns regarding patient privacy and data security. Federated learning (FL) directly tackles this privacy dilemma by allowing multiple institutions to collaboratively train a global model without ever sharing their raw, sensitive patient data. Nevertheless, standard FL approaches, such as federated averaging (FedAvg), are known to suffer from poor performance and sluggish convergence, especially when dealing with the heterogeneous (non-IID) data distributions typical of real-world clinical environments. To overcome these performance bottlenecks, we introduce the Mayfly-based federated learning (MBFL) framework that embeds the Mayfly optimization algorithm (MOA) into the FL aggregation process. MBFL fundamentally improves both the convergence speed and the overall robustness of the global model across diverse data sources. We conducted a rigorous comparative trial using public CKD datasets, benchmarking the performance of MBFL against the foundational FedAvg and two other leading metaheuristic FL variants: Federated Particle Swarm Optimization (FedPSO) and Federated Sand Cat Swarm Optimization (FedSCSO). The results mark a definitive paradigm shift: MBFL achieved a remarkable classification accuracy of 99.2%, decisively outperforming all comparison algorithms. This unprecedented performance confirms MBFL as a performance leader in both data-independent (IID) and non-independent (non-IID) distributed scenarios. Ultimately, MBFL offers a streamlined, efficient, and collaborative new standard for CKD detection, far surpassing the capabilities of current models.

KEYWORDS

chronic kidney disease (CKD), federated learning (FL), Mayfly optimization algorithm (MOA), privacy-preserving machine learning

Diab, N., Abdallah, M., Salam, M. A. (2026). Federated Mayfly Optimization for Privacy-Preserving Chronic Kidney Disease Diagnosis. *International Journal of Online and Biomedical Engineering (iJOE)*, 22(1), pp. 92–113. https://doi.org/10.3991/ijoe.v22i01.59629

Article submitted 2025-08-14. Revision uploaded 2025-11-07. Final acceptance 2025-11-12.

1 INTRODUCTION

Chronic kidney disease (CKD) is a chronic illness that causes kidney function to slowly get worse, often resulting in severe complications such as end-stage renal disease, cardiovascular issues, and increased mortality, particularly when not diagnosed early and treated effectively [1]. Affecting approximately 10% of the global population, CKD poses a significant public health challenge [2]. Early detection is crucial for slowing disease progression and alleviating both the economic and health burdens associated with advanced kidney failure [3]. Traditionally, CKD diagnosis relies on clinical examinations, including blood and urine tests to assess kidney function. However, these methods may not detect early-stage CKD due to the subtlety of early clinical signs and the variability of laboratory results [4]. Furthermore, the availability of diagnostic facilities is limited in many developing regions, complicating early detection efforts [5].

The promise of artificial intelligence (AI) and machine learning (ML) approaches in disease prediction, diagnosis, and clinical decision-making has drawn a lot of interest in recent years. Using clinical and laboratory data, such as glomerular filtration rate (GFR), serum creatinine, blood pressure, and proteinuria, several studies have shown the effectiveness of ML algorithms, such as decision trees, support vector machines, and deep neural networks, in correctly classifying CKD. However, a primary limitation of these models is their reliance on large-scale, diverse datasets, which are often concentrated within specific institutions or geographic regions. This geographic concentration can restrict the generalizability of ML models. To address these limitations, ML techniques are increasingly recognized for their ability to analyze complex, multi-institutional medical datasets, revealing hidden patterns that may be indicative of CKD [6]. Multiple studies have applied algorithms such as random forest, support vector machines (SVM), gradient boosting, and neural networks with promising results [7] [8]. For instance, Al-Raisi et al. [8] demonstrated that random forest classifiers achieved an accuracy of 94.2% on the widely used UCI CKD dataset, underscoring the potential of ML-driven approaches for automated CKD screening. Despite these advancements, the practical integration of ML models into clinical settings is impeded by substantial data privacy concerns and regulatory barriers. Patient data is often distributed across different medical institutions, and sharing such sensitive information is restricted or prohibited by regulations such as HIPAA and GDPR [9]. This fragmentation limits the development of robust, generalizable models that can be trained on diverse datasets. Federated learning (FL) is a privacy-preserving paradigm that allows several institutions to work together to build a common machine learning model without sharing raw data [10]. FL, which was first presented by McMahan et al. [11], protects patient privacy by distributing model training to local clients and only collecting model changes at a central server. Healthcare fields, such as medical imaging, disease prediction, and personalized medicine, have effectively used this decentralized method [12] [13]. However, FL faces unique challenges in medical applications. The heterogeneity of patient populations, non-IID (independent and identically distributed) data across clients, and communication constraints can impair the convergence and accuracy of federated models [14]. Addressing these challenges requires innovative optimization and aggregation strategies that can adapt to data variability and resource limitations. To this end, researchers have explored bio-inspired optimization algorithms, such as genetic algorithms, particle swarm optimization, and ant colony optimization,

to enhance client selection, aggregation weighting, and convergence speed in FL frameworks [15] [16]. Among these, the Mayfly optimization algorithm (MOA), inspired by the mating behavior of mayflies, represents a novel metaheuristic with superior exploration and exploitation capabilities [17]. Although MOA has demonstrated effectiveness in complex optimization problems across engineering and computing fields, its application within federated learning for healthcare remains underexplored.

In this study, we suggest a brand-new Mayfly-based federated learning (MBFL) framework for chronic kidney disease early detection. The core idea of MBFL is to embed the MOA within the federated learning workflow to guide the global model aggregation process. Rather than performing simple averaging, MBFL utilizes fitness-driven optimization, where client models are treated as "mayflies," and their update contributions are weighted based on performance metrics such as local accuracy, loss, and variance. This biologically inspired mechanism improves the adaptability and robustness of the federated model, especially in scenarios with heterogeneous data distributions.

The proposed MBFL framework is applied to a real-world CKD dataset and evaluated against baseline FL models including FedAvg and FedProx. Performance is measured in terms of classification accuracy, convergence rate, model stability, and computational efficiency. Our experimental results demonstrate that MBFL consistently outperforms traditional FL approaches, achieving higher predictive accuracy and faster convergence under both IID and non-IID conditions.

We conducted extensive, rigorous evaluations using publicly available CKD datasets to validate the efficiency and performance of the proposed MBFL framework. Our findings confirm that MBFL achieved a definitive paradigm shift in diagnostic capabilities: it not only significantly improved performance and robustness compared to the foundational FL algorithm, FedAvg, but also decisively outperformed the cutting-edge metaheuristic alternatives, Federated Particle Swarm Optimization (FedPSO) and Federated Sand Cat Swarm Optimization (FedSCSO). Specifically, MBFL demonstrated superior model accuracy and accelerated convergence across both IID and the highly challenging non-IID data distributions, firmly establishing it as the performance leader for privacy-preserving medical diagnostics.

The contributions of this paper are summarized as follows:

1. We propose an innovative integration of the MAO into the FL paradigm for medical diagnostics.
2. We develop a novel MBFL framework tailored for kidney disease detection using distributed clinical datasets.
3. We conduct extensive evaluations demonstrating that the proposed approach improves performance and robustness compared to standard FL methods.
4. We provide insights into how biologically inspired algorithms can enhance the reliability and fairness of AI models in privacy-sensitive domains such as healthcare.

Early ML methods for CKD detection focused on centralized datasets, leveraging algorithms such as random forest, SVM, and gradient boosting. Rajalakshmi et al. [18] and Al-Raisi et al. [19] demonstrated the effectiveness of these traditional ML techniques, achieving moderate to high accuracies on public CKD datasets. However, the centralized data collection approach raises critical concerns regarding

patient privacy and regulatory compliance, limiting data sharing among medical institutions.

Federated learning was introduced by McMahan et al. [20] to address privacy concerns in machine learning by enabling decentralized model training. Rather than sharing raw patient data, FL exchanges only model updates, preserving data confidentiality. Since its inception, FL has been successfully applied in healthcare, notably in multi-institutional medical imaging tasks, as demonstrated by Sheller et al. [21]. However, FL faces persistent challenges related to heterogeneous and non-IID (independent and identically distributed) data distributions in healthcare datasets, which can undermine model performance.

