

FedDermaSeg: Federated Learning for Dermatological Image Segmentation

Anabik Pal¹, Ganesh Patidar², and Bikash Santra³

¹ Department of Computer Science, Indian Institute of Science Education and Research Berhampur, India

² Department of Mathematics, Indian Institute of Technology Jodhpur, India

³ School of Artificial Intelligence and Data Science, Indian Institute of Technology Jodhpur, India

Corresponding author: anabikpal@iiserbpr.ac.in

Abstract. Skin cancer is a global health concern. Early detection of skin cancer lesions and treatment planning are important for reducing premature deaths. To deal with the dermatologists' limitations, automated systems are highly demanding for examining skin conditions. Lesion segmentation is one of the prerequisites for developing any computerised system. The commonly used deep segmentation networks are trained in a centralized manner, i.e. images, and their annotated segmentation masks are gathered in a centralized server, and the training is performed. This centralized training loses data privacy, which is an important concern for the medical domain. Moreover, a huge computation has to be conducted on a single server. Hence, the feasibility of applying federated learning is experimented with and analysed for skin segmentation model development. The train and valid set of the ISIC 2018 skin lesion segmentation challenge dataset is used to simulate the distributed learning environment and build a federated model. The test set of the ISIC 2018 skin lesion segmentation challenge dataset and PH^2 dataset are used for performance evaluation. As per experimental results, the federated model produces comparable performance with the centralized training and improves the local models.

Keywords: Dermatology · Image Segmentation · Federated Learning · Deep Learning · TransUNet.

1 Introduction

Skin cancer is a major threat to our society. The early detection can reduce premature death. The development of an automated system is required to reduce the experts' workload and enhance medicare for early detection of such skin conditions. Automated skin image segmentation algorithm development is an important pre-requisite for such automated systems [14]. Although several classical image processing techniques are proposed [1], their performance is limited and prone to misclassification of the image pixels. In contrast, the application of the latest deep learning-based approaches will be a better choice as they have proven to be efficient in many image analysis applications [8]. Broadly, the segmentation

network can be built with deep convolutional neural network (DCNN) based encoder-decoder-like architecture [13] or transformer-based architecture [3]. In deep networks of the first kind, the encoder obtains the low-dimensional feature representation of the input images, and the decoder finds the semantic maps of the pixels using the low-dimensional feature representation. On the other hand, transformer-based approaches find low-dimensional image representation with attention mechanisms, and by combining the feature representation with object queries, the object regions are obtained. Note that the DCNN-based architectures fail to capture the global context of an image during feature learning [16], and the transformer networks are not capturing local interactions in the feature maps [15]. Hence, several segmentation architectures are designed that hybridize the potential of both CNN and transformer networks [4] and applied in skin lesion segmentation purposes [6, 7].

The key source of high performance in the above-mentioned approaches lies in the use of a huge volume of data samples during training. Getting a good volume of curated and annotated dataset(s) from a single clinic or hospital is often challenging. The easiest solution is to use a single repository and gather data prepared by all collaborated sources. However, this may violate data privacy and clinical ethics. Hence, federated learning is used as an effective alternative to train a robust model in collaboration with multiple sources without violating data privacy. In spite of data privacy control, federated learning distributes the computation among the collaborated parties (sources) and makes the computation de-centralized [17]. In general setting, federated learning is performed with the help of a parameter server. First, the parameter server creates a deep model and initializes the weight matrices. Then the model is sent to the collaborated clients for weight update based on their locally available dataset. Then, the updated models are transferred to the server for aggregation, which produces a global collaborative model. This process continues until the model converges. The model aggregation is performed with the FedAvg algorithm, which uses a weighted average of the weight matrices [9].

