Federated Machine Learning for Detection of Skin Diseases and Enhancement of Internet of Medical Things (IoMT) Security

Md. Nazmul Hossen, *Graduate Student Member, IEEE*, Vijayakumari Panneerselvam, Deepika Koundal, Kawsar Ahmed, *Graduate Student Member, IEEE*, Francis M. Bui, *Member, IEEE*, and Sobhy M. Ibrahim

***Abstract*—Human skin disease, the most infectious dermatological ailment globally, is initially diagnosed by sight. Some clinical screening and dermoscopic analysis of skin biopsies and scrapings for accurate classification are medically compulsory. Classification of skin diseases using medical images is more challenging because of the complex formation and variant colors of the disease and data security concerns. Both the Convolution Neural Network (CNN) for classification and a federated learning approach for data privacy preservation show significant performance in the realm of medical imaging fields. In this paper, a custom image dataset was prepared with four classes of skin disease, a CNN model was suggested and compared with several benchmark CNN algorithms, and an experiment was carried out to ensure data privacy using a federated learning approach. An image augmentation strategy was followed to enlarge the dataset and make the model more general. The proposed model achieved a precision of 86%, 43%, and 60%, and a recall of 67%, 60%, and 60% for acne, eczema, and psoriasis. In the federated learning approach, after distributing the dataset among 1000, 1500, 2000, and 2500 clients, the model showed an average accuracy of 81.21%, 86.57%, 91.15%, and 94.15%. The CNN-based skin disease classification merged with the federated learning approach is a breathtaking concept to classify human skin diseases while ensuring data security.**

***Index Terms*—Benchmark algorithms, convolution neural network, federated learning framework, IoMT security, medical imaging, skin disease classification.**

Manuscript received 8 November 2021; revised 24 January 2022; accepted 30 January 2022. Date of publication 8 February 2022; date of current version 6 February 2023. This work was supported by Researchers Supporting under Project number RSP-2021/100 King Saud University, Riyadh, Saudi Arabia, and in part by the Natural Sciences and Engineering Research Council of Canada (NSERC). *(Corresponding author: Kawsar Ahmed.)*

Md. Nazmul Hossen is with the Department of Information and Communication Technology, Mawlana Bhashani Science and Technology University, Tangail 1902, Bangladesh (e-mail: nazmul.ict92@gmail.com).

Vijayakumari Panneerselvam is with the Department of Applied Electronics, Institute of ECE, Saveetha School of Engineering, SIMATS, Chennai 602105, India (e-mail: vijayakumarip.sse@saveetha.com).

Deepika Koundal is with the Department of Systemics and School of Computer Science, University of Petroleum and Energy Studies, Dehradun 248006, India (e-mail: dkoundal@ddn.upes.ac.in).

Kawsar Ahmed is with the Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada, and also with the Group of Bio-Photomatiχ, Department of ICT, MBSTU, Tangail 1902, Bangladesh (e-mail: k.ahmed@usask.ca).

Francis M. Bui is with the Department of Electrical and Computer Engineering, University of Saskatchewan, Saskatoon, SK S7N 5A9, Canada (e-mail: francis.bui@usask.ca).

Sobhy M. Ibrahim is with the Department of Biochemistry and College of Science, King Saud University, Riyadh 11451, Saudi Arabia (e-mail: syakout@ksu.edu.sa).

Digital Object Identifier 10.1109/JBHI.2022.3149288

I. Introduction

SKIN disease is a terrific and common illness worldwide. There are multiple agents that determine the influence of skin diseases, from environmental factors to genetic susceptibility. Multiple social factors such as poverty, affluence, inequality, education, and access to health care are all also responsible. According to the Global Burden of Disease (GBD) Study 2010 [1], in the cause of nonfatal diseases in the world as a burden, skin disease was fourth on the list of most common diseases. Skin diseases cause a wide range of problems in both low and high-income countries, including psychological and sociological issues [2]. Its psychological impact is damning. A person with skin disease has anxiety, depression, anger, social isolation, and low self-esteem [3]–[5]. There is an anticipation that skin disease is remediable and medicable if it is detected before a prolonged period. However, dermatologists have a difficult time detecting skin diseases because many of them have the same appearance in terms of color and anatomy [6]. But machine learning has enabled a staggering alteration in medical imaging, especially in disease detection. With the improvement of the processing power of computers and the immeasurable amount of data availability [7], machine learning models have shown human-level activities in medical science. For example, CNN has accelerated progress in medical image processing (e.g., CT scan, MRI) [8]. Clinical images are inadequate for research due to different resolutions, complex contexts, and privacy concerns, particularly with sensitive body part images. Besides, the image of the skin disease dataset is not clearly labeled with information. Moreover, the number of accessible datasets with labeled information is very low. So, research on skin images is troublesome for all the aforementioned conditions. But, there is a problem with machine learning. In it, all the data is gathered in one location, usually a data center, which might possibly breach user privacy and data confidentiality laws. An emerging concept called federated learning will address both of these problems. The data is disseminated across the clients in federated learning, and then a prototype of the central model is delivered to the clients. The transmitting model is trained with client data at each client site, and each client model transmits the update weights or gradient to the central model after a period of time. Finally, the central model is updated using the federated averaging process, and the revised model prototype is provided to clients.