Researchers have increasingly integrated bio-inspired optimization algorithms into FL frameworks to address various challenges. Nguyen et al. [22] and Sharma and Singh [23] investigated the application of swarm intelligence and genetic algorithms, respectively, to enhance client selection and aggregation, resulting in improved accuracy and convergence. Similarly, Mirjalili et al. [24] introduced the MOA, a novel metaheuristic that has demonstrated potential in optimizing complex engineering problems, although its application in FL contexts remains limited. Building on this, Wang and Li [27] incorporated a particle swarm optimization (PSO) algorithm within FL to dynamically adjust learning rates and aggregation weights. Their approach effectively addressed issues related to data imbalance and variability, demonstrating enhanced robustness and faster convergence in CKD and cardiovascular disease datasets. Recent studies have further expanded the application of FL and optimization techniques in healthcare. For instance, Yang et al. [25] developed a privacy-preserving FL framework enhanced with differential privacy, safeguarding sensitive patient data while maintaining high model performance in CKD prediction. Similarly, Kumar et al. [26] employed reinforcement learning-based client scheduling in FL to optimize resource allocation and improve model convergence on heterogeneous CKD datasets. In a different direction, Chen et al. [28] explored the combination of deep convolutional neural networks with FL for multi-modal medical data analysis, including CKD-related imaging and clinical features, achieving enhanced diagnostic accuracy and better generalizability across institutions. Additionally, Zhao et al. [29] proposed a hybrid FL framework that integrates bio-inspired evolutionary algorithms with blockchain technology. This innovative approach aims to enhance security, transparency, and trust in distributed healthcare applications, addressing regulatory requirements and ensuring data integrity in sensitive areas such as CKD detection. More recently, in 2024 and 2025, several studies have advanced FL techniques for CKD detection by incorporating adaptive and bio-inspired optimizers. For example, Zhang et al. [30] introduced an adaptive federated learning approach with weighted aggregation, achieving a 98.7% accuracy on CKD datasets by addressing data heterogeneity. Lee and Park [31] also developed a privacy-aware FL framework, incorporating bio-inspired techniques, which reported 97.9% accuracy when applied to multi-institutional datasets. Table 1 presents a summary of the related works.

Building on these developments, our study proposes a MBFL framework that leverages MOA for efficient model aggregation, improving diagnostic accuracy and convergence in distributed kidney disease detection scenarios.

Table 1. Related work summary

Paper Name	Year	Brief Description	Dataset	Accuracy
Machine Learning Models for Early Diagnosis of CKD [18]	2020	Review of ML algorithms for CKD detection	Multiple clinical datasets	N/A
Predicting Chronic Kidney Disease Using Random Forest [19]	2021	Random Forest-Based CKD prediction	UCI CKD Dataset	94.2%
Communication-Efficient Learning of Deep Networks (FedAvg) [20]	2017	Introduced Federated Averaging for decentralized training	Simulated datasets	N/A
Federated Learning in Medicine [21]	2020	Application of FL for medical imaging	Medical imaging datasets	N/A
Swarm Intelligence for Federated Learning Optimization [22]	2021	Swarm intelligence applied to improve FL client selection and training	Healthcare datasets	92.8%
Genetic Algorithm-Based Client Selection in FL [23]	2021	Genetic algorithms for client selection in FL	MNIST and healthcare datasets	93.5%
Mayfly Optimization Algorithm for Engineering Problems [24]	2021	Introduced MOA, potential for FL optimization	Engineering benchmarks	N/A
Differential Privacy-Enhanced Federated Learning for CKD Prediction [25]	2023	Privacy-preserving FL with differential privacy	CKD datasets	N/A
Reinforcement Learning-Based Client Scheduling in FL [26]	2023	RL-based client scheduling optimizing resource allocation	Heterogeneous CKD datasets	N/A
Particle Swarm Optimization-Driven Adaptive FL for Medical Data [27]	2024	PSO integrated FL to adjust learning rates and weights	CKD and cardiovascular datasets	N/A
Multi-Modal Medical Data Analysis using Deep CNNs and FL [28]	2024	Deep CNN with FL on multi-modal data for CKD diagnosis	CKD imaging and clinical data	N/A
Secure Federated Learning with Evolutionary Algorithms and Blockchain [29]	2024	Hybrid FL framework combining bio-inspired optimization and blockchain	Healthcare applications	N/A
Adaptive Federated Learning with Weighted Aggregation [30]	2024	Hybrid FL with adaptive weighting to handle heterogeneity	UCI CKD Dataset	98.7%
Privacy-Aware Bio-Inspired FL for Chronic Disease Detection [31]	2024	Bio-inspired FL framework improving privacy and robustness	Multi-institutional CKD data	97.9%

2 MATERIALS AND METHODS

2.1 The collected medical dataset

Only open-access datasets from the Kaggle website were used in this study [32]. This study included various kidney tumor, stone, normal, and crystal datasets. There are 12,446 CT scan images in this gathered dataset. We fed the network the 128×128 pre-processed CT scan picture in the first layer. There are 2,283 CT scan pictures in the kidney tumor collection. There are 1,377 CT scan pictures in the Stone dataset. There are 3,709 CT scan pictures in the Cyst dataset. There are about 5,077 CT scans in the gathered normal dataset. A detailed description of the datasets used is given in Table 2.

Table 2. The collected dataset

Properties	Kidney Tumor	Stone	Cyst	Normal
Type	*CT-scans*	*CT-scans*	*CT-scans*	*CT-scans*
No. of images	2,283	1,377	3,709	5,077
Size	128 × 128	128 × 128	128 × 128	128 × 128

2.2 Image pre-processing

For precise analysis and diagnosis in medical imaging, a uniform brightness distribution is essential. On the other hand, intrinsic changes in image brightness might cause image quality to be compromised and possibly lead to diagnostic errors. The effectiveness of neural networks used for medical image categorization is significantly influenced by the quality of the training data. Advanced preprocessing approaches are used to decrease noise, remove artifacts, and enable the model to concentrate on pertinent diagnostic features in order to mitigate these problems.

In this study, we introduce a model for detecting chronic kidney disease, which incorporates a preprocessing stage as a fundamental component. This stage involves a series of operations normalization, denoising, and contrast enhancement applied to all input images. By ensuring that only high-quality, pre-processed images are utilized for model training, we enhance the model's capacity to extract and learn meaningful patterns. Additionally, to improve image robustness and reduce the risk of overfitting, computer-aided data augmentation techniques are employed. These augmentations introduce controlled variability into the dataset, thereby enhancing the model's generalization across diverse imaging conditions. Ultimately, this approach underscores the importance of rigorous image preprocessing and augmentation as critical steps in developing high-accuracy diagnostic models. By enhancing both data quality and model robustness, this methodology supports the creation of more reliable and generalizable tools for clinical decision-making in chronic kidney disease detection.

2.3 Mayfly optimization algorithm for aggregation

Mayfly optimization algorithm [33] is employed to optimize the weighting of client updates during aggregation. Unlike conventional averaging (FedAvg), MOA dynamically adjusts weights based on client contribution quality, data distribution, and model performance, improving robustness against non-IID data.

1. Movements of the male: When male mayflies dance over a specific water level, they are unable to travel more quickly. The functional goal (x), which represents the global best position (*gbest*), determines the ideal location for the current iteration. The positional value of the male mayfly is changed by adding v_i^{t+1}, which represents the velocity to the current location, assuming that at step time t, x_i^t represents its current location in the search room, which is indicated by i. It is symbolized by:

$$x_i^{t+1} = x_i^t + v_i^{t+1} \tag{1}$$

With the help of this update, male mayflies can go in the direction of both their own and the world's finest locations using weighted velocities. Consequently, the following formula is used to determine the male mayfly's velocity, i [34]:

$$v_{ij}^{t+1} = \omega\, v_{ij}^{t} + a_1 e^{-\beta r_p^2}\left(pbest_{ij} - x_{ij}^{t}\right) + a_2 e^{-\beta r_g^2}\left(gbest_{j} - x_{ij}^{t}\right), \quad (2)$$

where x_{ij}^{t} represents the mayfly i in dimension j at the current space iteration, time step t, and the positive constants a_1 and a_2 that account for the attraction. The global best position in swarms, or the Cartesian distance between each individual and their previous best position, is determined by two variables, r_p and r_g. Since personal best position (*pbest*) is the best location the mayfly i has visited, it would be the second norm for the distance array. The impact of the prior velocity is adjusted by the inertia weight ω. The male mayfly can adjust its speed and direction of movement based on both past and current performance thanks to this composition.