This paper develops a deep segmentation model with federated learning with FedAvg algorithm [9] in the cross-silo setting, systematically evaluating skin lesions' performance and analyzing its performance trade-offs. Equal weights are given to the clients during model aggregation in the server. As we are interested in analysing the algorithmic effectiveness of the proposed approach, we create a synthetic environment to create the clients and the parameter server. ISIC 2018 skin lesion segmentation challenge training dataset [5] is used for training the federated segmentation model. After training, the performance of the federated model is analysed with ISIC 2018 skin lesion segmentation challenge test dataset and PH^2 dataset [10]. Literature shows that some works are available to develop deep skin image classification models [2] with federated learning. However, as per our survey, limited work attempts to develop a federated segmentation model for skin lesion segmentation. In contrast to that, this work employs a federated learning algorithm for building a skin lesion segmentation model, analyzes the

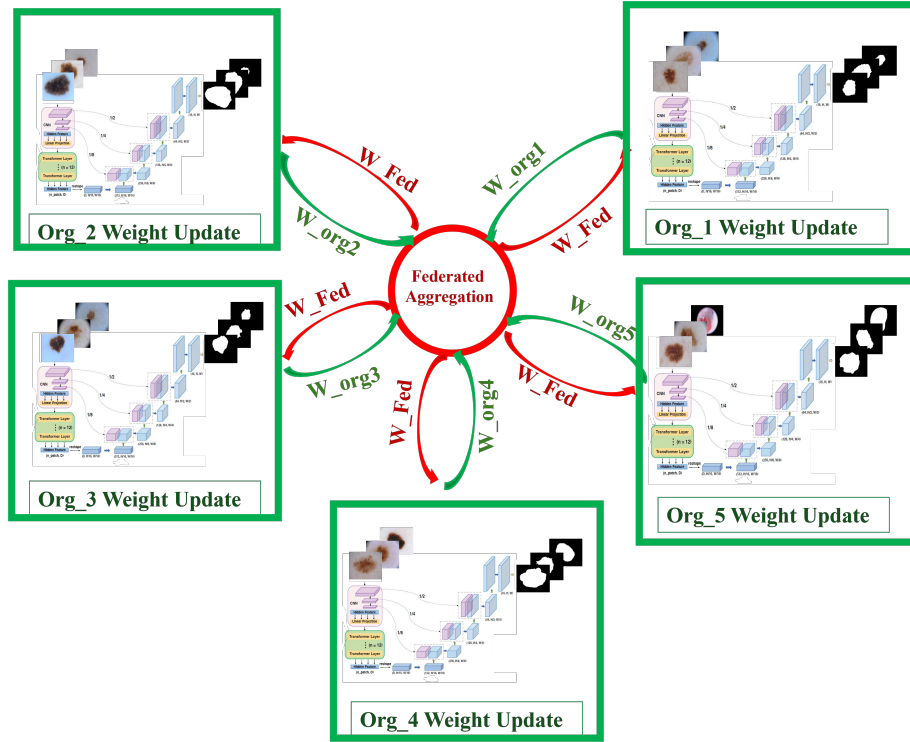


Fig. 1. The flow diagram of training segmentation network with federated learning. Org_i refers to i^{th} client; W_Org_i refers to local weight update in i^{th} client during a communication round; W_Fed refers to global model after local model aggregation.

performance of commonly used deep segmentation networks and analyzes the performance of commonly used segmentation loss functions.

The rest part of the paper is organized as follows: Section 2 discusses the proposed methodology; Section 3 presents the experimental setup; the experimental analysis is kept at Section 4; and finally, Section 5 concludes the paper.

2 Methodology

The block diagram of our work is presented in Fig 1. Fig 1 shows that there are five clients surrounded by green borders where local model training is happening. Note that in this paper, we are considering five clients; however, a finite number of clients can participate in the federated training process. The parameter server is shown in the middle part of this diagram and shown as a red-marked circle. At the beginning, a deep image segmentation network with initial weight matrices is sent to all clients. On receiving the model and its parameters, each client trains the model with their local annotated images. After E number of local model updates, the updated models are sent to the server for aggregation. The

server then aggregates the clients' models to have a single model. This process is continued for T rounds of communications. Note that we apply the federated average algorithm [9] for preparing the global model and consider equal weights for all clients. For the reader's convenience, the training mechanism is given in the Algorithm 1.

