



CD-FL: Cataract Images Based Disease Detection Using Federated Learning

Arfat Ahmad Khan¹, Shtwai Alsubai², Chitapong Wechtaisong^{3,*}, Ahmad Almadhor⁴,
Natalia Kryvinska^{5,*}, Abdullah Al Hejaili⁶ and Uzma Ghulam Mohammad⁷

¹Department of Computer Science, College of Computing, Khon Kaen University, Khon Kaen, 40002, Thailand

²College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Alkharj, Saudi Arabia

³School of Telecommunication Engineering, Suranaree University of Technology, Nakhon Ratchasima, 30000, Thailand

⁴Department of Computer Engineering and Networks, College of Computer and Information Sciences, Jouf University, Sakaka, 72388, Saudi Arabia

⁵Information Systems Department, Faculty of Management, Comenius University in Bratislava, Odbojárov, Bratislava, 440, Slovakia

⁶Faculty of Computers & Information Technology, Computer Science Department, University of Tabuk, Tabuk, 71491, Saudi Arabia

⁷Department of Computer Science and Software Engineering, International Islamic University, Islamabad, 44000, Pakistan

*Corresponding Authors: Chitapong Wechtaisong. Email: chitapong@g.sut.ac.th; Natalia Kryvinska.

Email: natalia.kryvinska@uniba.sk

Received: 21 January 2023; Accepted: 13 April 2023; Published: 28 July 2023

Abstract: A cataract is one of the most significant eye problems worldwide that does not immediately impair vision and progressively worsens over time. Automatic cataract prediction based on various imaging technologies has been addressed recently, such as smartphone apps used for remote health monitoring and eye treatment. In recent years, advances in diagnosis, prediction, and clinical decision support using Artificial Intelligence (AI) in medicine and ophthalmology have been exponential. Due to privacy concerns, a lack of data makes applying artificial intelligence models in the medical field challenging. To address this issue, a federated learning framework named *CD-FL* based on a VGG16 deep neural network model is proposed in this research. The study collects data from the Ocular Disease Intelligent Recognition (ODIR) database containing 5,000 patient records. The significant features are extracted and normalized using the min-max normalization technique. In the federated learning-based technique, the VGG16 model is trained on the dataset individually after receiving model updates from two clients. Before transferring the attributes to the global model, the suggested method trains the local model. The global model subsequently improves the technique after integrating the new parameters. Every client analyses the results in three rounds to decrease the over-fitting problem. The experimental result shows the effectiveness of the federated learning-based technique on a Deep Neural Network (DNN), reaching a 95.28% accuracy while also providing privacy to the patient's data. The experiment demonstrated that the suggested federated learning model outperforms other traditional methods, achieving client 1 accuracy of 95.0% and client 2 accuracy of 96.0%.



Keywords: Privacy-preserving; cataract disease; federated learning; fundus images; healthcare; smartphone applications; machine learning

1 Introduction

A cataract is one lenticular opacity that obscures the transparent lens in a person's eye. Usually, the lens directs light toward the retina. Poor vision results from this light being blocked by the cataract and not accessing the lens. The most significant eye problem worldwide that does not immediately impair vision regressively worsens over time. However, with time, it can obstruct eyesight and even result in loss of vision in persons over 40 [1,2]. Early detection may help patients prevent painful and expensive operations and avert blindness depending on the cataract severity [2]. The National Eye Institute estimates that 24.4 million Americans already have cataracts, which will rise to 38.7 million by 2030 [3]. The WHO (World Health Organization) reports around 285 million visually impaired persons worldwide, including 246 million moderates to severe blinds and 39 million blinds [4]. According to Flaxman et al. [5], in 2020, it was predicted that 38.5 million blind people would be globally, and 237.1 million MSVI (Moderate to Severe Vision Impairment) would suffer. 13.4 million (35%) and 57.1 million (24%) would suffer from cataracts. Comparing the findings of these reports demonstrates that throughout the past decade, there has only been a marginal improvement in the eye care system and the management of vision loss.

Early diagnosis and intervention can decrease the suffering of cataract patients and stop the growth from vision impairment to blindness. Three factors make providing an automatic system to detect cataracts challenging: the wide range of cataract lesions and human visual tones; the size, shape, and position of cataracts; and the dependence on age, gender, and eye type. The most significant blindness is caused by cataracts, glaucoma [6], corneal opacity, trachoma, and diabetic retinopathy [7]. It is regarded as one of the primary causes of blindness [5]. Based on the location and region in which it manifests, cataracts can be classified into three main categories nuclear cataracts [8], cortical cataracts [9], and Posterior Sub Capsular (PSC) cataracts [10]. These three distinct types of cataracts have all been attributed to aging, diabetes, and smoking, among other frequent causes [2]. Automatic cataract prediction based on various imaging technologies has been addressed recently [11]. The automatic cataract identification and categorization systems typically use four image types: retro illumination, slit lamp, ultrasonography, or fundus images. Because technologists or patients can use the fundus camera easily, fundus images have drawn much attention in this field among different imaging techniques [12,13]. In comparison, only ophthalmologists with extensive experience should use slit-lamp cameras. Consequently, the shortage of qualified ophthalmologists prevents appropriate treatments, particularly in underdeveloped countries [14]. Therefore, an automatic cataract diagnosis method based on fundus images is crucial to streamlining early cataract screening. To identify eye diseases, image processing and machine learning algorithms have been designed [15]. A computerized ocular diagnostic system that can diagnose several severe diseases like keratoconus, glaucoma, or diabetic retinopathy has also been proposed [16–19].

Similarly, smartphone apps have been utilized for health monitoring remotely, including remote eye treatment. Smartphones are now being used to diagnose cataracts by utilizing the color and texture of the lens [17,18,20]. Even though there have been several deep learning-based automatic cataract

detection systems reported in the literature, they still have drawbacks, such as poor detection accuracy, a large number of model parameters, and high computing costs [2,21,22]. Several techniques [6,16–18] are used based on the smartphone to detect cataracts. The accuracy of detecting cataracts in these studies is vulnerable to changes in smartphone models and environmental circumstances because these studies mainly depend on smartphone camera sensor properties, ambient light, environmental factors, and distance [20].

Motivation: The healthcare industry, like other industries, generates a ton of data. The information era uses data mining, machine learning, and deep learning to transform data into knowledge for the early identification of diseases. In recent years, advances in diagnosis, prediction, and clinical decision support using AI in medicine and ophthalmology have been exponential. The utilization of digital data has also ushered in demand for privacy-preserving systems to prevent threats like adversarial attacks and to maintain patient confidentiality [23]. However, due to privacy concerns, a lack of data makes applying artificial intelligence models in the medical field challenging. The need to protect image data from malicious entities has risen exponentially as digital technology evolves. Image data varies from textual data critically, including their vast data capacity, high redundancy, and significant correlation among adjacent pixels, making it impossible for traditional text encryption techniques to work on images. Since it is so easy for hackers to steal or modify data, whether on purpose or accidentally, private data cannot be guaranteed to be protected using machine learning algorithms [24]. These problems are addressed by Federated Learning (FL), which protects data privacy [25,26]. Artificial intelligence-based systems can learn about specific data using federated learning without sacrificing privacy [27–29]. By keeping their data private, end devices can take part in learning and disseminating the prediction model while disclosing their information. Researchers are currently concentrating on utilizing federated learning with health records [30–32]. Keeping in view the above limitations, this paper makes the following contributions:

- The proposed approach initially trains the local model before sending the parameters to the global model. After then, the global model combines the updated parameters and refines the procedure while preserving the confidentiality of each client's data.
- This study proposes a technique that utilizes the ocular disease recognition dataset (fundus images) to detect and diagnose cataract disease via federated learning.
- The in-depth features are extracted from Fundus images using the VGG16 model, while the min-max normalization technique normalizes the data.
- The experiment demonstrates that the federated learning model, based on a VGG16 deep neural network model, offers client data safety and improves accuracy for cataract disease by 95.28% compared to traditional methods.