II. Background Study

Abundant research articles have been published for skin disease detection and classification. Among them, many researchers have applied deep learning algorithms for skin disease classification [3]. For instance, Esteva *et al.* [9] showed good accuracy in classifying skin tumors using the V3 inception architecture in 2017. Two dermatological tests showed an accuracy of 55.0% and 53.3% for nine classes of tumors, respectively. The model showed an average accuracy of 55.4%. In the same year, Codella *et al.* studied a nonlinear support vector machine (SVM) algorithm for detecting melanoma, using 70% of the data for training and 30% of the data for testing [11]. The studies achieved an average accuracy of 76%. In 2018, Zhang *et al.* [10] also proposed the V3 inception architecture on dermoscopic images using the same network for classifying four common skin diseases such as psoriasis, SK, BCC, and melanocytic nevus. The outcomes achieved an accuracy of 87.25%. The accuracy of the aforementioned studies is never greater than 90%. Shanthi *et al.* classified diabetic retinopathy images of the Messidor dataset using an altered AlexNet architecture [12]. They have been classified based on four classes: healthy retina, DR stage 1, DR stage 2, and DR stage 3. The AlexNet model showed maximum accuracy of 96.6% in both the DR stage 1 and DR stage 3. Tushabe *et al.* studied different machine learning algorithms: k-nearest neighbor classifier (KNN), 2-norm support vector classifier (SVC), nave Bayes (NB), and two-layer perceptron neural network (NN) to detect skin disease, virus infection, and bacterial infection [13]. The KNN classifier showed a maximum accuracy of 100%, and the SVM classifier showed the second-highest accuracy of 92%. Although they have received maximum accuracy, some crucial metrics (e.g., f1 score, recall, precision) are not described. In addition, the number of applied datasets is undefined. Sheha *et al.* [14] proposed a classification study to detect melanoma based on dermoscopy images using a multi-layer preceptor (MLP) classier. The MLP showed a training accuracy of 100% and a testing accuracy of 92%. Gurovich *et al.* [15] proposed a CNN algorithm called DeepGestalt, and the model was trained using 17,000 face images of genetic syndromes. Their model can truly identify more than 200 images of genetic syndrome with high precision. To our best of our knowledge, up to now, there has been no designed model for skin disease classification considering the privacy of the images. However, some researchers are attempting to apply federated learning in medical health. Olivia *et al.* [16] introduced a federated learning framework capable of learning a global model using dispersed health data kept locally at many sites. They demonstrated the feasibility and efficacy of the federated learning framework using real-world electronic health data over 1 million patients while retaining the global model's relevance. On the BraTS dataset, Li *et al.* [17] evaluated the potential of using differential-privacy approaches to secure patient data in a federated learning setting for brain tumor segmentation. In 2020, Boyi *et al.* [18] presented federated learning for COVID-19 data training and undertake experiments to validate its performance. They also evaluated the effectiveness of four prominent models with and without the federated learning framework. However, in this study, we considered convolutional neural networks (CNN) for the classification of skin diseases and federated learning algorithms for the retention of data security. The novelties that were introduced in this paper are as follows:

- A custom image dataset was prepared with four distinct classes of skin disease. Moreover, we augmented the data to enhance the model generality as the amount of data is not sufficient.
- A novel CNN model was introduced to classify the four types of skin disease. To optimize the parameters, we employed a hyper-parameter tuning technique. Finally, we compared the proposed model with different CNN benchmark algorithms.
- Since some skin disease images, especially private body parts and organs, are extremely sensitive, we experimented with a federated learning approach to enhance the security of medical imaging using the custom dataset.

The skin disease images were collected from AtlasDerm [19], Derm101 [20], Dermnet [21]. There is limited data in those resources [19]–[21] for similar classes. We only found 849 image data from different accessible resources. Any researcher would struggle to perform research with the available data. Therefore, developing a specialized dataset of skin disease images in the future would be a research undertaking.

III. Methodology

Different CNN algorithms, such as VGG16 and AlexNet models, are quantitatively discussed and demonstrated in this section. Finally, the notion of a federated learning framework is discussed in this section. A neural network is a totally mathematical structure that has an input layer, many hidden layers, and an output layer. Each layer has a number of biological neurons that are linked by weight parameters. As a constraint, the loss function is utilized. The parameters and loss are optimized via the backpropagation approach. A CNN is a representation of a neural network with many layers: convolution layers, activation layers, pooling layers, and fully connected layers. Fig. 1 depicts the proposed CNN network topology, including the size and dimensions of each layer and filter.