Algorithm 1 Federated Learning Algorithm

Input: $\{D_i\}_{i=1}^N$: datasets for N clients; T : total rounds; E : local epochs; η : learning rate

Output: Global model M_G

```

1: Initialize global model  $M_G$  with random weights  $\theta_G$ 
2: for  $t = 1$  to  $T$  do
3:   Broadcast  $\theta_G$  to all clients
4:   for all clients  $i \in \{1, \dots, N\}$  (in parallel) do
5:     Initialize  $M_i \leftarrow \theta_G$ 
6:     for  $e = 1$  to  $E$  do
7:       Train  $M_i$  on  $D_i$  with learning rate  $\eta$  to get  $\theta_i$ 
8:     end for
9:   end for
10:  Aggregate:  $\theta_G \leftarrow \frac{1}{N} \sum_{i=1}^N \theta_i$ 
11: end for
12:
13: return  $M_G$ 

```

Regarding segmentation architecture, we experiment with two state-of-the-art deep image segmentation networks, namely UNet [13] and TransUNet [4]. In short, the UNet architecture contains a series of convolution layers and deconvolution layers. The role of convolutional layers is to find the lower-dimensional feature representation of the input images, and the deconvolutional layers are kept to obtain the segmentation masks in the original dimension. In UNet, there is neuronal interaction in the local regions only. In contrast, TransUNet uses Convolutional layers to find the lower-dimensional representation of the input image with local neuronal interaction. Then, the lower-dimensional representations are processed in the transformer layers. In the transformer layers, the final image embedding is prepared with global neuronal interactions. The learnt feature is then processed with deconvolutional layers. Both of the networks are trained with the above-mentioned federated learning algorithms and compared.

As per our workflow, the training or weight update happens only for the clients. We use a combination of the following state-of-the-art loss functions: (i) Binary Cross Entropy (BCE), (ii) Dice Loss, (iii) Focal Loss. Note that the lesional areas are much less than the normal skin regions in the input images; hence, the BCE loss can be biased towards detecting normal regions. This motivate us to incorporate Dice Loss and Focal Loss, which have some capability to deal with this bias. The mathematical definition of those loss functions are given below:

$$\text{BCE}(y, \hat{y}) = -(y \cdot \log(\hat{y}) + (1 - y) \cdot \log(1 - \hat{y})) \quad (1)$$

$$\text{Dice Loss}(y, \hat{y}) = 1 - \frac{2 \sum_{i=1}^N y_i \hat{y}_i}{\sum_{i=1}^N y_i + \sum_{i=1}^N \hat{y}_i + \epsilon} \quad (2)$$

$$\text{Focal Loss}(y, \hat{y}) = -\alpha (1 - \hat{y})^\gamma y \cdot \log(\hat{y}) - (1 - \alpha) \hat{y}^\gamma (1 - y) \cdot \log(1 - \hat{y}) \quad (3)$$

In the above equations, y is the ground-truth mask; $\hat{y}$ is the predicted mask; N is the total number of pixels; α, γ are scalar parameters; ϵ is a small number which is dealing with division by zero error.

3 Experimental Setup

3.1 Datasets

We use the following two publicly available dermoscopic image datasets for our research:

ISIC: The International Skin Imaging Collaboration (ISIC) prepared multiple dermoscopy image datasets and made them publicly available as challenge datasets [5]. We use the **ISIC 2018 skin image segmentation challenge** dataset as it contains the maximum number of skin lesion segmentation ground-truths. The dataset is organised into three subsets: training (2,594), validation (100), and test (1,000). We use available lesion segmentation ground truth annotations for our research. Note that the images have varying appearances and contain diverse clinical conditions, such as melanoma, melanocytic nevus, basal cell carcinoma, squamous cell carcinoma, vascular lesions, dermatofibroma, and benign keratosis.