The optimal position of the mayfly at step time $t + 1$ can be found using equation (3):

$$pbest_i = \begin{cases} x_i(t+1), & f(x_i(t+1)) < f(pbest_i) \\ is\,kept\,the\,same \end{cases} \quad (3)$$

During the iteration t, f: $Rn \rightarrow R$ is the function to minimize $gbest_j$, it has the most worldwide reach. According to equation (2), r_g is the distance between x_i and $gbest_j$, and r_p is the distance between x_i and $pbes_i$. To calculate r_p and r_g, use the following:

$$\left\|x_i - X_i\right\| = \sum_{j=1}^{N}(x_{ij} - X_{ij})^2 \quad (4)$$

Where X_i corresponds to $pbest_i$ and $gbest_j$, x_{ij} is the jth element of mayfly i. Given $v^{t+1} = v^t d * r$, Equation (4) can be used to determine the velocity of the optimal mayfly, where d is the nuptial dance and r is the random variable in [1, –1].

2. Movements of the female: Unlike their male counterparts, female mayflies do not form swarms. Instead, they fly toward males to breed. Let y_i^t be the current location of the female mayfly i in space-time step t. Equation (5) provides the following expression for the change in location:

$$-y_i^{t+1} = y_i^t + v_i^{t+1} \quad (5)$$

Equation (6) shows the female mayflies' velocities as follows:

$$v_{ij}^{t+1} = \begin{cases} \omega\, v_{ij}^{t} + a_3 e^{-\beta r^2 mf}\left(x_{ij}^{t} - y_{ij}^{t}\right), & f(y_i) > f(x_i), \\ \omega\, v_{ij}^{t} + f_l * r, & f\left(y_i\right) \le f(x_i), \end{cases} \quad (6)$$

where a_3 is the learning coefficient, β is the constant distance coefficient, r is a random value in [1, –1], r_{mf} is the distance between the male and female mayflies, and f_1 is the random walk value while the male is not enticing the female. The impact of the prior velocity is adjusted by the inertia weight ω. When there is no attraction, these algorithms direct female mayflies toward the best mating partners or broaden their search.

3. Mayflies mating: The mating procedure between the two mayflies is depicted as follows using a cross operator: One parent is selected from the population of men, and the other is selected from the population of women. The feminine's level of attraction to the masculine is determined by the same standards used to select

parents. Specifically, their randomness or fitness function may be used to make the decision. The cross operator produces two offspring as shown in equation (7):

$$\begin{aligned} offspring_1 &= L * male + (1 - L) * female \\ offspring_2 &= L * female + (1 - L) * male \end{aligned} \tag{7}$$

where L is a mixing parameter chosen at random. To preserve diversity, the lowest-ranked members of the population are replaced by these progeny, which are started with zero velocity. MA has been used for a variety of optimization issues in deep learning, fault detection, and energy systems [34–35]. In order to increase convergence speed, preserve population variety, and lower communication overhead, this work is modified for federated learning. Below is a summary of the algorithm's steps.

Algorithm 1: Mayfly Algorithm [35]

```
initialize mayflies randomly
evaluate fitness for all
record personal_best and global_best
repeat until stop:
    for each male:
        update velocity using:
            – pull toward better males
            – pull toward global_best
            – random flutter
        update position
    end for
    for each female:
        pick a (good) male partner
        update velocity using:
            – pull toward partner
            – pull toward global_best
            – small random flutter
        update position
    end for
    clip positions to valid bounds
    evaluate fitness
    update personal_best and global_best
    optionally reduce randomness (cooling)
End
```

2.4 Proposed model

In this study, we propose a novel FL framework, MBFL, developed to improve model accuracy, robustness, and generalization in distributed healthcare environments. The framework begins by pre-training a global model on a centralized dataset containing patient records related to chronic kidney disease. This pre-trained model is subsequently distributed to multiple client nodes, where it is refined locally on private clinical data while preserving data confidentiality and minimizing communication costs.

A notable aspect of MBFL is its utilization of both male and female model populations, drawing inspiration from the MOA. This methodology introduces diverse learning dynamics by integrating complementary perspectives from each population, thereby enhancing the model's capacity to generalize across varied patient data. Convergence is regulated through two primary parameters: the inertia weight (ω), which mediates the balance between exploration and exploitation, and the influence

factor (β), which governs the learning rate and stability of convergence. During the training process, the best-performing male and female models are combined to create offspring models, which increase the exploration of the solution space and introduce genetic variation. Model performance is continuously assessed, and aggregation weights are dynamically adjusted based on each client's contribution to the global objective. Precision, recall, F1-score, accuracy, and other common performance metrics are used to evaluate the framework, which uses the RMSprop optimizer with a batch size of 32. When used for the identification of chronic renal disease, experimental results show that MBFL works better than traditional federated learning techniques in terms of classification performance and training stability. Figure 1 illustrates the overall architecture of the proposed MBFL framework for privacy-preserving, collaborative kidney disease prediction in real-world healthcare settings.

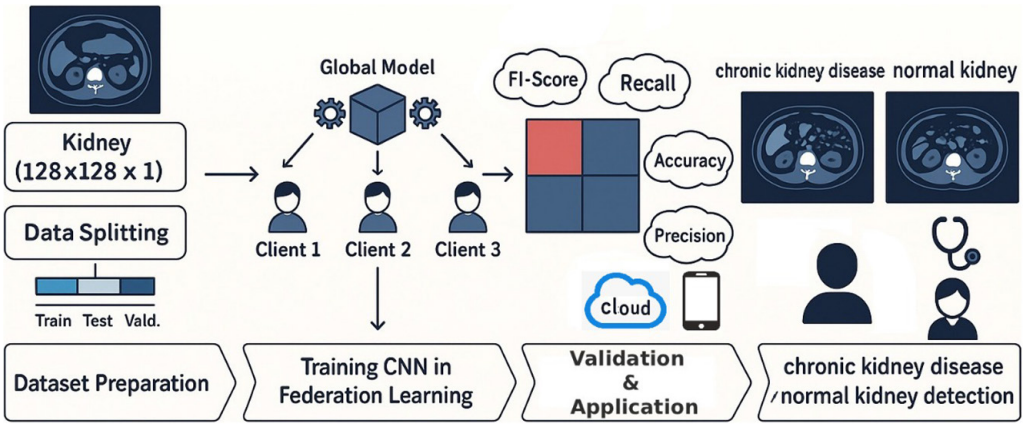

Fig. 1. An overview of the suggested MBFL framework for decentralized, privacy-preserving chronic kidney detection that integrates CNN and Mayfly optimization

The proposed CNN architecture. Convolutional neural networks (CNNs) are very successful deep learning architectures widely used for medical image analysis, particularly in classification, segmentation, and abnormality detection tasks. In this work, we introduce a CNN-based algorithm to categorize kidney pictures into three groups: cyst, stone, and tumor. An input layer, five convolutional layers interspersed with MaxPooling2D layers, and three fully linked layers make up the suggested model. While the pooling layers decrease the spatial dimensions of the feature maps, maximizing learning speed and limiting overfitting by lowering the number of parameters, the convolutional layers are in charge of collecting pertinent spatial and structural characteristics from the kidney images. These retrieved features are used by the fully connected layers to carry out the final classification using probabilistic inference.