The following sections comprise the structure of the paper: The work on machine learning, deep learning, and federated learning techniques for identifying cataract disease and healthcare systems is covered in [Section 2](#). The proposed methodology is detailed in [Section 3](#), including dataset selection, data preparation, feature extraction, a federated learning framework, and model design. The results are described and discussed in [Section 4](#). The study's conclusion and suggestions for additional research are presented in [Section 5](#). [Table 1](#) provides the list of abbreviations.

Table 1: List of abbreviations

No.	Abbreviation	Description
1	AI	Artificial Intelligence
2	ODIR	Oracal Disease Intelligence Recognition
3	DNN	Deep Neural Network
4	WHO	World Health Organization
5	MVSI	Moderate to Serve Vision Impairment
6	PSC	Posterior Sub Capsular
7	BPNN	Back Propagation Neural Network
8	RF	Random Forest
9	CNN	Convolutional Neural Network
10	NB	Naïve Bayes
11	SVC	Support Vector Classifier
12	SVM	Support Vector Machine
13	LPS	Lens Partition Strategies
14	OCC	Operating Characteristic Curve
15	SRD	Standard Reference Diagnosis
16	FEA	Federated Edge Aggregator
17	DP	Distributed Protection
18	QDA	Quadratic Discriminant Analysis
19	PCA	Principal Component Analysis
20	DNN	Deep Neural Network
21	AUC	Area Under the Curve
22	GA	Gestational Age
23	BW	Birth Weight
24	VSS	Vascular Severity Score

2 Related Works

This section provides the background of cataract disease using machine and deep learning techniques. Additionally, included is background information on healthcare systems that use federated learning.

Machine Learning Techniques: Machine learning and artificial intelligence are significant for the earlier identification and detection of healthcare [33–35], including human neuropsychiatric disease [36], behavioral abnormality [37], stress illness [38] and brain disorder [39,40]. Several techniques [6,16–18] are used based on the smartphone to detect cataracts. The authors in [14] proposed an ensemble learning-based technique, SVM and Back Propagation Neural Network (BPNN), to increase diagnosis accuracy. The outcomes show a clear advantage for the ensemble classifier over the individual learning model, highlighting the potency of the suggested strategy. Regarding average accuracy for cataract identification, the ensemble classifier performs best, with a score of 84.5%.

In a study [16], researchers looked into the viability of smartphone cataract detection. Using a cutting-edge luminance-based feature extraction method, authors have suggested an accurate, transportable healthcare solution to identify cataracts. Images taken from eye models were modeled to duplicate different cataract diseases and healthy eyes. The suggested method sought to identify condensation and chromatic aberration in the eye simulation as indicators of cataracts. Using various smartphone images of 100 eye models, 50 of which suffered from cataracts and 50 of which were healthy, the authors put the suggested strategy into practice. The suggested technique was assessed in several scenarios to determine how proximity, light source, elevation, and various smartphone camera parameters affected the research result. The technique was able to identify sick eyes significantly with 96.6% accuracy, 93.4% specificity, and 93.75% sensitivity from all the images provided in the system.

A practical method for diagnosing cataract disease utilizing smartphone brightness features was proposed in [20]. In the beginning, smartphones take images of eyes and are cropped to isolate the lens. The images are then preprocessed with a median filter and a watershed transformation to eliminate extraneous background and noise. Then, a unique luminance conversion from the pixel brightness approach is provided to extract lens image features. This stage allows for correctly acquiring various cataract disease images' luminance and textural characteristics. Finally, the classification algorithm uses Support Vector Machines (SVM) to identify cataract eyes. The proposed method has a 96.6% accuracy, 93.4% specificity, and 93.75% sensitivity for diagnosing damaged eyes from all the images fed into the system. A smartphone-based android application was created utilizing the suggested approach by the authors in [41], which can be used to identify the existence of a cataract in a person's eye. The proposed approach is developed using machine learning and image processing techniques like KNN (K-Nearest Neighbor), SVM, and Naive Bayes. The proposed method has an 83.07% accuracy, an 83.18% specificity, an F-score of 82.97%, and an 82.7% recall for diagnosing cataracts.

Deep Learning Techniques: Deep learning has been utilized recently to analyze clinical data across several sectors, and it excels in tasks like segmentation and recognition. Edge-detecting filters and several mathematical methods are the foundation of the traditional approach to image classification. Different deep learning-based automatic cataract detection systems were reported in [2,21,22]. The authors in [2] suggested using a deep neural network called Cataract Net to detect cataracts automatically in fundus images. Cataract Net has a much lower computing cost and a shorter average operating time than certain pre-trained Convolutional Neural Network (CNN) algorithms. According to experimental findings, the proposed approach performs better than the most advanced cataract detection techniques, with 99.13% average accuracy.

Authors in [21] utilized convolutional neural networks to categorize cataract disease using an accessible image dataset. In this experiment, the Tensor Flow object identification framework is used to apply four distinct convolutional neural network (CNN) meta-architectures, including Xception, InceptionResnetV2, InceptionV3, and DenseNet121. The study can diagnose cataract disease at the cutting edge because of InceptionResnetV2. This framework accurately predicted cataract disease on the dataset with 1.09% training loss, 99.54% training accuracy, 6.22% validation loss, and 98.17% validation accuracy. This model's sensitivity and specificity are 96.55% and 100%, respectively. The model also significantly reduces training loss while improving accuracy. To enhance the transferability of infantile cataract detection, the paper's authors [22] created two Lens Partition Strategies (LPSs) using deep learning Faster R-CNN and Hough transform. To assess the effectiveness of LPSs, 1,643 images of multicenter slit-lamp from five ophthalmology clinics are collected. By gradually incorporating multicenter images into the training dataset, the transferability of Faster R-CNN for testing and evaluation is investigated. The Faster R-CNN can achieve a mean intersection over union for the abnormal and normal lens partitions of 0.9419 and 0.9107, respectively, with average precision

more significant than 95%. The Faster R-CNN for opacity area evaluation's sensitivity, accuracy, and specificity is enhanced by 3.29%, 5.31%, and 8.09%, respectively.