PH²: PH² dataset was made publicly available in 2013 by the Universidade do Porto, Instituto Superior Técnico Lisboa, and the Hospital Pedro Hispano Dermatology Service [10]. It contains 200 annotated (lesion ground-truth) dermoscopy images having the following three categories: nevus (80), atypical (80), and melanomas (40).

3.2 Data Simulation for FL Setup

This study simulates the decentralized learning environment. We consider five clients so we need to have five disjoint datasets that will serve as local datasets. For that, we split the **training set** of the ISIC 2018 skin lesion segmentation challenge dataset into five disjoint sets- four sets containing 519 images and the remaining one containing 518 images. We hypothesize that simulating datasets through random splitting introduces a certain degree of non-IID data distribution across clients, mirroring what is commonly observed in real-world distributed learning scenarios. The **validation set** of the ISIC 2018 skin lesion segmentation

Table 1. Performance in the test set of ISIC 2018 skin lesion segmentation challenge dataset

Network	Model	Dice	IoU	Accuracy	Precision	Recall	F1
UNet	local-1	0.882	0.789	0.937	0.934	0.835	0.882
	local-2	0.883	0.791	0.936	0.92	0.849	0.883
	local-3	0.880	0.786	0.935	0.912	0.850	0.880
	local-4	0.877	0.782	0.933	0.909	0.848	0.877
	local-5	0.881	0.787	0.935	0.911	0.853	0.881
	global-L	0.919	0.850	0.950	0.908	0.930	0.919
	global-FL	0.899	0.816	0.945	0.934	0.866	0.899
TransUNet	local-1	0.885	0.793	0.936	0.901	0.869	0.885
	local-2	0.883	0.791	0.936	0.905	0.862	0.883
	local-3	0.888	0.799	0.938	0.904	0.873	0.888
	local-4	0.878	0.783	0.934	0.923	0.838	0.878
	local-5	0.878	0.783	0.934	0.923	0.838	0.878
	global-L	0.927	0.864	0.959	0.933	0.921	0.927
	global-FL	0.905	0.828	0.946	0.904	0.907	0.905

challenge is used to assess the performance of the local training updates as well as the global model updates after federated aggregation. Both the **test set** of the ISIC 2018 skin lesion segmentation challenge and the PH^2 **dataset** are used for the performance evaluation of the trained model.

3.3 Implementation Details and Performance Metrics

The PyTorch deep learning framework [12] is used for implementing the proposed algorithm. The following deep segmentation networks: UNet [13] and TransUNet [7] are considered for experimental evaluation. For federated learning, the local models update weights for K interaction numbers before aggregating in the central server. The whole weight update is performed for T rounds. The value of K , T is decided experimentally as presented in Section 4. The models are trained with a stochastic gradient descent algorithm with an initial learning rate of 0.001 and momentum of 0.9. An experiment is carried out to determine the right combination of losses in our ablation study presented in Section 4. For comparison with the local models built with the local dataset and the centralized training having the whole dataset, an initial learning rate of 0.001 and momentum of 0.9 is used.

In this work, the skin lesion segmentation models are evaluated with the following five commonly used metrics for performance analysis: Dice similarity coefficient (Dice), intersection-over-union (IoU) also known as Jaccard index or Jaccard similarity coefficient, accuracy, precision, recall, and F1 score [11].