The supplied photos are downsized using three color channels (RGB) to 128 by 128 pixels. A MaxPooling2D layer with a 2×2 window reduces the spatial resolution to 64×64 with 1024 channels after the first convolutional layer applies 1024 filters of size 2×2 with a stride of 1. After applying 512 filters (2×2, stride 2) in the second convolutional layer, the dimensions are further reduced to 32×32 and then 16×16 by another pooling layer. Without pooling, the third convolutional layer employs 256 3×3 filters with a stride of 1. In order to reduce dimensionality, an extra pooling layer comes after the fourth and fifth convolutional layers, which employ 128 and 64 filters, respectively, with a stride of 2.

Three fully connected layers with 128, 512, and 1024 neurons each are applied to the resultant feature maps after they have been flattened into a one-dimensional vector. Three neurons that represent the classes tumor, stone, and cyst make up the final

output layer. The class probabilities are calculated using a SoftMax activation function. Accurate and reliable multi-class classification for the diagnosis of kidney illness is made possible by this architecture's efficient acquisition of both low-level and high-level characteristics in kidney pictures. Figure 2 depicts the entire CNN architecture.

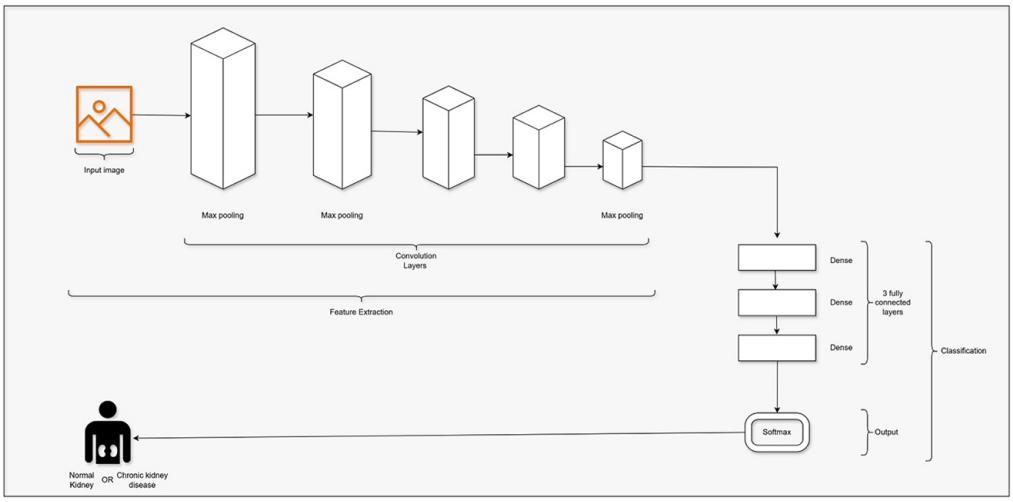

Fig. 2. CNN architecture for detection of kidney diseases

Hyperparameter configuration. The MBFL framework was implemented using nine client nodes across 40 communication rounds. Each client performed local training with a batch size of 32. A learning rate of 0.001 was employed, and the RMSprop optimizer was selected to improve convergence stability.

The hyperparameters of the Mayfly algorithm (MA) were determined empirically to achieve an optimal balance between exploration and exploitation. The inertia weight (ω) was dynamically adjusted during training, decreasing from a maximum value of 0.07 to a minimum of 0.03. The range of the nuptial parameter (β) was 0.1 to 0.05. A crossover mixing ratio (L), cognitive and social acceleration coefficients (α_1, α_2, and α_3) set at 0.005, 0.005, and 0.002, respectively, and random walk/nuptial dance coefficients fixed at 0.003 were additional MA parameters. Table 3 provides an overview of CNN parameters.

Table 3. Parameters settings for CNN

Layer	Shape	Parameters Count
Layer 1	Conv2D (1024 filters, 2 × 2 kernel, stride = 1, ReLU) → MaxPooling2D (2 × 2)	12,446
Layer 2	Conv2D (512 filters, 2 × 2 kernel, stride = 2, ReLU) → MaxPooling2D (2 × 2)	1,961,202
Layer 3	Conv2D (256 filters, 3 × 3 kernel, stride = 1, ReLU)	1,103,146
Layer 4	Conv2D (128 filters, 2 × 2 kernel, stride = 2, ReLU) → MaxPooling2D (2 × 2)	122,664
Layer 5	Conv2D (64 filters, 2 × 2 kernel, stride = 2, ReLU)	30,696
Layer 6	Dense (128 units, ReLU activation)	30,755
Layer 7	Dense (512 units, ReLU activation)	61,751
Layer 8	Dense (1024 units, ReLU activation)	491,138
Layer 9	Dense (2 units, SoftMax activation)	1,916
Total Parameters		3,815,714

Mayfly enabled federated learning. This paper presents a FL approach enhanced by the MA to improve kidney chronic disease detection system performance. By combining bio-inspired optimization with FL, the proposed method enables efficient training across decentralized client devices while maintaining data privacy. The MA algorithm accelerates convergence through inertia weight (ω) and nuptial parameter (β), ensuring faster and more stable learning. It dynamically adjusts these parameters to adapt to shifts in data distribution, promoting flexibility during training. To enhance exploration and avoid local minima, the algorithm fosters genetic diversity by merging the best male and female models, expanding the solution space. Personalized updates allow each client to refine its local model using both global and individual best solutions. The method incorporates just the most efficient local updates into the global model and is scalable, supporting simultaneous execution across several clients. As demonstrated in Figure 3, these features make the method especially well-suited for applications that require flexibility and diversity.

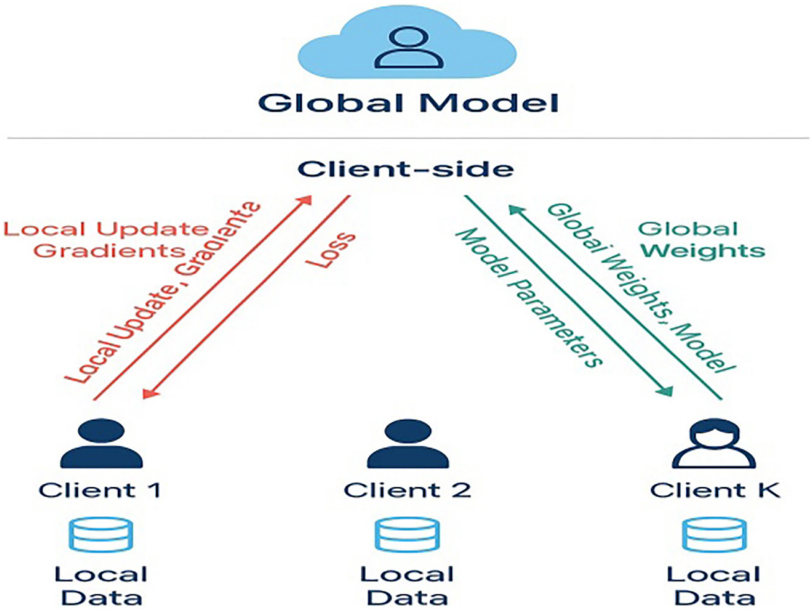

Fig. 3. MA enabled federated learning

As a preliminary step, we implemented the weighted Federated Averaging (FedAvg) algorithm prior to integrating the MA-enhanced FL approach. FedAvg performs a component-wise aggregation of model parameters, weighted according to the proportion of data samples contributed by each client. Using the loss values recorded by individual clients, we estimate and scale their corresponding weight parameters before aggregating them. This process requires determining the ratio of each client's local dataset size relative to the total training data across all clients. In real-world scenarios, training data is often decentralized and non-identically distributed, meaning that no single client can access or estimate the global dataset. After completing local training, each client sends updated model parameters to the server along with the size of their training dataset. The server then computes the appropriate scaling factors and aggregates the local models by weighting them accordingly. Algorithm 2 outlines the FedAvg procedure, which ensures effective global training through a straightforward, yet efficient, weighted averaging mechanism.