A. Convolution Layer (ConVo)

A kernel is utilized in the convolution layer to identify spatial patterns such as edges in the input picture data using matrix multiplication (see Fig. 2). Predefined kernel sizes are 3×3, 5×5, and 7×7. After the multiplication procedure, a feature matrix is created. (1) is used to compute the dimension of the feature matrix. This layer is a conventional neural network that accepts

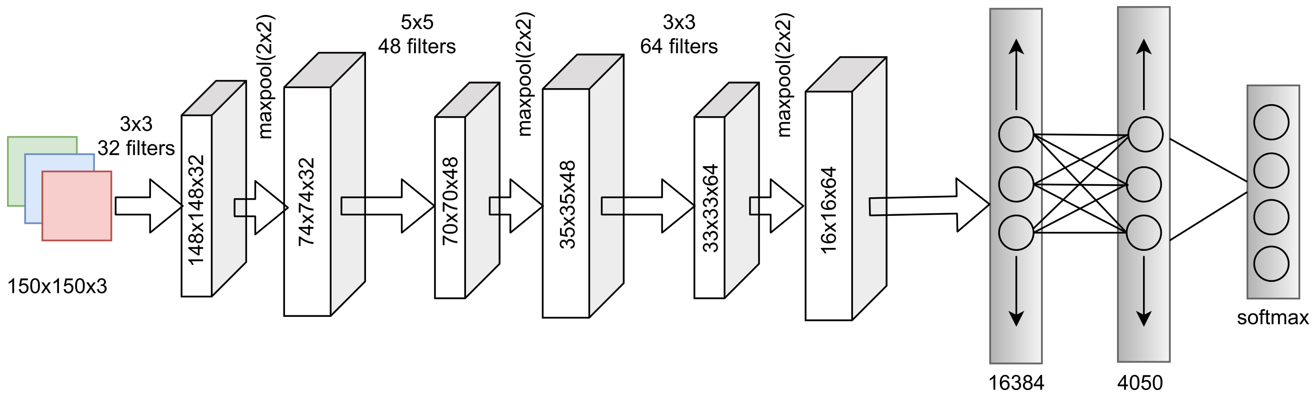

Fig. 1. Proposed model structure for the skin disease classification.

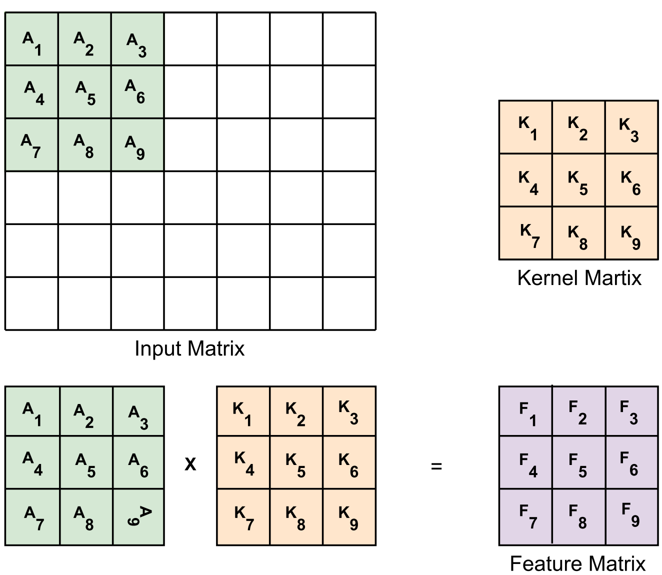

Fig. 2. Convolution layer operation.

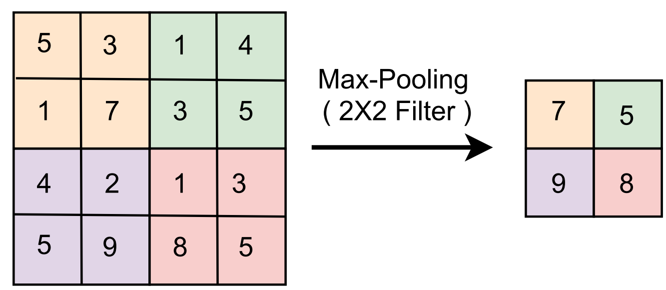

Fig. 3. Pooling layer operation.