Table 2. Results of various models trained with ISIC 2018 and tested with the PH^2 dataset

Network	Model	Dice	IoU	Accuracy	Precision	Recall	F1
UNet	local-1	0.854	0.745	0.911	0.905	0.808	0.854
	local-2	0.851	0.741	0.907	0.88	0.824	0.851
	local-3	0.858	0.751	0.911	0.887	0.831	0.858
	local-4	0.847	0.735	0.903	0.861	0.833	0.847
	local-5	0.846	0.733	0.905	0.888	0.808	0.846
	global-L	0.869	0.769	0.915	0.862	0.877	0.869
	global-FL	0.877	0.781	0.922	0.893	0.861	0.877
TransUNet	local-1	0.886	0.796	0.927	0.886	0.887	0.886
	local-2	0.847	0.735	0.899	0.83	0.866	0.847
	local-3	0.879	0.784	0.921	0.875	0.882	0.879
	local-4	0.883	0.791	0.928	0.928	0.842	0.883
	local-5	0.885	0.794	0.926	0.892	0.878	0.885
	global-L	0.894	0.809	0.930	0.872	0.918	0.894
	global-FL	0.885	0.793	0.924	0.872	0.898	0.885

4 Results and Discussions

This section discusses the experimental results obtained from our research. We are interested in comparing the performances of the federated learnt model with the following two scenarios: (i) the local models with their own data and (ii) a single model prepared with centralized training having all images and their ground truths. The first competing approach will advocate the efficacy of the federally learnt model, and the second one can provide an idea of how much the performance is dropped when centralised training is impossible. The performance with respect to the test set of ISIC 2018 skin lesion segmentation is available at Table 1 and the performance with the images available in PH^2 dataset is available in Table 2. In the rest of this section, the models trained locally with its own dataset (as mentioned in Section 3.2) are denoted by ‘**local-1**’, ‘**local-2**’, ‘**local-3**’, ‘**local-4**’, and ‘**local-5**’; the model trained with federated learning is denoted as ‘**global-FL**’; finally, the models trained centrally with the full dataset is denoted by ‘**global-L**’.

Table 1 shows that for both UNet and TransUNet, the local models (local-1 to local-5) trained with local datasets perform inferior to the centrally trained global model (global-L), which has access to all client’s images. This justifies that if a centralized model-building set-up exists, the centrally trained global model (global-L) would perform better than an individual client for the chosen dataset. The probable reason is that such a model would have access to more images, which generalized the input in a better way. The federated global model (global-FL) performs better than the best local model for both segmentation networks. This advocates that federated learning can capture the data diversity available to different clients without accessing the raw images. However, the

performance of the federated global model (global-FL) is comparable to the centrally trained global model (global-L).

As per our visual qualitative segmentation performance evaluation, it is difficult to perceive the segmentation networks' performance gap between the federated global model (global-FL) and the centrally trained global model (global-L) for the majority of the images. For visual representation, we select three images and shows their expert annotated ground truth, and their segmentation outcome with the federated global model (global-FL) and global model (global-L) in Fig. 2. It is evident that among these three images, the third image is not segmented well.

Our second experiment validated the performances of the trained models obtained from the ISIC 2018 skin lesion segmentation challenge dataset with

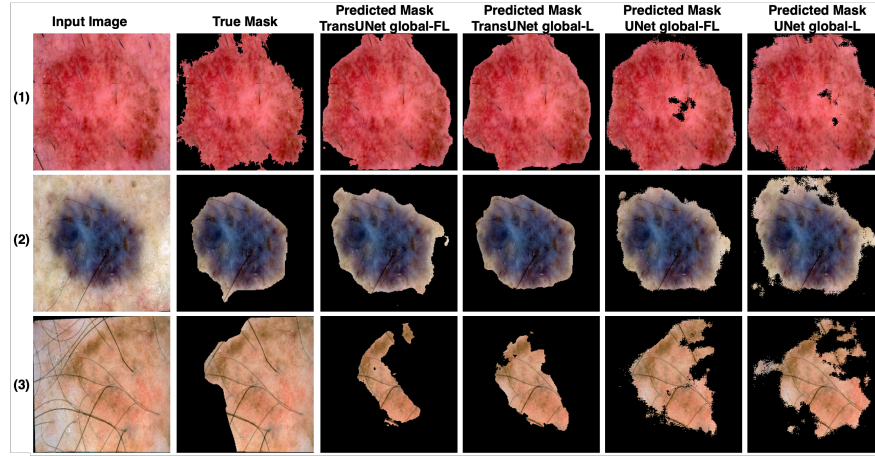


Fig. 2. Pictorial representation of segmentation results generated by the global-L and global-FL models for both TransUNet and UNet tested with ISIC 2018 dataset