Federated Learning Techniques: FL techniques provide data privacy [25,26]. The author in [30] proposed the DEEP-FEL federated edge learning system, which permits medical devices in many institutions to train a global model cooperatively without exchanging raw data. First, creates a hierarchical ring topology to reduce the centralization of the traditional training framework. Next, formulates the ring's creation as an optimization problem that a robust heuristic algorithm can resolve. Then, to create a new global model, develops an effective parameter aggregation technique for remote medical institutions, where the amount of data transferred by N nodes is only $2/N$ times that of the conventional approach. The proposed model accuracy varies between 0.72% to 0.75%. Additionally, data security across various medical institutions is improved by incorporating artificial randomness into the edge model.

The authors of [31] used FL, a technique for jointly training deep learning algorithms without disclosing patient records, to distinguish between inter-institutional diagnostic patterns and disease epidemiology in retinopathy of prematurity (ROP). From patient examinations in seven different institutions' newborn critical care units, 5,245 retinal images are acquired. Images are classified using the bedside Clinical Grading (CL) of plus disease (no plus, pre-plus, plus), as well as the clinical diagnosis and a Standard Reference Diagnosis (RSD) that has been agreed upon by three graders wearing masks. It has been discovered that the performance of central trained models on medical labels and federated learning models equals ROC by $0.93 \hat{A} \pm 0.06$ vs. 0.95 ± 0.03 , $p = 0.0175$. To protect Smart Healthcare Systems at the edge from such privacy assaults, the authors in [32] described a Federated Edge Aggregator (FEA) system with Distributed Protection (DP) employing IoT technologies. The study uses artificial noise functions and an iterative Conventional Neural Network (CNN) model to balance model performance with privacy protection.

In summary, various machine learning and deep learning techniques have been researched and developed for predicting cataract disease. Nevertheless, they did not provide comforting proof of improved accuracy, and they did not place a high priority on data security. In this study, we develop a federated learning framework to protect data privacy that machine and deep learning models do not consider. The federated learning study that was previously discussed focused on addressing numerous healthcare challenges. The related work summary is provided in [Table 2](#).

Table 2: Summary of related work

Ref.	Techniques	Results	Limitations
[14]	SVM and BPNN	84.5% accuracy	Low performance and do not provide privacy to patient data
[22]	R-CNN	95% precision	Do not provide privacy to patient data.
[30]	DEEP-FEL	75% accuracy	Low Performance
[32]	FEA	90% accuracy	Low performance
[41]	ML techniques	83.07% accuracy	Low performance and do not provide privacy to patient data

3 Proposed Methodology

This section details the proposed methodology, including dataset, data preprocessing, feature extraction, federated learning framework, and model architecture. Fig. 1 represents the different steps of the proposed methodology. In the first step, take the ocular disease recognition dataset, preprocess the data by min-max normalization technique and extract the significant features in the second step. The preprocessed dataset is divided into two windows for two clients in the next step. The preprocessed dataset is then used to train the local model using a VGG16 deep neural network. In the last step, results from both clients are combined and updated in the global model.

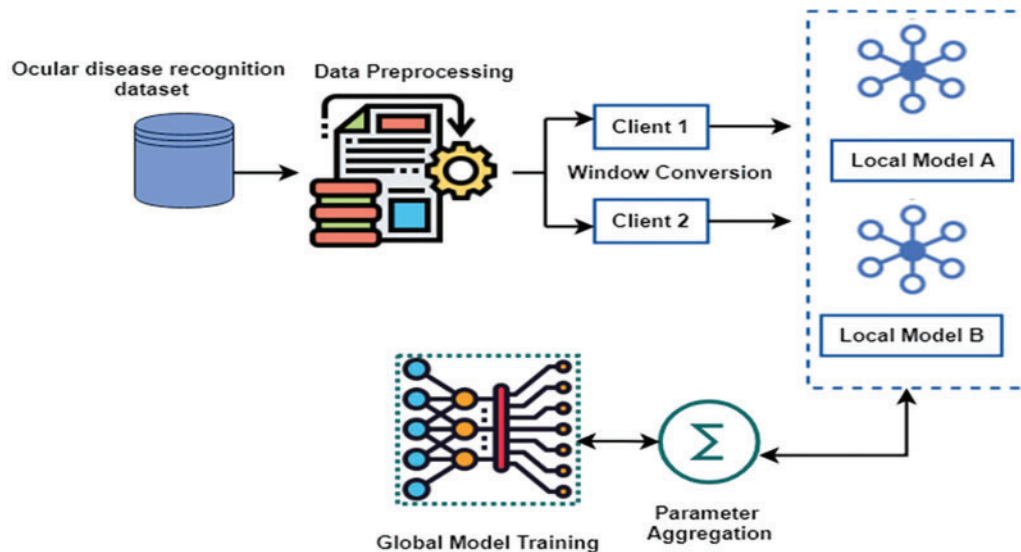


Figure 1: Overview of proposed methodology

3.1 Dataset Selection

An appropriate dataset with many samples must be used to improve validation and training for deep learning-based classification. The ocular disease recognition dataset is used in this study and is available on Kaggle. The organized ophthalmic database Ocular Disease Intelligent Recognition (ODIR) contains information on 5,000 patients, including their age, the color of their fundus images in both right and left eyes, and the keyword of their doctor's diagnostic. This dataset is intended to reflect a "real-life" collection of patient records that Shang gong Medical Technology Co., Ltd. has gathered from various hospitals and medical facilities in China. These institutions use a variety of cameras available on the market, including Kowa, Canon, and Zeiss, to acquire fundus images, producing images with different qualities. Quality assurance management educates human readers to assign labels to the annotations. Patients are divided into eight groups according to their diagnoses: normal, diabetes, hypertension, glaucoma, cataract, age-related macular degeneration, pathological myopia, and other illnesses/abnormalities. There are 6392 training data in total, including cataracts and normal eyes. The dataset samples in Fig. 2 represent whether the human eye has cataracts. Because this study uses two clients, each with its dataset, the dataset is split into two windows. The dataset is divided between training and testing, with 75% of the data used to train the models and 25% used to test the models.