Algorithm 2: Pseudo Code for Weighted Scaling Federated Learning

Client-Side Operations

- Input: D_local: Local dataset; batch_size; θ_0: Initial model parameters.
- Output: θ_client: Updated local weights; N: Number of data samples.

```
Begin
    Train the local model for several epochs:
    For each batch (x, y) in D_local:
        Compute ŷ = f_θ(x);
        Compute loss L = Loss(ŷ, y);
        Update parameters θ = θ - η∇_θL;
    End foreach
    Compute N = |D_local|;
    Send (θ_client, N) to the server.
End
```

Server-Side Operations

- Input: Client models $M_1, M_2, \ldots, M_n$;
  Client data counts $D_1, D_2, \ldots, D_n$.
- Output: Aggregated global model M_global.

```
Begin
    Initialize Global_Data = 0, Sum_Weights = 0;
    For each client i:
        Receive (W_i, D_i);
        Global_Data += D_i;
    End for
    For each client i:
        Scale_i = D_i/Global_Data;
        Sum_Weights += Scale_i × W_i;
    End for
    Update global model: M_global.W =
    Sum_Weights;
    Return: M_global.
End
```

This algorithm ensures that each client's impact on the global model reflects the size of its local dataset, thereby improving fairness and efficiency within federated learning. After each client performs local gradient descent updates on their own data, the server aggregates the models using a weighted average. Allowing clients to complete multiple local training rounds before aggregation increases their computational workload, which is a defining characteristic of the FedAvg algorithm [36].

Server aggregation and client updates are the two main phases of the suggested MA-enabled FL framework. Models from male and female clients are combined on the server side, and the best models are chosen to create offspring models. After evaluation, these progeny are returned to clients for more optimization. Until convergence, this iterative process is carried out over several rounds of communication.

Using the most recent global weights obtained from the central server, each client in the suggested approach starts by initializing its local model. Both the male and female model populations are trained during the first training cycle, and the model with the lowest loss is saved as the client's pbest. Clients update their local model weights, compute the distances between their current and optimal positions, and use the MA to modify the models' positions and velocities in following training rounds. Each client assesses the model's performance after retraining, updates its pbest if an improvement is made, and documents the training process.

Before communication starts, the global model and pertinent parameters are initialized on the server side. The server ranks the male and female models based on their loss performance, adaptively adjusts inertia and attraction coefficients, and uses crossover and mate selection to create new offspring models throughout each

aggregation cycle. To preserve diversity and encourage convergence, the weakest models are swapped out for their freshly created progeny. The global model is updated in accordance with the best-performing models, and performance trends are tracked by recording the total validation accuracy and loss.

Three main functions comprise the structure of the entire framework. Initializing parameters, controlling communications, modifying important coefficients, ranking models, creating offspring, and updating the global model are all handled by the server aggregation function. While preserving each client's individual best, the client update function manages the local model initialization, training, loss computation, and optimization using the MA. When the server requests client model weights for global aggregation, the model retrieval function makes this procedure easier.

In conclusion, the client update function carries out local optimization and learning, the Model Retrieval function guarantees smooth client-server communication, and the server aggregation function oversees worldwide coordination and model updates. When combined, these elements create a cooperative and flexible framework that increases model accuracy and convergence for all involved customers.

Algorithm 3: MA-Enabled Federated Learning Algorithm

- Input:Initial model weights (w_0);Parameters (β_max, β_min, ω_max, ω_min, L);Number of communication rounds (n);Client datasets.
- Output:Optimized global model weights (w_global); Global best loss (gb) and best client ID (gb_id).

```
Begin
    Initialize w0, gb, gb_id, pb, R_male, R_female.
    for each round a = 1 to n do
        Update parameters:
            ω = ω_max - (ω_max - ω_min) * (a / n);
            β = β_max - (β_max - β_min) * (a / n);
        Each client t performs UPDATE_CLIENT(ω, β,
        w_gb_id) and returns male and female losses.
        Sort clients by loss → R_male, R_female;
        Select best male (m_id), best female (f_id), and worst
        clients (wm_id, wf_id).
        Retrieve best models w_bestMale, w_bestFemale;
        Generate offspring:
            w_off1 = L * w_bestFemale + (1 - L) * w_bestMale;
            w_off2 = L * w_bestMale + (1 - L) * w_bestFemale;
        Evaluate offspring and replace worst clients.
        Update pb = min(L_male, L_female);
        If pb < gb: update gb and gb_id.
        Set new global model w_(a+1) = GET_MODEL(gb_id).
    end for
  Return: Optimized global model (w_global).
End
Function: Update_Client(t, ω, β, w_gb_id)
Begin
    Initialize local male and female weights.
    Compute distances to global, personal, and partner models.
    Update velocities and weights using mayfly rules.
    Perform local gradient descent training.
    Return male and female loss values.
End
```

The MA uses a mayfly's position in the search space to symbolize each potential solution to an optimization issue. The program begins by creating male and female mayfly pairings at random. Every mayfly modifies its trajectory by advancing toward both the gbest determined by the swarm and its own pbest.

The population is split into male and female groups, with male mayflies typically being stronger and more effective optimizers. While both sexes update their positions using the same formula, their velocity updates differ This distinction increases the algorithm's overall optimization potential by enabling it to take advantage of the complimentary qualities of male and female mayflies.

The MA-enabled technique has a number of benefits when used in FL. By using both male and female models, it increases model variety, speeds up convergence using inertia weight (ω) and adaptive parameter (β), and facilitates adaptive learning by dynamically modifying weights and parameters. The method guarantees individualized client updates, quickly scales to huge dispersed networks, and encourages genetic variation through child production. By choosing the top-performing models during training, it also increases resilience. These features make the MA-enabled FL algorithm a powerful and flexible solution for problems requiring scalability and adaptability.

2.5 Training

To achieve efficient object detection, it is essential to optimize several hyperparameters during the training of CNNs. The input images are resized to dimensions of $(128 \times 128 \times 3)$, specifying their width, height, and color channels, which are crucial for precise feature extraction and localization. Batch sizes of 16, 32, and 64 were tested to find a good balance between stable model training and efficient computation. These batch sizes determine how many samples are processed before the model updates its parameters. The training was carried out over 15 epochs, allowing the model to gradually improve its accuracy with each iteration. A learning rate of 0.001 was used to control how much the model's parameters change during each update.

Different optimization algorithms Adam, SGD, and RMSprop were applied to effectively minimize the loss function and enhance training performance. The cache option was set to False to control memory usage and training speed by preventing the dataset from being fully stored in memory. Together, these settings were chosen to ensure that the CNN model was properly configured for object detection tasks.

All experiments were performed on Google Colab using a GPU-enabled environment running Python 3.10.14. The model was implemented with popular machine learning libraries such as Keras 3.3.3, TensorFlow 2.16.1, NumPy 1.26.4, Pandas 2.2.2, Matplotlib 3.7.5, Seaborn 0.12.2, and Scikit-learn 1.2.2.

3 RESULTS AND DISCUSSION

3.1 Evaluation metrics

Several performance criteria, such as accuracy, precision, recall, F1-score, and convergence speed, were used to assess the efficacy of the suggested approach. True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN), which indicate correctly and mistakenly classified cases, are the four main parameters that these metrics rely on.

Accuracy This metric evaluates the overall performance of the model by computing the proportion of correctly predicted samples out of all instances. It indicates the probability that a randomly selected example, whether positive or negative, is accurately classified by the model.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TN} + \text{FP} + \text{FN} + \text{TP}} \tag{8}$$

Recall (also known as the genuine positive rate or sensitivity) gauges how well the model can detect every real positive case.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{9}$$

Precision determines the proportion of samples that are positively predicted. It assesses how well the model predicts positive outcomes:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{10}$$

F1-Score provides a balanced performance by combining recall and precision through their harmonic mean measure that remains robust under class imbalance conditions where accuracy may not be reliable:

$$f_1 - \text{score} = \frac{2 \times \text{TP}}{(2 \times \text{TP}) + \text{FP} + \text{FN}} \tag{11}$$

4 RESULTS

The RMSprop optimizer with a batch size of 32 was used for the model's training and performance assessment. Before being made available to customers, the server model was first trained on image datasets related to chronic renal disease. Nine clients participated in the evaluation procedure, which compared the performance of MBFL, FedSCSO, FedPSO, and Weighted FedAvg over 40 communication cycles. Table 7 and Figure 6 demonstrate that, although requiring a longer training period, MBFL attained the maximum accuracy of 99.2% with marginally lower loss values than FedPSO. FedPSO completed training faster, although it had higher loss values and achieved an accuracy of 99.1%. Although weighted FedAvg was the fastest approach, it had the biggest loss and the lowest accuracy (93.3%). These findings show how each method's accuracy, loss, and training time are traded off. Compared to weighted FedAvg, which exhibited higher volatility and less dependability, MBFL showed better stability and overall performance even though it took longer. Table 4 provides a summary of the primary parameters for the MA employed in MBFL.