TABLE I
PROPOSED MODEL SUMMARY FOR SKIN DISEASE CLASSIFICATION

Layer	Size	Filter size	Stride	Activation
Input	150 x 150 x 3	-	-	-
Convolution-1	148 x 148 x 32	5 x 5	1	ReLu
Maxpooling	74 x 74 x 32	2 x 2	1	ReLu
Convolution-2	70 x 70 x 48	3 x 3	1	ReLu
Maxpooling	35 x 35 x 48	2 x 2	1	ReLu
Convolution-3	33 x 33 x 64	5 x 5	1	ReLu
Maxpooling	16 x 16 x 64	2 x 2	1	ReLu
FC	16384	-	-	ReLu
FC	4050	-	-	ReLu
FC	4050	-	-	ReLu
FC	4	-	-	Softmax

a 1D array as input and computes class-wise probabilities before classifying the correct class. Table I shows the model summary for the introduced CNN model, which comprises 9 layers.

$$\text{Output} = \frac{(\text{In} - \text{Fs} + 2 \times \text{Pd})}{\text{St}} \quad (1)$$

where, In, Fs, Pd, and St are the input image, the filter size, the padding, and the stride, respectively.

B. Pooling Layer

In this layer, a filter operates upon the feature matrix independently to form a new dimension-reduced feature matrix. Two types of pooling layers, average-pooling and maximum-pooling, are used to form a new feature matrix. In max-pooling, the new feature matrix for each set is the maximum value of each patch selected by the filter, and the max-pooling approach is shown in Fig. 3. But in avg-pooling, the new set values are the average values of each patch. The newly formed feature matrix dimension is calculated using (2).

$$\text{Output} = \frac{(\text{Nh} - \text{Fs} + 1)}{\text{St}} \times \frac{(\text{Nw} - \text{Fs} + 1)}{\text{St}} \times \text{Nc} \quad (2)$$

where, Nh, Nw, and Nc are the height, width, and the number of channels in the feature map, respectively.

C. Activation Function:

The excitement of a neuron is assembled into a function of an artificial neural network (ANN) to learn the complex patterns of data. This function is called the "activation function". The activation function decides what is to be fired to the next neuron. This is the particular task that the activation function actually does. Various broadly used activation functions include ReLu (Rectified Linear Unit), Sigmoid function, and Tanh. Equation

TABLE II
VGG16 NETWORK SUMMARY FOR SKIN DISEASE CLASSIFICATION

Layer	Size	Filter size	Stride	Activation
Input	224 x 224	-	-	-
2xConvolution	224 x 224 x 64	3 x 3	1	ReLu
Maxpooling	112 x 112 x 64	3 x 3	2	ReLu
2xConvolutio	112 x 112 x 128	3 x 3	1	ReLu
Maxpooling	56 x 56 x 128	3 x 3	2	ReLu
2xConvolution	56 x 56 x 256	3 x 3	1	ReLu
Maxpooling	28 x 28 x 256	3 x 3	2	ReLu
3xConvolution	28 x 28 x 512	3 x 3	1	ReLu
Maxpooling	14 x 14 x 512	3 x 3	2	ReLu
3xConvolution	28 x 28 x 512	3 x 3	1	ReLu
Maxpooling	7 x 7 x 512	3 x 3	2	ReLu
FC	25088	-	-	ReLu
FC	4096	-	-	ReLu
FC	4096	-	-	ReLu
FC	4	-	-	Softmax

(3) represents an activation function used in the proposed model.

$$f(x) = \begin{Bmatrix} 0 \text{ if } x \leq 0 \\ x \text{ if } x > 0 \end{Bmatrix} \quad (3)$$

D. Fully Connected Layer

This layer is a regular neural network that takes input as a 1D array and computes class-wise probability and finally classifies the right class based on the probability. The introduced CNN model has 9 layers, and the model summary is indicated in Table I.

E. VGG16

VGG16 is a 16-layer convolutional neural network. This network is kind of enormous, with 138 million parameters. The inventors of VGG16 had the same problem in that they concentrated on the convolution layers with 3×3 filters and stride 1 rather than a large number of hyper-parameters. VGG16 has the same padding and a maximum pool of 2×2. Convolution layers, max-pooling layers, and two fully connected layers with a softmax layer are the layers in VGG16's layer order. Table II depicts the VGG16 network summary. In 2014 [22], simionian *et al.* proposed the VGG16 network with a large scale image recognition setting, and it obtained first and second place in localization and classification at the ImageNet challenge in 2014.

F. Alexnet

For computer vision applications, the AlexNet structure works well. It contains three completely linked layers and five convolution layers. In Table III, you'll find a structured overview of AlexNet. To extract features, many convolution layers are employed. The output was injected into the overlapping max-pooling layer by the first two convolution layers. The following three layers are then joined in a straight line. An overlapping max-pooling layer follows the neighboring fifth convolution layer. The inputs are then sent straight to the fully connected layers. Finally, a softmax layer is utilized to classify the data.