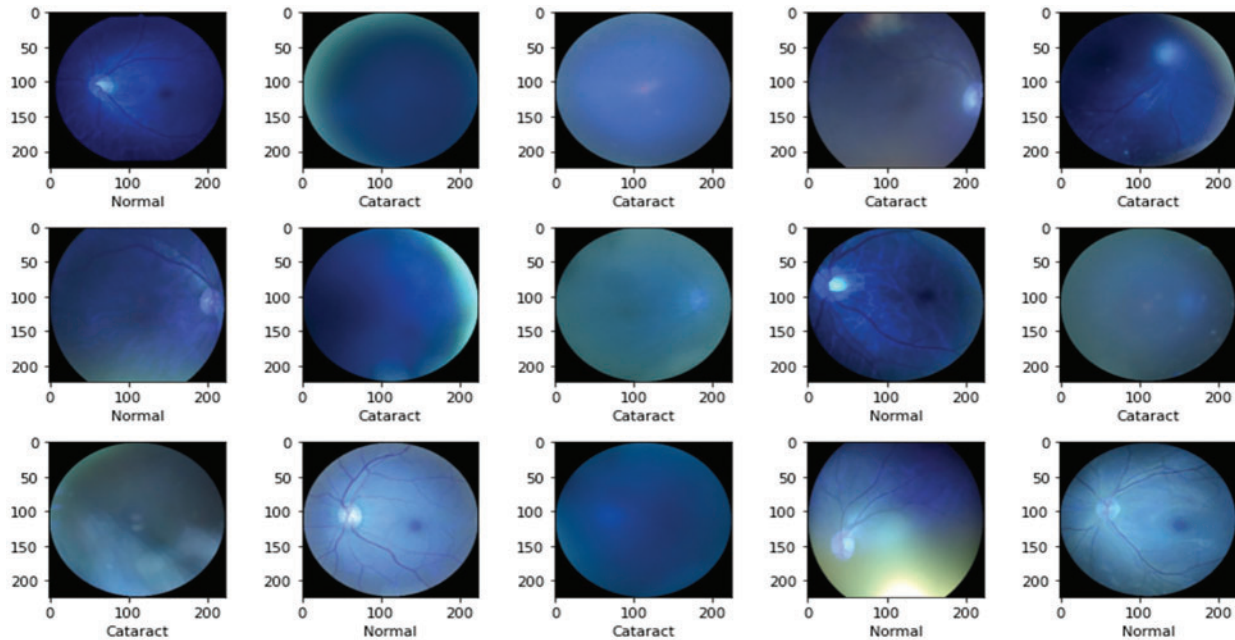


Figure 2: Samples of normal and cataract eye

3.2 Data Preprocessing and Feature Extraction

In machine learning, data preprocessing entails converting unstructured data into a format that can be utilized to create and enhance machine learning models. The initial stage in machine learning before building a model is data preprocessing. Data preparation is crucial in increasing data reliability to extract valuable information. Actual data is frequently unreliable, inaccurate (contains outliers or errors), incomplete, and devoid of particular attribute values or patterns. In this situation, data preparation is essential because it makes it easier to organize, filter, and present raw data in a format that machine learning models can use. The min-max normalization technique is used for data preprocessing in this study.

Min-Max Normalization: The low variance, the ambiguous dataset is structured, and data integrity is maintained using min-max scaling for normalizing the features. A model that relies on the magnitude of values has to scale the input attributes. Because of this, normalizing describes the discrete range of real-valued numerical properties between 0 and 1. Eq. (1) is being utilized to normalize the data.

$$Y_{norm} = \frac{Y_i - Y_{min}}{Y_{max} - Y_{min}} \quad (1)$$

After preprocessing, the dataset is split into two sections: a training dataset and a testing dataset, with 25% of the dataset used as testing data to assess the proposed model and 75% used to train the model.

Feature Extraction: The ocular disease recognition dataset is trained after data preprocessing. The VGG16 deep neural network model extracts the critical features without human oversight. Convolutional filters extract features from the training dataset following the benefit of deep learning. The VGG16 deep neural network model is used in this study to classify different ocular types and extract the finer details from an image. The extracted features are then sent to a VGG16 model with a fully connected (FC) layer in the deep neural network model.

3.3 Federated Learning Framework

A federated learning system comprises a client(s) and a server. The cloud-based federated learning server analyses key data types for the target application and trains hyper-parameters like learning rate, number of epochs, activation function, and Adam optimizer. Three crucial steps are included in the federated learning paradigm. The initialization of training is the first step. Additionally, a global model is first developed by the federated learning server.

The VGG16 deep neural network model has specified client requirements and multiple hyper-parameters. It is important to note that the federated learning server determines the model's epoch and learning rate. Training the VGG16 deep neural network model at the second level is required. Every client starts by acquiring updated information and modifying the (M^y_x) local model parameter, which depends on the global model (G^y), where y is the index for the next iteration. Each client seeks for the optimal situation to minimize the loss. Ultimately, provide the federated learning server with the new parameters regularly. The global model's integration is the third level. Once the results from various clients have been combined on the server side, send the updated parameters to each client. The global mean loss function is the primary objective of the federated learning server, which uses Eq. (2).

$$Loss(G^y) = \frac{1}{M} \sum_{x=M}^{x=1} Loss(M^y_x) \quad (2)$$

In Algorithm 1, firstly, take the ocular disease recognition dataset D_s as input and predict the cataract. In the second step, the dataset is preprocessed D_p and partitioned as the training dataset for two clients. Then extract the features f_e using the VGG16 model and normalize the data using the min-max normalization technique. The dataset is divided between training and testing, with 75% of the data used to train the models and 25% used to test the models. Update the global model and initialize the model weight w_0 . ti denotes the current round of the model, T is the total round of the local model. ci is the current client, and C is the total clients. Update the local model for each client according to the current iteration/round. Calculate the current iteration weight by the sum of the weight of the client's dataset and the current client iteration. The following model parameters, such as an epoch value, activation function, and batch size, are used to calculate the loss of the local model of each client. Update the local model by calculating the loss function $F_i(w)$. The procedure is repeated until the requisite accuracy is attained or the loss function is constantly minimized.

Algorithm 1: Proposed Federated Learning Framework for Cataract Image-based Disease Detection

- 1: **Input:** D_s (Ocular diseases recognition dataset)
 - 2: **Output:** Cataract prediction
 - 3: D_p (Data preprocessing)
 - 4: f_e (Feature extraction)
 - 5: $Y_{norm} = Y_i - Y_{min} / Y_{max} - Y_{min}$ (data Normalization)
 - 6: $x_{train}, x_{test}, y_{train}, y_{test} \leftarrow train-test\ split\ dataset$ (dataset splitting)
 - 7: **function** global model upgrade
 - 8: **weight initialization** w_0
 - 9: **for** ($t_i = 1$) to T **do**
 - 10: **for** ($ci = 1$) to C **do**
 - 11: $w_{ti+1}^{ci} = local\ model\ upgrade(c_i, w_{ti})$
 - 12: $w_{ti+1} = \sum_{ci=1}^C w_{D_s}^{ci} * w_{ci+1}^{ti}$
 - 13: **function** local model upgrade (ci, w_{ti})
 - 14: **for** ($Epoch = 1$) to E_p **do**
-

(Continued)

Algorithm 1 (continued)

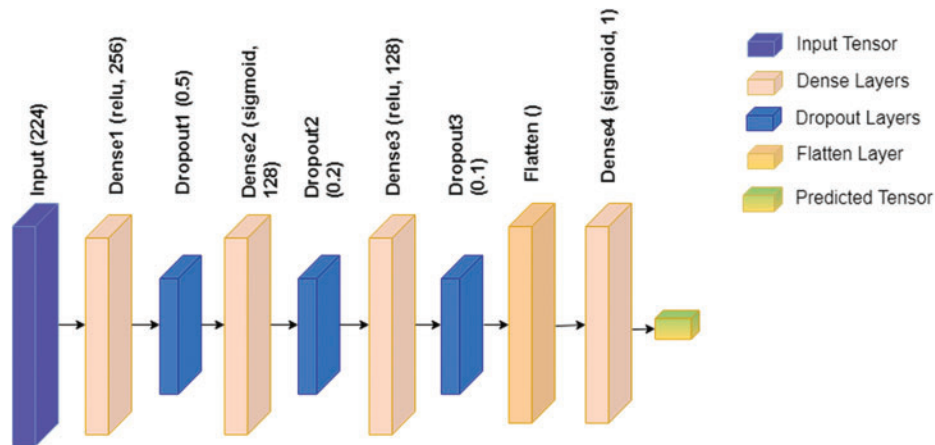
```

15:   for ( $B_s \in \text{BatchSize}$ ) do
16:      $w = w - \nabla F_i(w)$ 
17:   return  $w$ 