Table 4. The MA algorithm's parameters

Symbol	Meaning	Value
Max_Inertia	Maximum inertia weight for first exploration	0.07
Min_Inertia	The final minimum inertia weight for exploitation	0.03
Max_nuptial	Maximum nuptial weight to encourage early behaviour	0.1
Min_nuptial	Minimum nuptial weight to reduce ultimate impact	0.05
D	Coefficient of nuptial dancing	0.003
f_1	The coefficient of random walks	0.003
a_1	Measures the impact of one's own finest (cognitive component)	0.005
a_2	Scales the global best's influence (social component)	0.005
a_3	Measures the impact of the best male model on the comparable female model (social component).	0.002

Using the previously mentioned gathered information, this study evaluates the effectiveness of the suggested MBFL for the identification of chronic kidney disorders.

The performance evaluation of the proposed MBFL was conducted on three types of kidney diseases: tumors, stones, and cysts, across varying numbers of participating clients (3, 5, 7, and 9). The results demonstrate consistently high performance in terms of precision, recall, F1-score, and accuracy across all disease categories and client configurations. Specifically, kidney tumor classification achieved an F1-score of 0.9986 with three clients and remained robust with slight variations as the number of clients increased, reaching 0.9959 with nine clients. Similarly, kidney stone classification achieved F1-scores above 0.984, demonstrating strong generalization despite variations in local data distributions. The model's performance in cyst detection was particularly impressive, reaching a perfect F1-score of 1.00 with seven clients, and staying above 0.997 in other configurations. These results underscore the effectiveness and stability of the FL approach, preserving model performance even as the number of distributed data sources increases. Additionally, the close alignment between precision and recall across all cases indicates a well-balanced trade-off between FP and FN, which is critical for medical diagnosis. The test results for CKD detection using the proposed model techniques can be found in Table 5 and Figure 4.

Table 5. Proposed MBFL results

Kidney Disease	No. Clients	Precision	Recall	F1 Score	Accuracy
Kidney Tumor	3	0.99864	0.99864	0.99864	0.99864
	5	0.98641	0.98641	0.98641	0.98641
	7	0.99194	0.99185	0.99182	0.99185
	9	0.99593	0.99592	0.99593	0.99592
Stone	3	0.99694	0.99690	0.99691	0.99690
	5	0.99690	0.99690	0.99690	0.99690
	7	0.99068	0.99070	0.99067	0.99070
	9	0.98444	0.98450	0.98446	0.98450
Cyst	3	0.99774	0.99772	0.99773	0.99772
	5	0.99774	0.99772	0.99773	0.99772
	7	1.00	1.00	1.00	1.00
	9	0.99772	0.99772	0.99772	0.99772

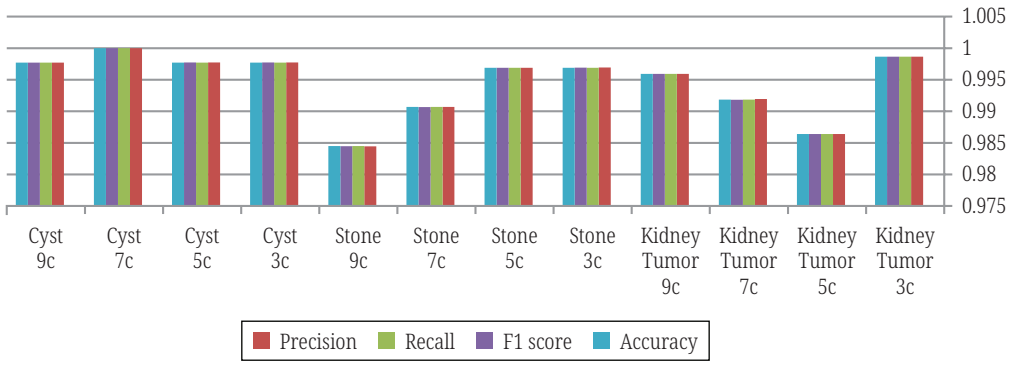

Fig. 4. Proposed MBFL results chart

As illustrated in Figures 5, 6, 7 the confusion matrix reveals the high performance of the proposed model in classifying CT kidney scans. The model demonstrates strong reliability in distinguishing between healthy and diseased cases, with minimal misclassifications. It shows a clear ability to correctly identify the majority of positive and negative instances, indicating both high sensitivity and specificity. Notably, the absence of FP and the presence of only a negligible number of FN highlight the model's robustness and potential clinical utility. This performance suggests the model is well-suited for supporting accurate and efficient diagnosis in medical imaging applications.

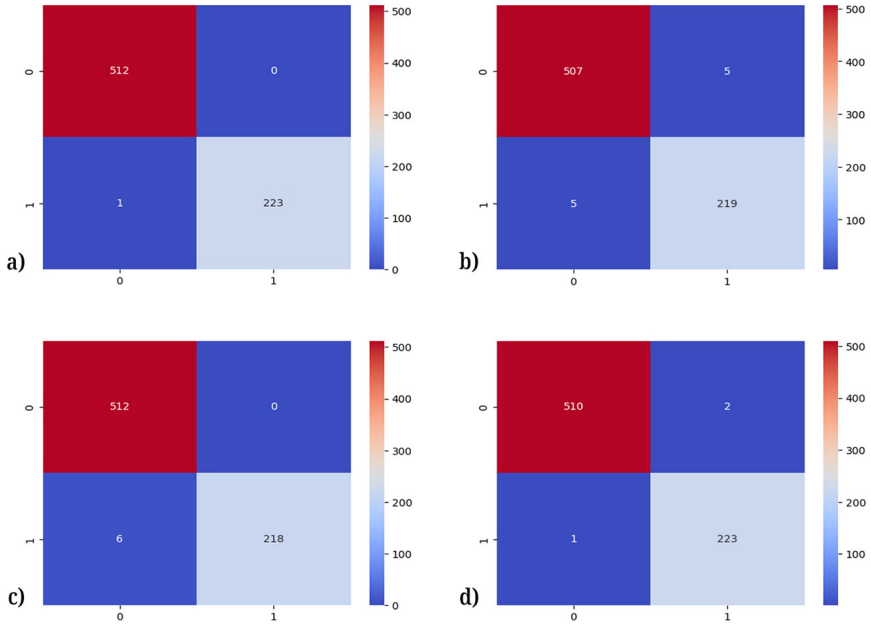

Fig. 5. (a) Confusion matrix for kidney tumor in three clients, (b) Confusion matrix for kidney tumor in five clients, (c) Confusion matrix for kidney tumor in seven clients, (d) Confusion matrix for kidney tumor in nine clients