TABLE III
ALEXNET NETWORK SUMMARY FOR SKIN DISEASE CLASSIFICATION

Layer	Size	Filter size	Stride	Activation
Input	227 x 227 x 3	-	-	-
Convolution-1	55 x 55 x 96	11 x 11	4	ReLu
Maxpooling	27 x 27 x 96	3 x 3	2	ReLu
Convolution-2	27 x 27 x 256	5 x 5	1	ReLu
Maxpooling	13 x 13 x 256	3 x 3	2	ReLu
Convolution-3	13 x 13 x 384	3 x 3	1	ReLu
Convolution-4	13 x 13 x 384	3 x 3	1	ReLu
Convolution-5	13 x 13 x 256	3 x 3	1	ReLu
Maxpooling	6 x 6 x 256	3 x 3	2	ReLu
FC	9216	-	-	ReLu
FC	4096	-	-	ReLu
FC	4096	-	-	ReLu
FC	4	-	-	Softmax

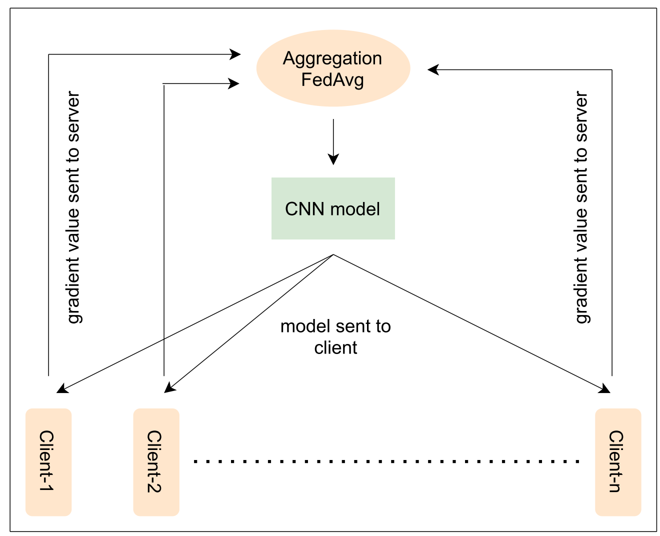

Fig. 4. The Flowchart of the federated machine learning (FedML) framework for the proposed model.

ALex *et al.* [23] developed the network which competed in the ImageNet Large Scale Visual Recognition Challenge and showed top-5 errors of 15.3% more than 10.8%.

G. Federated Learning

For data privacy, federated learning models, which are trained with decentralized data, are an emerging concept. They are able to maintain anonymity while minimizing latency. A central model copy is delivered to all devices in federated learning. Each device's user input data is used to train the models. The trained results are then forwarded to the server, where they are aggregated and the central model is updated. The functioning principle of the federated learning framework is depicted in Fig. 4. A federated learning approach, especially for medical data, has a significant impact on security and data confidentiality. In 2019, augenstein *et al.* [23] created a generative model based on federated learning when data cannot be viewed directly for troubleshooting many common data issues.

Fig. 5. Some sample images of different class.

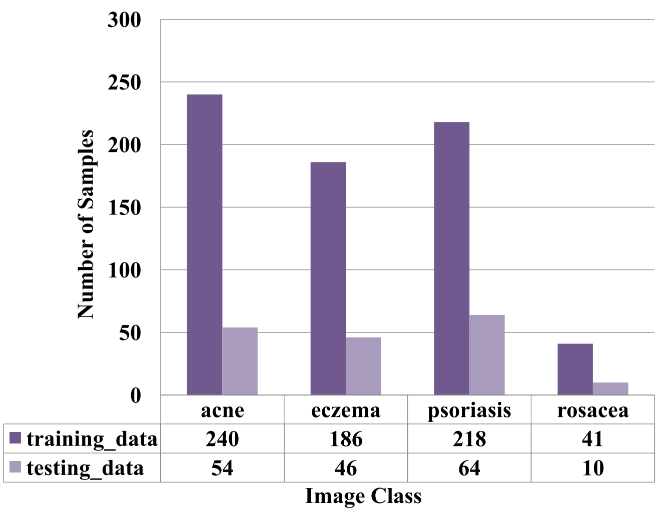

Fig. 6. Dataset distribution into training data and testing data.

IV. Image Dataset

First, a dataset comprising four distinct classes of skin diseases such as acne, psoriasis, eczema, and rosacea was created for this study. Fig. 5 presents a range of image data samples. The dataset's images were all gathered from the DermNet image library. New Zealand dermatologists contribute to DermNet, the world's biggest picture database. Five dermatologists assessed all of the assembled image data as well as the level of each data point. The prepared dataset comprises 849 images divided into two categories: training data (685 images) and testing data (164 images). The data distribution ratio between the training and testing sets in each class is 7:3. In addition, Fig. 6 depicts the distribution of the dataset into training and testing data for each class. Data is considered non-id in the federated learning method, and the entire dataset is spread across the clients at random.