```

3.4 Model Architecture

The VGG16 deep neural network model is used for cataract disease prediction. When simulating the structure of neural networks, the neurons and the number of layers are essential. The number of neurons utilized as input and output in a deep neural network model depends on the training data size. A deep convolution neural network model with much success in computer vision is the VGG16. The architecture of the VGG16 model is provided in Fig. 3. In general, the VGG16 model has three layers: an input layer, several hidden layers (such as dropout, dense, flatten, etc.), and the output layer. The sequential VGG16 model used in this study has a single input layer. The input layer has a shape of 224 and uses the relu activation function. The next layer is the hidden layer, consisting of four dense layers and three dropout layers. The three dropout layers are employed to prevent the overfitting of the model. The values of the dropout layers are 0.5, 0.2, and 0.1, respectively. The four dense layers comprise 256 and 128, and 1 unit and the activation functions are the relu and sigmoid. The flattened layer, typically used in the transition from the convolution layer to the fully connected layer, is the next layer to reduce the multidimensional input to one dimension. The output layer comes next; it is the fully connected layer utilized for binary classification problems. The VGG16 model uses binary cross-entropy and Adam as an optimizer to calculate and reduce the loss. To address the binary classification problem, every dense layer uses relu and sigmoid activation functions along with a fully connected layer.

**Figure 3:** Proposed architecture of VGG16 model**4 Experimental Results and Discussion**

This section presents the experimental results and an evaluation of the proposed methodology. The study investigates the effects of broader approach technique parameters. The experiment is conducted on the ocular disease recognition dataset containing the 6392 fundus images of both left and right eyes. One server and two clients are involved in the experiment. The two clients use the ocular disease recognition dataset to train the VGG16 deep neural network model.

The ocular disease recognition image dataset trains the VGG16 deep neural network model and constructs the random weight. The section regarding model architecture provides the basics of the VGG16 deep neural network model. First, the study acquired 6392 fundus images from the ophthalmic database Ocular Disease Intelligent Recognition (ODIR) and extracted the specific features that may accurately identify cataracts. The dataset is divided into two portions: 25% is used for testing the model, and 75% is utilized for training the model. All values are normalized using the min-max normalization to fit inside the range [0, 1]. The label is translated into a machine-readable format using the label encoder approach. The experiment utilizes a variety of evaluation criteria, including recall, f1-score, precision, and accuracy, to evaluate the federated learning-based approach's ability for prediction. The results for each client are verified three times for loss prevention. The server end aggregates the client's results. The equations of the evaluation metrics are given below. The accuracy of the model is evaluated by using Eq. (3). The precision of the model is assessed by using Eq. (4). The recall and f1-score of the model are evaluated by Eqs. (5) and (6).

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (3)$$

$$Precision = \frac{TP}{TP + FP} \quad (4)$$

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

$$F1 = \frac{2 * TP}{2 * (TP + FP + FN)} \quad (6)$$

4.1 Server-Based Training Using Log Data

The VGG16 deep neural network model is trained using data logged on servers. The server decides which client or node is used at the initial level of the model training stage and accumulates any modifications received. Before the training, all individually identifying information is eliminated from the logs, and they are made anonymous. There are three rounds after the federated learning begins. After initializing a few server-side parameters, we set the round count to three and evaluated the tests three times. The federated learning model consists of two phases for each round: the fit round and the evaluate round. During the fit-round, the clients transmit the learning outcomes to the server. Both clients send the test results to the server when the findings are combined during the evaluation stage. The VGG16 deep neural network model has the highest accuracy, rating 95.28%, according to the server's analysis of the information from N clients.

4.2 Client 1 Federated Training and Testing

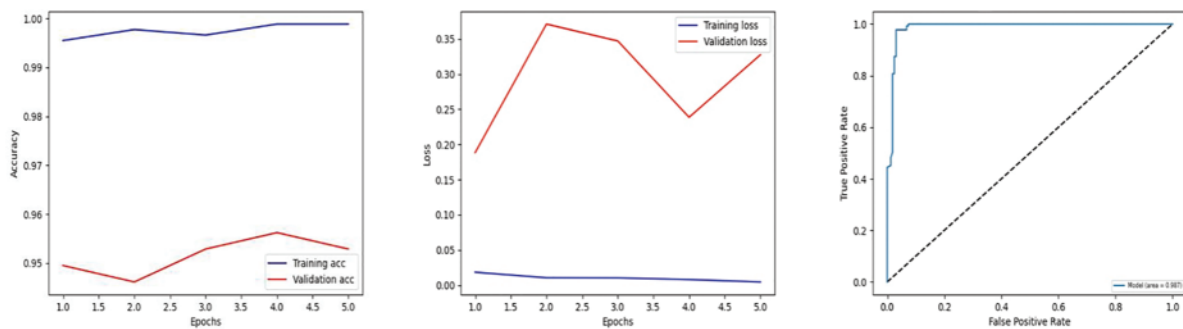
Client 1 conducts the experiment using a sequential VGG16 deep neural network model with four dense, three dropouts, and one flatten layer as hidden layers. The input shape of the input layer is 224 with the relu activation function. Adam optimizer calculates and reduces the loss; the VGG16 model uses binary cross-entropy. Three rounds of experiments are executed by client 1, and the results are provided in two rounds: the evaluation and the fit rounds. The VGG16 deep neural network model is employed in this experiment, which includes three test rounds. The experiment evaluates evaluation metrics in three rounds: recall, accuracy, precision, and f1-score. The experimental results for client 1 are shown in Table 3. The VGG16 model is employed in the study's round 1 analysis, and the results are successful in terms of accuracy, recall, precision, and F1-score. The VGG16 model provides 94.0% accuracy, 91.0% precision, 96.0% recall, and 93.0% F1-score in the first round. Analyze the results to

avoid the model overfitting once more. The accuracy, precision, recall, and F1-score for the VGG16 model in round 2 are 95.0%, 92.0%, 99.0%, and 95.0%, respectively. The VGG16 model achieves an accuracy of 95.0%, a precision of 91.0%, a recall of 100.0%, and an F1-score of 95.0% in round 3. Round 3 yields the best result for client 1.