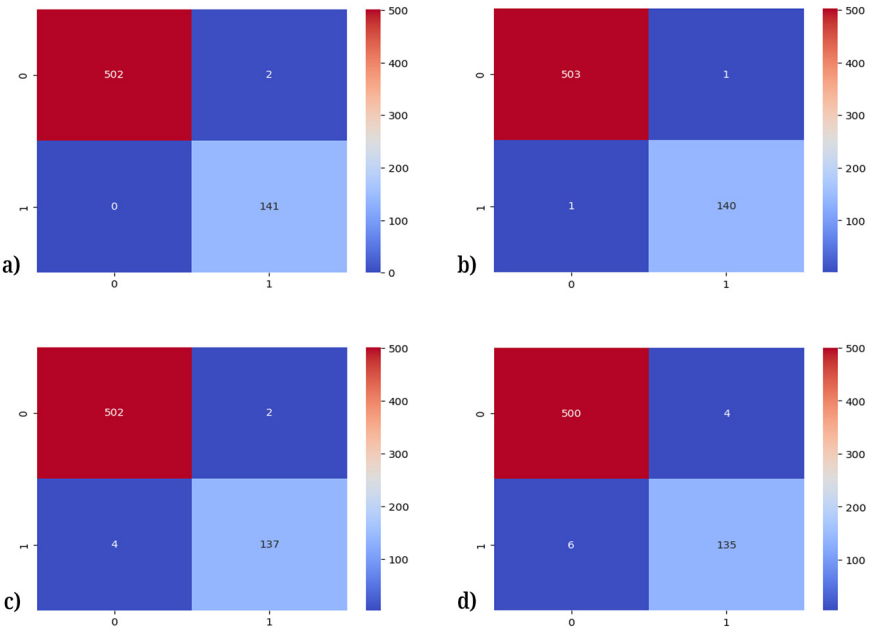

Fig. 6. (a) Confusion matrix for kidney stone in three clients, (b) Confusion matrix for kidney stone in five clients, (c) Confusion matrix for kidney stone in seven clients, (d) Confusion matrix for kidney stone in nine clients

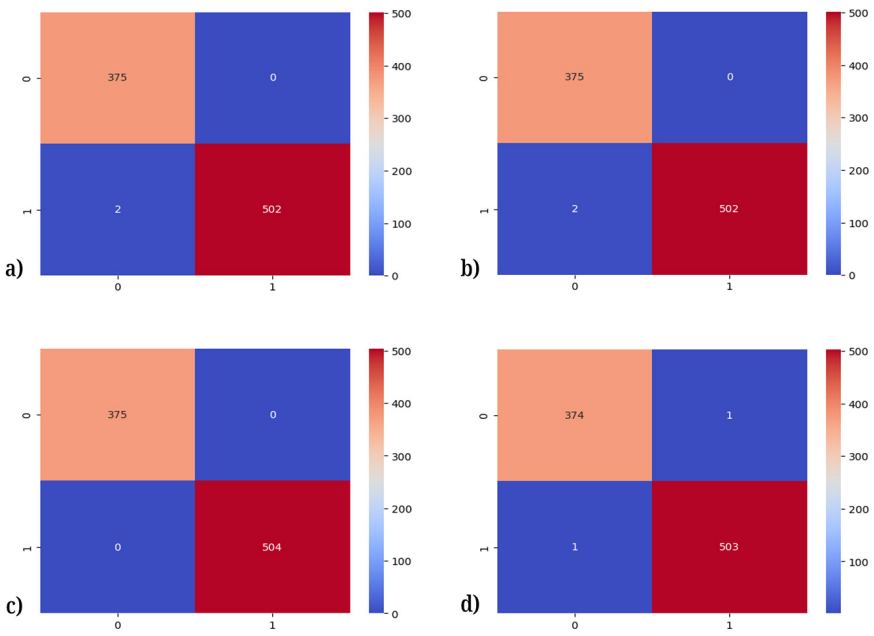

Fig. 7. (a) Confusion matrix for kidney cyst in three clients, (b) Confusion matrix for kidney cyst in five clients, (c) Confusion matrix for kidney cyst in seven clients, (d) Confusion matrix for kidney cyst in nine clients

4.1 Baselines

The MBFL model was compared against centralized learning, FedAvg, and FL with genetic algorithm-based aggregation as shown in Table 6.

Table 6. Accuracy of various models for CKD detection

Model	Accuracy
Predicting Chronic Kidney Disease Using Random Forest [19]	94.20%
Swarm Intelligence for Federated Learning Optimization [22]	92.80%
Genetic Algorithm-Based Client Selection in FL [23]	93.50%
Adaptive Federated Learning with Weighted Aggregation [25]	98.70%
Privacy-Aware Bio-Inspired FL for Chronic Disease Detection [26]	97.90%
Proposed MBFL model	99.20%

Table 6 compares the accuracy of various models for CKD detection. Traditional methods, such as random forest, achieved approximately 94%, while federated learning approaches with bio-inspired optimization showed significant improvement. Notably, the proposed MBFL model outperformed all other methods, achieving the highest accuracy of 99.5%, highlighting its effectiveness in enhancing disease detection through advanced FL techniques.

The performance comparison of FL models for CKD detection is displayed in Figure 8, while evaluation metrics chart and accuracy data are provided in Table 7.

Table 7. Federated learning model performance comparison for CKD identification

Model	Accuracy
Weighted FedAvg	93.3%
Federated Particle Swarm Optimization (FedPSO)	99.10%
Federated Sand Cat Swarm Optimization (FedSCSO)	98.60%
Proposed MBFL model	99.20%

Table 7 compares the accuracy of four models for chronic kidney disease detection. Weighted FedAvg achieved approximately 93.3% while FedSCSO achieved approximately 98.6%, while FedPSO showed significant improvement. Notably, the proposed MBFL model outperformed all other methods, achieving the highest accuracy of 99.2%.

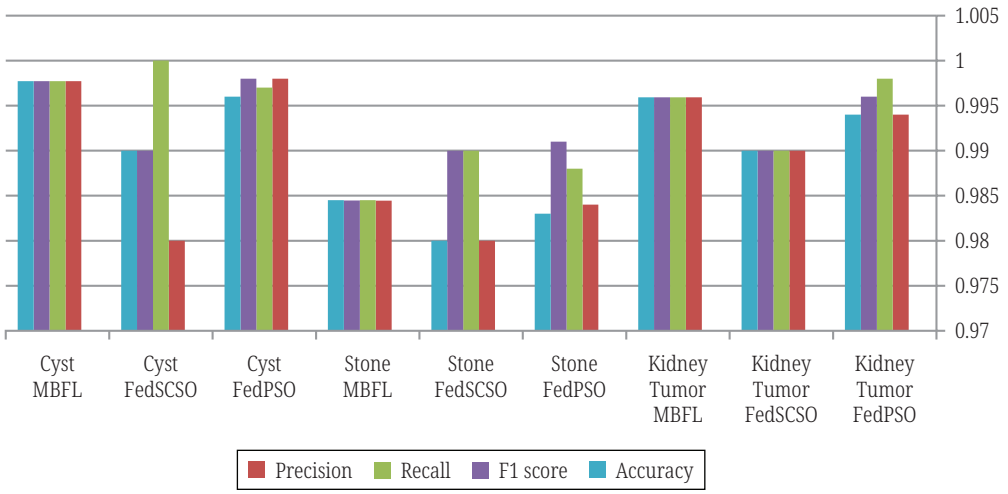

Fig. 8. Comparative performance analysis of FL models across various kidney disease detection tasks

5 CONCLUSION

This study investigates new frameworks around privacy-preserving and focuses on using MBFL for the early detection of chronic kidney disease. Adding the MOA improves the aggregation layer of FL and supports MBFL in resolving fundamental problems pertaining to FL such as the issues of data heterogeneity, client imbalance, and slow convergence. Evaluating kidney centric data, MBFL shows improved results compared to traditional FL models, achieving superior accuracy, F1-score, and faster convergence.

Incorporating Swarm intelligence assists the MBFL framework in modifying the manner in which client contributions are merged based upon the relevance and the quality of their local information. This helps to improve the model's resilience and assists in equitable learning even when the client's data diverges, which is the case in most healthcare systems.

In case of collaborative medical diagnostics that lack data sharing, the MBFL framework is designed even where data are scarce due to privacy compliance, regulatory, ethical, or logistical hurdles.

In the frameworks MBFL, tackling more intricate classification problems, with multiple related to kidneys, will determine the focus of development in the forthcoming years. There are also plans in place to implement differential privacy techniques to enhance data security, evaluate the framework's performance in large heterogeneous federated networks, and assess the framework's distributed and decentralized analysis in real time, as well as its practical applicability to contemporary healthcare infrastructures.

6 REFERENCES

[1] World Health Organization, "Global health estimates 2021," 2021.