V. Results and Discussion

The outcomes of this research can be divided into two categories: (1) performance evaluation of CNN algorithms and two other benchmark algorithms 2) Evaluation of federated learning performance on the same datasets. The introduced model was trained using 685 images of the prepared dataset and tested with 174 images. In addition, during training, the image augmentation approach is used to enlarge the training images and to generalize the model performance. Epochs 500 and batch size of 50 are set during the training period. Three convolution layers are used to extract features, and two fully connected layers are used for training. In the first convolution layer, 48 filters with dimensions of 5×5 are used. The second convolution layer has 32 filters with a 3×3 dimension, and the last convolution layer possesses 32 filters with a 5×5 dimension. The default stride 1 and the same padding are used in all three layers. The first fully connected layer among the two has 16384 neurons, and the second fully connected has 4050 neurons. Finally, a softmax layer with four neurons is used for four class classifications.

TABLE IV
Performance Metrics of Proposed Model

Metrices	Acne	Eczema	Psoriasis	Rosacea	Average
Precision(%)	86.0	43.0	60.0	0.00	47.3
Recall(%)	67.0	60.0	60.0	0.00	46.8
Accuracy(%)	-	-	-	-	83.0

All of these aforementioned parameters had been set through a hyper-parameter tuning strategy, and it had taken a long time. In the proposed model, the input image dimension is 150×150. The output dimension of the first convolution layer followed by a 2×2 max-pooling layer is 74×74 which is used as the input of the second convolution layer. And, the second convolution layer output dimension followed by a 2×2 max-pooling layer is 35×35. The third convolution layer produces 16×16 dimensional output followed by a 2×2 max-pooling layer. After flattening the third convolution layer output into 1D, it is inserted into the first fully connected layer which is connected with the second fully connected layer. The final layer, softmax has connected with the second fully connected layer. The ultimate output dimension of the proposed model is 4×1 which is the output of the softmax layer. The proposed model performance matrix is shown in Table IV, which reveals that the average precision and recall are 47.3% and 46.8%, respectively, although the value of rosacea is 0.00% in both cases, as the number of rosacea images is too small.

Next, another two benchmark algorithms (AlexNet and VGG16) were also trained and tested with the same dataset. But they did not show better performance for acne, eczema, and psoriasis compared with the proposed model. Nevertheless, they showed a little bit better performance for rosacea only. The comparison metrics among different benchmark algorithms and proposed models are displayed in Table V. In Table V, we notice that the introduced model presents higher precision and recall for acne, eczema, and psoriasis compared with benchmark algorithms. Besides, accuracy is also higher for the proposed model compared with the other two algorithms. The learning parameter of the proposed model is around 1.5 million, which is smaller than VGG16 and AexNet. Moreover, the proposed model loss is also evaluated, which is lower compared with the

TABLE V
PERFORMANCE METRICS OF DIFFERENT NETWORK AND PROPOSED MODEL

Network	Metrices	Acne	Eczema	Psoriasis	Rosacea	Average
AlexNet	Precision(%)	50.0	40.0	33.0	5.00	32.0
	Recall(%)	70.0	45.0	20.0	10.2	36.3
VGG16	Precision(%)	33.0	38.0	10.0	0.50	20.3
	Recall(%)	25.0	18.0	20.0	24.2	21.8
Proposed network	Precision(%)	86.0	43.0	60.0	0.00	47.3
	Recall(%)	67.0	60.0	60.0	0.00	46.8

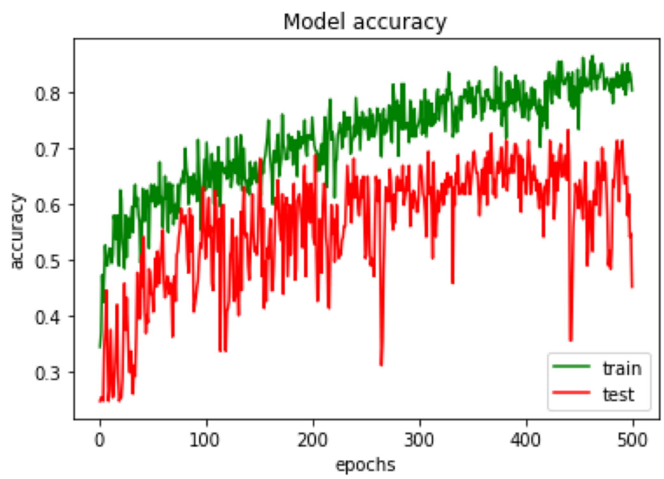

Fig. 7. Training accuracy and testing accuracy of proposed model (learning rate=0.01, batch size=50 and epochs=500).