Table 3: Proposed model result of client 1

Rounds	Accuracy	Precision	Recall	F1-score
Round 1	94.0%	91.0%	96.0%	93.0%
Round 2	95.0%	92.0%	99.0%	95.0%
Round 3	95.0%	91.0%	100.0%	95.0%

The highest results are illustrated in Fig. 4. The graph in Fig. 4a shows the training and validation accuracy, with round 3 of client 1 attaining higher validation accuracy than training accuracy. A blue line shows the training accuracy curve, while an orange line shows the validation accuracy curve. At the 1st epoch, the training accuracy is 0.99%; after various fluctuations between falls and gains, it reached about 1.00% accuracy at the 5th epoch. At the 1st epoch, validation accuracy is 0.95%; it then fluctuates between drops and gains before returning to that value at the 5th epoch. The training and validation loss is depicted on the graph in Fig. 4b. During the training phase, the loss fluctuates at each epoch. The blue line represents the training loss curve, while the orange line represents the validation loss in the loss curve. Training loss initiated from 0.02% at the 1st epoch and decreased to 0.01 at the 5th. Validation loss initiated from 0.19% at 1st epoch and increased to 0.32% at the 5th epoch. The Receiver Operating Characteristic (ROC) is depicted in the graph in Fig. 4c. The experiment uses two classes, and the proposed model performs better on the used dataset with ROC scores of 0.987%. The ROC curves close to the top-left corner demonstrate better performance. Fig. 6a, which graphically illustrates the proposed technique's confusion matrix for client 1, gives a general concept of how a classification algorithm functions. Because it has more continuous, better true positive and negative results and fewer false positive and negative values, the proposed technique performs better. The proposed model predicted 100% of cataract cases accurately for client 1.



(a) Round 3 accuracy of validation and training (b) Round 3 loss of validation and training (c) Round 3 score of ROC

Figure 4: Representation of client 1 highest-scoring results

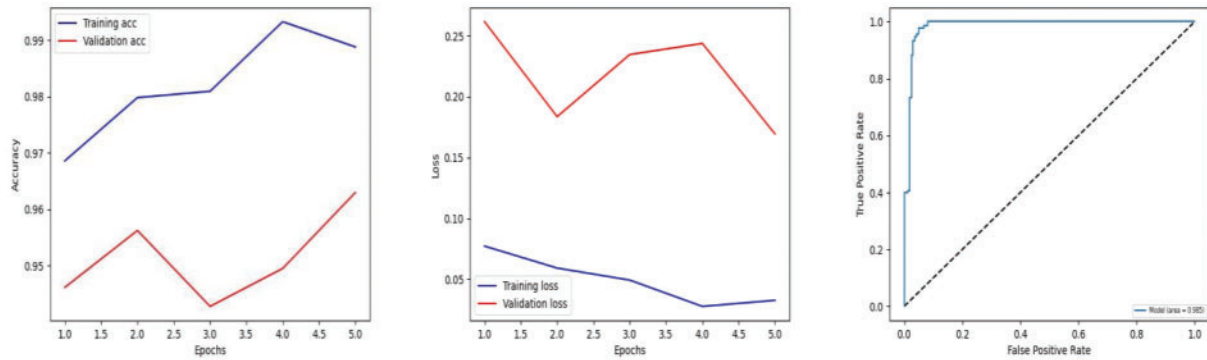
4.3 Client 2 Federated Training and Testing

Client 2 conducts the experiment using a sequential VGG16 deep neural network model with four dense, three dropouts, and one flatten layer as hidden layers. The input shape of the input layer is 224 with the relu activation function. Adam optimizer calculates and reduces the loss; the VGG16 model uses binary cross-entropy. Three rounds of experiments are executed by client 2, and the results are provided in two rounds: the evaluation and the fit rounds. The VGG16 deep neural network model is employed in this experiment, which includes three test rounds. The experiment evaluates evaluation metrics in three rounds: recall, accuracy, precision, and f1-score. The experimental results for client 2 are shown in Table 4. The VGG16 model is employed in the study's round 1 analysis, and the results are successful in accuracy, recall, precision, and F1-score. The VGG16 model provides 94.0% accuracy, 90.0% precision, 97.0% recall, and 93.0% F1-score in the first round. Analyze the results to avoid the model overfitting once more. The accuracy, precision, recall, and F1-score for the VGG16 model in round 2 are 96.0%, 94.0%, 98.0%, and 96.0%, respectively. The VGG16 model achieves an accuracy of 95.0%, a precision of 90.0%, a recall of 100.0%, and an F1-score of 95.0% in round 3. Round 2 yields the best result for client 2.

Table 4: Proposed model result of client 2

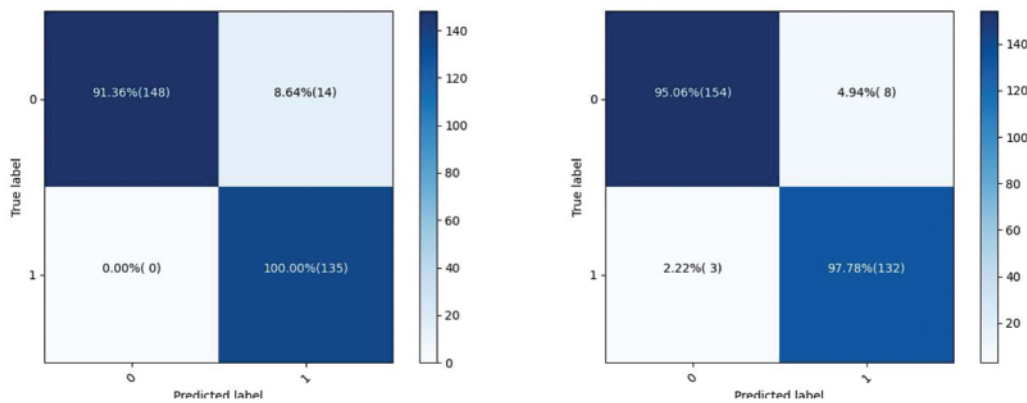
Rounds	Accuracy	Precision	Recall	F1-score
Round 1	94.0%	90.0%	97.0%	93.0%
Round 2	96.0%	94.0%	98.0%	96.0%
Round 3	95.0%	90.0%	100.0%	95.0%

The highest results are illustrated in Fig. 5. The graph in Fig. 5a shows the training and validation accuracy, with round 2 of client 2 attaining higher validation accuracy than training accuracy. A blue line shows the training accuracy curve, while an orange line shows the validation accuracy curve. At the 1st epoch, the training accuracy is 0.97%; after various fluctuations between falls and gains, it reached about 0.99% accuracy at the 5th epoch. At the 1st epoch, validation accuracy is 0.93%; however, it fluctuates between drops and gains until reaching 0.96% accuracy at the 5th epoch. The training and validation loss is depicted on the graph in Fig. 5b. During the training phase, the loss fluctuates at each epoch. The blue line represents the training loss curve, while the orange line represents the validation loss in the loss curve. Training loss initiated from 0.75% at 1st epoch and decreased to 0.03 at the 5th epoch. Validation loss initiated from 0.26% at 1st epoch and decreased to 0.17% at the 5th epoch. The proposed model outperforms as training and validation loss is decreased at the final epoch, enhancing the model's performance. The Receiver Operating Characteristic (ROC) is depicted in the graph in Fig. 5c. The experiment uses two classes, and the proposed model performs better on the used dataset with ROC scores of 0.985%. The ROC curves close to the top-left corner demonstrate better performance. Fig. 6b, which graphically illustrates the proposed technique's confusion matrix for client 2, gives a general concept of how a classification algorithm functions. Because it has more continuous, better true positive and negative results and fewer false positive and negative values, the proposed technique performs better. The proposed model diagnosed 97.78% of cataract cases accurately for client 2.