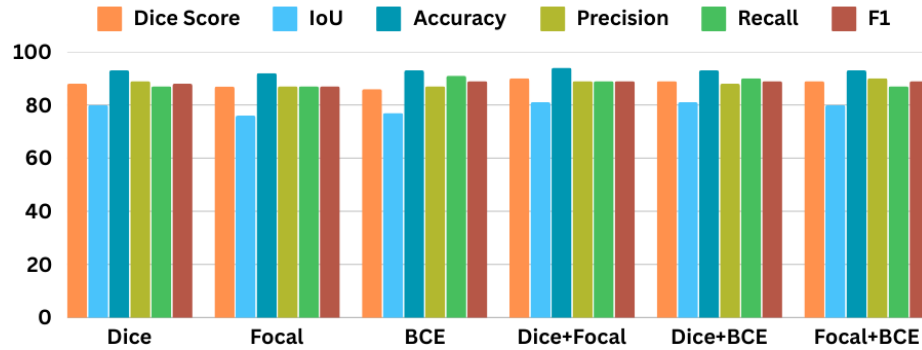


Fig. 3. Results of the global-FL TransUNet model trained and tested on ISIC 2018 dataset considering Dice, Focal, BCE, Dice+Focal, Dice+BCE, and Focal+BCE losses

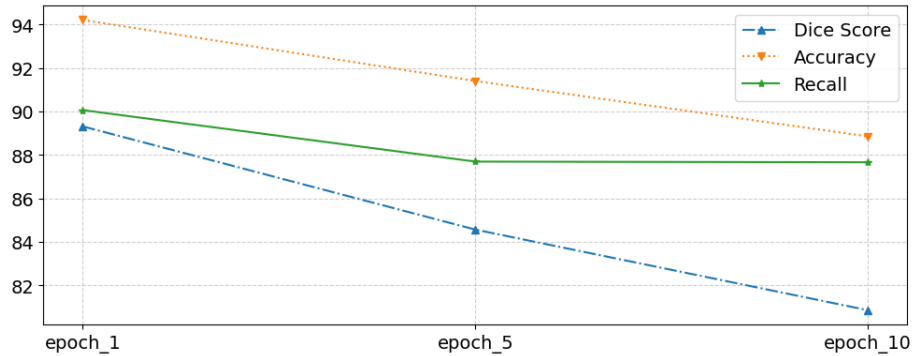


Fig. 4. Performances yield by the global-FL TransUNet model trained and tested on the ISIC 2018 dataset, when the #epochs in each iteration during federated learning is varied from 1 (epoch_1) to 5 (epoch_5) and 10 (epoch_10).

respect to PH^2 dataset (an external dataset). This will highlight the relative performance gap of the models and justify the importance of the federated global model. Table 2 contains the performance observed in PH^2 dataset. We find that the performances are degraded for all models trained with ISIC 2018 dataset. The probable reason is that the PH^2 dataset was prepared early, and the imaging device was not that much improved; thus, they might have more imaging artefacts, so segmentation is relatively difficult than the ISIC dataset. However, the relative performance gap among the models is quite similar to our first experiment. The global-FL model is performing much better than the best local model. The global-L and global-FL models for both UNet and TransUNet achieve very similar performances in terms of all of the performance metrics. Moreover, in some cases, global-FL beats global-L (for example, dice score in the case of UNet). In both our first and second experiments, TransUNet outperformed UNet by a significant margin, as shown in Tables 1 and 2. For example, the Dice score for TransUNet is $\sim 3\%$ higher than the Dice for UNet (look at the rows for the global-L of TransUNet and the global-L of UNet in Table 2). These results clearly show that TransUNet is more effective in segmenting skin lesions. Due to this, the rest of our experiments are only conducted for TransUNet.