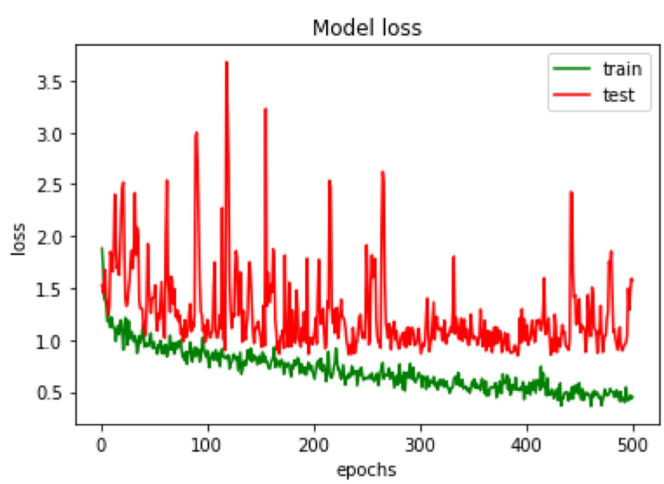

Fig. 8. Training loss and testing loss of proposed model (learning rate=0.01, batch size=50 and epochs=500).

other two algorithms. Initially, in the proposed model, the loss is higher, but after some epochs, the loss is decreasing. The introduced model accuracy and loss are shown in Figs. 7 and 8, apart. We also assessed the confusion matrix of the proposed model, which is shown in Fig. 9. The confusion matrix of the introduced model shows that the model prediction is worst for the rosacea class because of the small amount of data. The worst accomplishment of the proposed model for the rosacea class can be improved by increasing the amount of data. In the future,

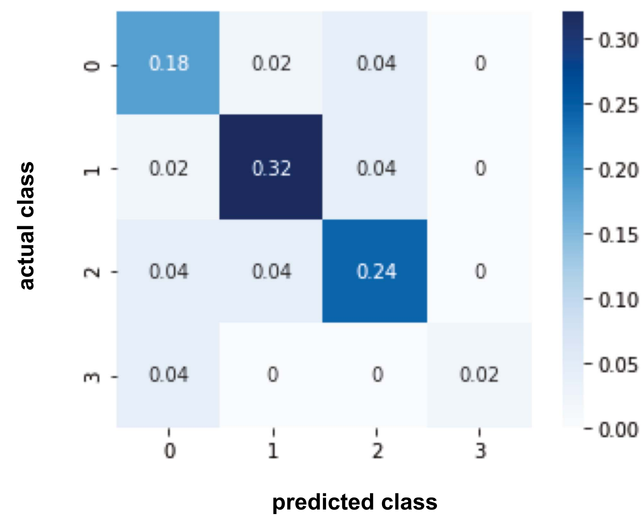

Fig. 9. Confusion matrix of proposed model and 0–3 represent acne, eczema, psoriasis and rosacea respectively.

TABLE VI
PERFORMANCE METRICS OF FEDERATED LEARNING MODEL

No. of clients	Average accuracy(%)	Average loss(%)
1000	81.21	0.33
1500	86.57	0.21
2000	91.15	0.17
2500	94.15	0.13

we will be trying to improve the overall model performance in all four classes, especially for rosacea. From a privacy point of view, we also demonstrated the federated learning algorithm with the same dataset and calculated the average accuracy and average loss by varying the number of clients. We also used the FedAvg generalized approach to update the central model, and in FedAvg, the client model sends updated weights rather than gradients after a series of batch updates on the client device. The FedML model metrics are shown in Table VI. From Table VI, it is clearly visible that the accuracy is increasing with the enlargement of clients or devices and conversely with the average loss profile. This proposed model showed a maximum accuracy of 94.15% for 2500 clients. However, the FedML showed lower accuracy than CNN algorithms. But, the accuracy of the FedML will increase day by day after increasing the training images. The most beneficial point of FedML is privacy. It helps to conserve data privacy across various clients because only the updated weights are shared with the centralized model. So anyone can train a model using personal data without sharing the original data from their own devices. To the best of my knowledge, this is the first time we have applied the federated learning concept to skin disease classifications. So, this approach will reveal a new dimension in technological society.

The proposed model can only classify four types of skin diseases, although a huge number of skin diseases are seen in humans. Moreover, the model cannot measure the degree of severity of the disease. In addition, due to the lack of image sources, we trained our model for a small amount of data. So, performance could be varied after training more data. In the future, we will try to increase the amount of data in the dataset and enhance the federated learning model's performance from several perspectives.

The proposed work will assist dermatologists in correctly detecting skin diseases in a short period of time. The federated learning approach for skin disease classification is helpful for patients without sending private data to the server.

VI. Conclusion

In this study, a custom image dataset was prepared, a novel CNN model was proposed, and we explored a federated learning technique to address data privacy. The proposed CNN model is compared to the performance of two other benchmark CNN algorithms (AlexNet and VGG16). All three models, including the proposed model, are trained and tested using a learning rate of 0.01, batch size of 50, and epochs of 500 with the custom dataset. The proposed model showed a higher precision of 86%, 43%, and 60% for acne, eczema, and psoriasis. In addition, the model also showed a higher recall of 67%, 60%, and 60% for acne, eczema, and psoriasis. Overall, the performance of the proposed model is privileged, which is helpful for dermatologists to diagnose diseases, but some improvement is necessary for overall performance. In contrast, the experimented federated learning algorithm showed a maximum average accuracy of 94.15% when the number of clients was set at 2500. The proposed model would be more efficient and could classify more diseases if the training data and disease class were increased. Furthermore, the CNN and federated learning-based skin disease classification models will be improved until a standard model is developed for the classification and privacy preservation of all skin diseases with greater accuracy.