(a) Round 2 accuracy of validation and training (b) Round 2 loss of validation and training (c) Round 2 score of ROC

Figure 5: Visualization of client 2 highest-scoring results



(a) Confusion matrix of client 1

(b) Confusion matrix of client 2

Figure 6: Confusion matrices for both clients using the proposed method

The comparative analysis of the suggested approach with existing techniques [14,30,32,41] is provided in Table 5. The comparison is provided regarding the accuracy, precision, recall, and F1-score. The proposed approach outperforms as associated to the existing techniques.

Table 5: Comparison with existing techniques

References	Techniques	Accuracy	Precision	Recall	F1-score
[14]	SVM and BPNN	84.5%	NA	NA	NA
[30]	DEEP-FEL	75%	NA	NA	NA
[32]	FEA	90%	NA	NA	NA
[41]	ML techniques	83.07%	NA	82.7%	82.97%
Proposed approach	CD-FL (VGG-16)	95.0%	91.0%	100.0%	95.0%

4.4 Research Limitations

The sustainability of healthcare systems relies on using user data to train ML models while raising severe privacy and security concerns. This research proposes an FL-based VGG16 model to recognize and classify Cataract image-based disease to address this issue. Since FL prioritizes protecting patient data, no direct data exchange occurs in this collaborative process. Although FL permits more collaborative ML (with privacy protection at its core), it also presents many problems, including those mentioned here. Communication is a significant hurdle in FL networks, as the data created by each device is kept locally. Designing models that allow effective communication while limiting the number of cycles to train a model using device-generated input is crucial. Although FL's privacy features are a big positive, they also prevent data analysts from seeing user data in its unprocessed form. To detect missing data, eliminate extraneous information, and identify the data points that the system should be trained on, they cannot clean up the data resulting in the system's poor performance. Additionally, it should incrementally communicate discrete model upgrades rather than providing the complete data set during training.

5 Conclusion and Future Directions

A cataract is one of the most severe ocular diseases affecting people worldwide, which gradually deteriorates over time without instantly impairing eyesight. Cataract disease is the most severe disease for which a patient needs special care, including protecting the confidentiality of their medical records. However, using artificial intelligence models in the medical industry is challenging due to privacy concerns. Federated learning (FL), which protects data privacy, is proposed in this study to deal with this issue using a VGG16 deep neural network model which detects and diagnoses cataract disease. The study collects data from the Ocular Disease Intelligent Recognition (ODIR) database containing 5,000 patient records. The significant features are extracted and normalized using the min-max normalization technique. Because this study uses two clients, each with its dataset, the dataset is split into two windows to construct the training model. The VGG16 model is trained on two clients, and the aggregate server end accuracy is 95.28%. Each client reviewed the results three times to reduce the over-fitting element. The best result of round 3 of client 1 was obtained with a 95.0% accuracy, and client 2 of round 2 obtained the best result with a 96.0% accuracy. The VGG16 model's ROC curve average of 0.985% demonstrates the proposed method's excellent performance on the tested dataset. The research results demonstrated that federated learning effectively protects client data privacy. According to a system efficiency study, side training times and storage costs favor medical devices with constrained resources. Furthermore, it enables concurrent model training on each piece of data while storing the patient data in several locations. Compared to training on a personal processor, researchers may train more quickly while requiring less computing or storage capacity from each location. In the future, we plan to investigate this trend more thoroughly by training more systems using various smartphone device combinations and expanding our research by applying new deep learning algorithms with multiple datasets and employing different statistical tests, such as Wilcoxon and ANOVA, to ensure the quality of the proposed algorithm.

Acknowledgement: The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia, for funding this research work through Project Number 959.

Funding Statement: The authors extend their appreciation to the Deputyship for Research & Innovation, Ministry of Education in Saudi Arabia, for funding this research work through Project Number 959.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] J. Rana and S. M. Galib, "Cataract detection using smartphone," in *2017 3rd Int. Conf. on Electrical Information and Communication Technology (EICT)*, IEEE, Khulna, Bangladesh, pp. 1–4, 2017.
- [2] M. S. Junayed, M. B. Islam, A. Sadeghzadeh and S. Rahman, "Cataractnet: An automated cataract detection system using deep learning for fundus images," *IEEE Access*, vol. 9, pp. 128799–128808, 2021.
- [3] N. NIH, "Eye disease statistics," *National Eye Institute*, 2019. [Online]. Available: <https://www.nei.nih.gov/>
- [4] D. Allen and A. Vasavada, "Cataract and surgery for cataract," *Bmj*, vol. 333, no. 7559, pp. 128–132, 2006.
- [5] S. R. Flaxman, R. R. Bourne, S. Resnikoff, P. Ackland, T. Braithwaite *et al.*, "Global causes of blindness and distance vision impairment 1990–2020: A systematic review and meta-analysis," *The Lancet Global Health*, vol. 5, no. 12, pp. e1221–e1234, 2017.
- [6] S. Hu, X. Wang, H. Wu, X. Luan, P. Qi *et al.*, "Unified diagnosis framework for automated nuclear cataract grading based on smartphone slit-lamp images," *IEEE Access*, vol. 8, pp. 174169–174178, 2020.
- [7] L. Qiao, Y. Zhu and H. Zhou, "Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms," *IEEE Access*, vol. 8, pp. 104292–104302, 2020.
- [8] Y. C. Liu, M. Wilkins, T. Kim, B. Malyugin and J. S. Mehta, "Cataracts," *The Lancet*, vol. 390, no. 10094, pp. 600–612, 2017.
- [9] C. Niya and T. Jayakumar, "Analysis of different automatic cataract detection and classification methods," in *2015 IEEE Int. Advance Computing Conf. (IACC)*, IEEE, Bangalore, India, pp. 696–700, 2015.
- [10] B. Raju, N. Raju, J. D. Akkara and A. Pathengay, "Do it yourself smartphone fundus camera–diyretcam," *Indian Journal of Ophthalmology*, vol. 64, no. 9, pp. 663, 2016.
- [11] A. R. Javed, F. Shahzad, S. ur Rehman, Y. B. Zikria, I. Razzak *et al.*, "Future smart cities requirements, emerging technologies, applications, challenges, and future aspects," *Cities*, vol. 129, pp. 103794, 2022.
- [12] W. Fan, R. Shen, Q. Zhang, J. J. Yang and J. Li, "Principal component analysis based cataract grading and classification," in *2015 17th Int. Conf. on E-Health Networking, Application & Services (HealthCom)*, IEEE, Boston, MA, pp. 459–462, 2015.
- [13] C. Costanian, M. J. Aubin, R. Buhrmann and E. E. Freeman, "Interaction between postmenopausal hormone therapy and diabetes on cataract," *Menopause*, vol. 27, no. 3, pp. 263, 2020.
- [14] J. J. Yang, J. Li, R. Shen, Y. Zeng, J. He *et al.*, "Exploiting ensemble learning for automatic cataract detection and grading," *Computer Methods and Programs in Biomedicine*, vol. 124, pp. 45–57, 2016.
- [15] F. Grassmann, J. Mengelkamp, C. Brandl, S. Harsch, M. E. Zimmermann *et al.*, "A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography," *Ophthalmology*, vol. 125, no. 9, pp. 1410–1420, 2018.
- [16] B. Askarian, J. W. Chong and F. Tabei, "Diagnostic tool for eye disease detection using smartphone," *IEEE Journal of Translational Engineering in Health and Medical*, uS Patent 10,952,604, vol. 9, 3800110, 2021.
- [17] B. Askarian, F. Tabei, G. A. Tipton and J. W. Chong, "Novel keratoconus detection method using smartphone," in *2019 IEEE Healthcare Innovations and Point of Care Technologies (HI-POCT)*, IEEE, Bethesda, MD, USA, pp. 60–62, 2019.