Ablation Study. The ablation study is conducted to choose the right loss function for training the networks. In this study, we train only the model global-FL to decide the loss function by selecting an optimal combination of different loss components. The loss components are the conventional segmentation losses: Dice, Focal, and BCE. Thus, the global-FL is trained for each of the following (joint) losses: Dice, Focal, BCE, Dice + Focal, Dice + BCE, and Focal + BCE. All the models are trained with the training set of the ISIC 2018 skin lesion segmentation dataset, followed by the testing with its test set. Considering the various performance metrics, the results of this experiment are presented in a bar plot in Fig. 3. It can be seen that Dice + Focal loss shows superior or similar performances compared to other losses. However, Dice + Focal loss yields $\sim 1\%$

higher accuracy than others. Due to this, the combination of Dice and Focal losses is chosen as the loss function in our implementation.

Selecting Number of Epochs (E) in An Iteration of Federated Learning.

We also experiment to determine an optimal number of epochs in an iteration of the federated learning to train the global model (global-FL) of TransUNet. The number of epochs per iteration is set to 1 (see epoch_1), 5 (epoch_5), and 10 (epoch_10), and evaluated the performances of the global-FL TransUNet model. The Dice, accuracy, and F1 scores are shown in Fig. 4, where the model is trained and tested on the ISIC 2018 dataset. The best performance is seen when the number of epochs is set to one per iteration. Although setting one epoch per iteration is computationally expensive, we set this value in our implementation to obtain higher performances.

5 Conclusion and future scope of Work

This paper experimented with the federated learning algorithm for dermoscopic image segmentation. Two commonly known deep segmentation networks are trained with the federated learning algorithm in a simulated environment. ISIC 2018 skin lesion segmentation challenge dataset is used for training the model and testing. The performance is also analyzed with respect to the PH^2 dermoscopic image dataset. We find that the TransUNet is more suitable than UNet for skin lesion segmentation. The experimental outcome advocates that the global model trained with federated learning improves the segmentation performance in data restriction scenarios. However, in most cases, we find some performance gaps when a model is trained in a data restriction scenario rather than having full data access for centralized training. The performance obtained from the centrally trained global model shows that there is a huge scope to improve the segmentation performance so that the global model obtained from the federated learning will perform human-like.

The imaging artifact, the presence of hair in the skin, creates dermatological images that are noisy and can mislead the training. Producing a reliable, trained model with such noisy images will be the immediate future work of this research. Developing a better deep network and utilising visual foundation models for skin lesion segmentation is important for future work. Moreover, the development of a federated learning algorithm for segmenting skin lesion attributes like pigmented networks, globules, etc., is to be conducted next. In future work, we will evaluate model performance under more nuanced non-IID conditions by employing advanced data partitioning strategies, such as the Dirichlet distribution-based sampling protocol. Enhancing security during model aggregation with privacy-preserving techniques like Differential Privacy, Homomorphic Encryption is another important future scope of work.

Acknowledgements This work is partially supported by the Institute Research Initiation Grant (RIG) of the Indian Institute of Technology Jodhpur. We thank

Rahul Kumar and Aayushman for helping us organize the data and implement the proposed technique.