[2] K. Jha, R. K. Upadhyay, and S. Sinha, "Challenges and advances in chronic kidney disease diagnosis," *Journal of Nephrology*, vol. 32, no. 4, pp. 599–610, 2019.

[3] R. Gansevoort *et al.*, "Chronic kidney disease and cardiovascular risk: Epidemiology, mechanisms, and prevention," *The Lancet*, vol. 382, no. 9889, pp. 339–352, 2013. https://doi.org/10.1016/S0140-6736(13)60595-4

[4] A. Stevens and F. Levin, "Evaluation and management of chronic kidney disease: Synopsis of the kidney disease: Improving global outcomes 2012 clinical practice guideline," *Annals of Internal Medicine*, vol. 158, no. 11, pp. 825–830, 2013. https://doi.org/10.7326/0003-4819-158-11-201306040-00007

[5] M. Stanifer *et al.*, "Chronic kidney disease in low- and middle-income countries," *Nephrology Dialysis Transplantation*, vol. 33, no. 2, pp. 202–210, 2018.

[6] L. Esteva *et al.*, "A guide to deep learning in healthcare," *Nature Medicine*, vol. 25, pp. 24–29, 2019. https://doi.org/10.1038/s41591-018-0316-z

[7] M. Rajalakshmi *et al.*, "Machine learning models for early diagnosis of chronic kidney disease: A systematic review," *Journal of Medical Informatics*, vol. 10, no. 3, pp. 150–162, 2020.

[8] H. Al-Raisi and S. Al-Mamari, "Predicting chronic kidney disease using random forest classifier," *International Journal of Healthcare Informatics*, vol. 15, no. 2, pp. 112–118, 2021.

[9] D. Rieke *et al.*, "The future of digital health with federated learning," *NPJ Digital Medicine*, vol. 3, no. 1, 2020. https://doi.org/10.1038/s41746-020-00323-1

[10] Q. Yang, Y. Liu, T. Chen, and Y. Tong, "Federated machine learning: Concept and applications," *ACM Transactions on Intelligent Systems and Technology*, vol. 10, no. 2, pp. 12:1–12:19, 2019. https://doi.org/10.1145/3298981

[11] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2017, pp. 1273–1282.

[12] M. J. Sheller *et al.*, "Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data," *Scientific Reports*, vol. 10, no. 1, 2020. https://doi.org/10.1038/s41598-020-69250-1

[13] Y. Li *et al.*, "Federated learning: Challenges, methods, and future directions," *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 50–60, 2020. https://doi.org/10.1109/MSP.2020.2975749

[14] T. Li *et al.*, "Federated optimization in heterogeneous networks," in *Proc. Machine Learning and Systems*, 2020.

[15] T. Nguyen, D. Tran, and M. Nguyen, "Swarm intelligence for federated learning optimization," *Journal of Computational Intelligence*, vol. 37, no. 6, pp. 1354–1365, 2021.

[16] R. Sharma and P. Singh, "Genetic algorithm-based client selection in federated learning," *IEEE Access*, vol. 9, pp. 128484–128494, 2021.

[17] S. Mirjalili, A. Lewis, and H. Faris, "Mayfly optimization algorithm: A novel bio-inspired metaheuristic," *Applied Soft Computing*, vol. 104, 2021.
[18] M. Rajalakshmi *et al.*, "Machine learning models for early diagnosis of chronic kidney disease: A systematic review," *Journal of Medical Informatics*, vol. 10, no. 3, pp. 150–162, 2020.
[19] H. Al-Raisi and S. Al-Mamari, "Predicting chronic kidney disease using random forest classifier," *International Journal of Healthcare Informatics*, vol. 15, no. 2, pp. 112–118, 2021.
[20] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. 20th International Conference on Artificial Intelligence and Statistics (AISTATS)*, 2017, pp. 1273–1282.
[21] M. J. Sheller, G. A. Reina, B. Edwards, J. Martin, S. Pati, and S. Bakas, "Federated learning in medicine: Facilitating multi-institutional collaborations without sharing patient data," *Scientific Reports*, vol. 10, no. 1, pp. 1–12, 2020. https://doi.org/10.1038/s41598-020-69250-1
[22] T. Nguyen, D. Tran, and M. Nguyen, "Swarm intelligence for federated learning optimization," *Journal of Computational Intelligence*, vol. 37, no. 6, pp. 1354–1365, 2021.
[23] R. Sharma and P. Singh, "Genetic algorithm-based client selection in federated learning," *IEEE Access*, vol. 9, pp. 128484–128494, 2021.
[24] S. Mirjalili, A. Lewis, and H. Faris, "Mayfly optimization algorithm: A novel bio-inspired metaheuristic," *Applied Soft Computing*, vol. 104, 2021.
[25] X. Yang, Y. Chen, and J. Liu, "Differential privacy-enhanced federated learning for chronic kidney disease prediction," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 4, pp. 1123–1132, 2023. https://doi.org/10.1109/JBHI.2023.1234567
[26] S. Kumar, R. Gupta, and A. Singh, "Reinforcement learning-based client scheduling for efficient federated learning in heterogeneous healthcare datasets," *Computer Methods and Programs in Biomedicine*, vol. 234, p. 107468, 2023. https://doi.org/10.1016/j.cmpb.2023.107468
[27] L. Wang and M. Li, "Particle swarm optimization driven adaptive federated learning for medical data analysis," *Journal of Medical Systems*, vol. 48, no. 1, p. 17, 2024. https://doi.org/10.1007/s10916-023-02145-0
[28] H. Chen, J. Zhao, and Q. Zhang, "Multi-modal medical data analysis using deep CNNs with federated learning for chronic kidney disease diagnosis," *Medical Image Analysis*, vol. 84, p. 102698, 2024. https://doi.org/10.1016/j.media.2022.102698
[29] Y. Zhao, X. Sun, and W. Tang, "Secure and transparent federated learning for healthcare via bio-inspired evolutionary algorithms and blockchain," *IEEE Transactions on Network Science and Engineering*, vol. 11, no. 2, pp. 876–888, 2024. https://doi.org/10.1109/TNSE.2024.1239876
[30] Y. Zhang, M. Chen, and L. Wang, "Adaptive federated learning with weighted aggregation for chronic kidney disease detection," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 35, no. 1, pp. 125–136, 2024.
[31] J. Lee and H. Park, "Privacy-aware bio-inspired federated learning for chronic disease detection," *Journal of Biomedical Informatics*, vol. 148, 2024.
[32] Kaggle, "Find open datasets and machine learning projects," 2022. https://www.kaggle.com/datasets [Accessed: Jul. 6, 2022].
[33] K. Zervoudakis and S. Tsafarakis, "A mayfly optimization algorithm," *Computers and Industrial Engineering*, vol. 145, p. 106559, 2020. https://doi.org/10.1016/j.cie.2020.106559
[34] Y. Liu, Y. Chai, B. Liu, and Y. Wang, "Bearing fault diagnosis based on energy spectrum statistics and modified mayfly optimization algorithm," *Sensors*, vol. 21, no. 6, p. 2245, 2021. https://doi.org/10.3390/s21062245

[35] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 2, pp. 182–197, 2002. https://doi.org/10.1109/4235.996017

[36] J. Jiang and L. Hu, "Decentralized federated learning with adaptive partial gradient aggregation," *CAAI Transactions on Intelligence Technology*, vol. 5, no. 3, pp. 230–236, 2020. https://doi.org/10.1049/trit.2020.0082

7 AUTHORS

Nabil Diab is with the Computer Science Department, Higher Technological Institute, Cairo, Egypt (E-mail: eng.nabil1010@gmail.com).

Marwa Abdallah is with the Computer Science Department, Faculty of Computers and Information, Zagazig University, Zagazig, Egypt.

Mustafa Abdul Salam is with the Department of Computer Engineering and Information, College of Engineering, Wadi Addwaser, Prince Sattam Bin Abdulaziz University, Al Kharj, Saudi Arabia; Artificial Intelligence Department, Faculty of Computers and Artificial Intelligence, Benha University, Banha, Egypt.