- [18] B. Askarian, F. Tabei, A. Askarian and J. W. Chong, "An affordable and easy-to-use diagnostic method for keratoconus detection using a smartphone," in *Medical Imaging 2018: Computer-Aided Diagnosis*, vol. 10575. SPIE, pp. 238–243, 2018.
- [19] J. Tekli, "An overview of cluster-based image search result organization: Background, techniques, and ongoing challenges," *Knowledge and Information Systems*, pp. 1–54, 2022.
- [20] B. Askarian, P. Ho and J. W. Chong, "Detecting cataract using smartphones," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 9, pp. 1–10, 2021.
- [21] M. K. Hasan, T. Tanha, M. R. Amin, O. Faruk, M. M. Khan *et al.*, "Cataract disease detection by using transfer learning-based intelligent methods," *Computational and Mathematical Methods in Medicine*, vol. 2021, 2021.
- [22] J. Jiang, S. Lei, M. Zhu, R. Li, J. Yue *et al.*, "Improving the generalizability of infantile cataracts detection via deep learning-based lens partition strategy and multicenter datasets," *Frontiers in Medicine*, vol. 8, pp. 664023, 2021.
- [23] J. S. Lim, M. Hong, W. S. Lam, Z. Zhang, Z. L. Teo *et al.*, "Novel technical and privacy-preserving technology for artificial intelligence in ophthalmology," *Current Opinion in Ophthalmology*, vol. 33, no. 3, pp. 174–187, 2022.
- [24] F. A. KhoKhar, J. H. Shah, M. A. Khan, M. Sharif, U. Tariq *et al.*, "A review on federated learning towards image processing," *Computers and Electrical Engineering*, vol. 99, pp. 107818, 2022.
- [25] A. R. Javed, M. A. Hassan, F. Shahzad, W. Ahmed, S. Singh *et al.*, "Integration of blockchain technology and federated learning in vehicular (IOT) networks: A comprehensive survey," *Sensors*, vol. 22, no. 12, pp. 4394, 2022.
- [26] A. Rehman, I. Razzak and G. Xu, "Federated learning for privacy preservation of healthcare data from smartphone-based side-channel attacks," *IEEE Journal of Biomedical and Health Informatics*, 2022.
- [27] S. AbdulRahman, H. Tout, A. Mourad and C. Talhi, "Fedmccs: Multicriteria client selection model for optimal IOT federated learning," *IEEE Internet of Things Journal*, vol. 8, no. 6, pp. 4723–4735, 2020.
- [28] O. A. Wahab, A. Mourad, H. Otok and T. Taleb, "Federated machine learning: Survey, multi-level classification, desirable criteria and future directions in communication and networking systems," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 1342–1397, 2021.
- [29] S. AbdulRahman, H. Tout, H. Ould-Slimane, A. Mourad, C. Talhi *et al.*, "A survey on federated learning: The journey from centralized to distributed on-site learning and beyond," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5476–5497, 2020.
- [30] Z. Lian, Q. Yang, W. Wang, Q. Zeng, M. Alazab *et al.*, "Deep-fel: Decentralized, efficient and privacy-enhanced federated edge learning for healthcare cyber physical systems," *IEEE Transactions on Network Science and Engineering*, 2022.
- [31] A. H. Hanif, C. Lu, K. Chang, P. Singh, A. S. Coyner *et al.*, "Federated learning for collaborative clinical diagnosis and disease epidemiology in retinopathy of prematurity," *Investigative Ophthalmology & Visual Science*, vol. 63, no. 7, pp. 2329–2329, 2022.
- [32] M. Akter, N. Moustafa and T. Lynar, "Edge intelligence-based privacy protection framework for iotbased smart healthcare systems," in *IEEE INFOCOM 2022-IEEE Conf. on Computer Communications Workshops (INFOCOM WKSHPS)*, IEEE, New York, NY, USA, pp. 1–8, 2022.
- [33] P. Monga, M. Sharma and S. K. Sharma, "A comprehensive meta-analysis of emerging swarm intelligent computing techniques and their research trend," *Journal of King Saud University-Computer and Information Sciences*, 2021.
- [34] U. G. Mohammad, S. Imtiaz, M. Shakya, A. Almadhor and F. Anwar, "An optimized feature selection method using ensemble classifiers in software defect prediction for healthcare systems," *Wireless Communications and Mobile Computing*, vol. 2022, pp. 14, 2022.
- [35] M. Sharma, "Design of brain-computer interface-based classification model for mining mental state of COVID-19 afflicted marinerâs," *International Maritime Health*, vol. 71, no. 4, pp. 298–300, 2020.

- [36] M. Sharma, "Research and google trend for human neuropsychiatric disorders and machine learning: A brief report," *Psychiatria Danubina*, vol. 33, no. br. 3, pp. 354–357, 2021.
- [37] P. Monga, M. Sharma and S. K. Sharma, "Performance analysis of machine learning and soft computing techniques in diagnosis of behavioral disorders," in *Electronic Systems and Intelligent Computing: Proc. of ESIC 2021*, Springer, Cham, pp. 85–99, 2022.
- [38] S. Sharma, G. Singh and M. Sharma, "A comprehensive review and analysis of supervised-learning and soft computing techniques for stress diagnosis in humans," *Computers in Biology and Medicine*, vol. 134, pp. 104450, 2021.
- [39] R. Gautam and M. Sharma, "Prevalence and diagnosis of neurological disorders using different deep learning techniques: A meta-analysis," *Journal of Medical Systems*, vol. 44, no. 2, pp. 49, 2020.
- [40] S. Alsubai, H. U. Khan, A. Alqahtani, M. Sha, S. Abbas *et al.*, "Ensemble deep learning for brain tumor detection," *Frontiers in Computational Neuroscience*, vol. 16, pp. 14, 2022.
- [41] V. Agarwal, V. Gupta, V. M. Vashisht, K. Sharma and N. Sharma, "Mobile application based cataract detection system," in *2019 3rd Int. Conf. on Trends in Electronics and Informatics (ICOEI)*, IEEE, Tirunelveli, India, pp. 780–787, 2019.