References

[1] R. J. Hay *et al.*, "The global burden of skin disease in 2010: An analysis of the prevalence and impact of skin conditions," *Invest. Dermatol.*, vol. 134, no. 6, pp. 1527–1534, 2014.

[2] A. Tuckman, "The potential psychological impact of skin conditions," *Dermatol. Ther.*, vol. 7, no. 1, pp. 53–57, 2017.

[3] A. Bewley, "The neglected psychological aspects of skin disease," *Brit. Med. J.*, vol. 358, Jul. 2017, doi: 10.1136/bmj.j3208.

[4] W. Chen *et al.*, "Polymorphisms of SLCO1B1 rs4149056 and SLC22A1 rs2282143 are associated with responsiveness to acitretin in psoriasis patients," *Sci. Rep.*, vol. 8, no. 1, pp. 1–9, 2018.

[5] X. Zhou *et al.*, "Frizzled-related proteins 4 (SFRP4) rs1802073 G allele predicts the elevated serum lipid levels during acitretin treatment in psoriatic patients from Hunan, China," *PeerJ*, vol. 13, no. 6, 2018, Art. no. e4637.

[6] R. B. Roslan *et al.*, "Evaluation of psoriasis skin disease classification using convolutional neural network," *IAES Int. J. Artif. Intell.*, vol. 9, no. 2, pp. 349–355, 2020.

[7] J. Deng, W. Dong, R. Socher, L. -J. Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2009, pp. 248–255.

[8] Z. Wu *et al.*, "Studies on different CNN algorithms for face skin disease classification based on clinical images," *IEEE Access*, vol. 7, pp. 66505–66511, 2019.

[9] A. Esteva *et al.*, "Dermatologist-level classication of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, Feb. 2017.

[10] X. Zhang *et al.*, "Towards improving diagnosis of skin diseases by combining deep neural network and human knowledge," *Med. Informat. Decis. Mak.*, vol. 18, no. 2, pp. 69–76, 2018.

[11] F. Codella *et al.*, "Deep learning ensembles for melanoma recognition in dermoscopy images," *IBM J. Res. Develop.*, vol. 61, no. 4/5, pp. 1–28, 2017.

[12] T. Shanthi and R. S. Sabeenian, "Modified Alexnet architecture for classification of diabetic retinopathy images," *Comput. Electron. Eng.*, vol. 76, pp. 56–64, 2019.

[13] F. Tushabe, E. Mwebaze, and F. N. Kiwanuka, "An image-based diagnosis of virus and bacterial skin infections," in *Proc. Int. Conf. Complications Interventional Radiol.*, 2011, pp. 1–7.

[14] M. A. Sheha, M. S. Mabrouk, and A. Sharawy, "Automatic detection of melanoma skin cancer using texture analysis," *Int. J. Comput. Appl.*, vol. 42, no. 20, pp. 22–26, 2012.

[15] Y. Gurovich *et al.*, "Identifying facial phenotypes of genetic disorders usingdeep learning," *Nature Med.*, vol. 25, no. 1, pp. 60–64, Jan. 2019.

[16] O. Choudhury *et al.*, "Differential privacy-enabled federated learning for sensitive health data," 2019, *arXiv:1910.02578*.

[17] W. Li *et al.*, "Privacy-preserving federated brain tumour segmentation," in *Proc. Int. Workshop Mach. Learn. Med. Imag.*, 2019, pp. 133–141.

[18] B. Liu *et al.*, "Experiments of federated learning for COVID-19 chest X-ray images," 2020, *arXiv:2007.05592*.

[19] S. F. D. Silva and D. B. Calheiros, "Dermatology atlas," Accessed: Feb. 2, 2022. [Online]. Available: http://www.atlasdermatologico.com.br

[20] Derm101, Accessed: Feb. 2, 2022. [Online]. Available: https://www.derm101.com/imagelibrary/?match=IN

[21] T. Habif, "Dermnet," Accessed: Feb. 2, 2022. [Online]. Available: http://www.dermnet.com

[22] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*.

[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Accessed: Feb. 2, 2022. [Online]. Available: https://blog.karatos.in/a?ID=00450-9b1db208-6e87-450e-bc1d-78f9c16d5996

[24] S. Augenstein *et al.*, "Generative models for effective ML on private, decentralized datasets," 2019, *arXiv:1911.06679*.