References

1. Barata, C., Celebi, M.E., Marques, J.S.: A survey of feature extraction in dermoscopy image analysis of skin cancer. *IEEE Journal of Biomedical and Health Informatics* **23**(3), 1096–1109 (2019). <https://doi.org/10.1109/JBHI.2018.2845939>
2. Bdair, T., Navab, N., Albarqouni, S.: Fedperl: Semi-supervised peer learning for skin lesion classification. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part III*. p. 336–346. Springer-Verlag, Berlin, Heidelberg (2021)
3. Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., Zagoruyko, S.: End-to-end object detection with transformers. In: *European conference on computer vision*. pp. 213–229. Springer (2020)
4. Chen, J., Mei, J., Li, X., Lu, Y., Yu, Q., Wei, Q., Luo, X., Xie, Y., Adeli, E., Wang, Y., Lungren, M.P., Zhang, S., Xing, L., Lu, L., Yuille, A., Zhou, Y.: Transunet: Rethinking the u-net architecture design for medical image segmentation through the lens of transformers. *Medical Image Analysis* **97**, 103280 (2024). <https://doi.org/https://doi.org/10.1016/j.media.2024.103280>, <https://www.sciencedirect.com/science/article/pii/S1361841524002056>
5. Codella, N., Rotemberg, V., Tschandl, P., Celebi, M.E., Dusza, S., Gutman, D., Helba, B., Kalloo, A., Liopyris, K., Marchetti, M., et al.: Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). *arXiv preprint arXiv:1902.03368* (2019)
6. Ding, Y., Yi, Z., Xiao, J., Hu, M., Guo, Y., Liao, Z., Wang, Y.: Cth-net: A cnn and transformer hybrid network for skin lesion segmentation. *iScience* **27**(4), 109442 (2024). <https://doi.org/https://doi.org/10.1016/j.isci.2024.109442>, <https://www.sciencedirect.com/science/article/pii/S2589004224006631>
7. HUANG, J., LAN, Q., DENG, W., HUANG, C., ZHANG, S.: Sfe-transunet: A transformer-based u-net with skipped features enhancer for medical image segmentation. *Journal of Mechanics in Medicine and Biology* **0**(0), 2440050 (2024). <https://doi.org/10.1142/S0219519424400505>, <https://doi.org/10.1142/S0219519424400505>
8. Kshatri, S.S., Singh, D.: Convolutional neural network in medical image analysis: a review. *Archives of Computational Methods in Engineering* **30**(4), 2793–2810 (2023)
9. McMahan, H.B., Moore, E., Ramage, D., Hampson, S., y Arcas, B.A.: Communication-efficient learning of deep networks from decentralized data. In: *International Conference on Artificial Intelligence and Statistics* (2016), <https://api.semanticscholar.org/CorpusID:14955348>
10. Mendonça, T., Celebi, M., Mendonca, T., Marques, J.: Ph2: A public database for the analysis of dermoscopic images. *Dermoscopy image analysis* **2** (2015)
11. Müller, D., Soto-Rey, I., Kramer, F.: Towards a guideline for evaluation metrics in medical image segmentation. *BMC Research Notes* **15**(1), 210 (2022)
12. Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., et al.: Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems* **32** (2019)

13. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Medical image computing and computer-assisted intervention—MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III 18. pp. 234–241. Springer (2015)
14. Tschandl, P., Rosendahl, C., Kittler, H.: The ham10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific Data* **5**(1), 180161 (Aug 2018). <https://doi.org/10.1038/sdata.2018.161>, <https://doi.org/10.1038/sdata.2018.161>
15. Xin, C., Liu, Z., Ma, Y., Wang, D., Zhang, J., Li, L., Zhou, Q., Xu, S., Zhang, Y.: Transformer guided self-adaptive network for multi-scale skin lesion image segmentation. *Comput. Biol. Med.* **169**(C) (Apr 2024). <https://doi.org/10.1016/j.compbiomed.2023.107846>, <https://doi.org/10.1016/j.compbiomed.2023.107846>
16. Yanagisawa, Y., Shido, K., Kojima, K., Yamasaki, K.: Convolutional neural network-based skin image segmentation model to improve classification of skin diseases in conventional and non-standardized picture images. *Journal of Dermatological Science* **109**(1), 30–36 (2023). <https://doi.org/https://doi.org/10.1016/j.jdermsci.2023.01.005>, <https://www.sciencedirect.com/science/article/pii/S0923181123000051>
17. Yang, Q., Liu, Y., Cheng, Y., Kang, Y., Chen, T., Yu, H.: Federated learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning* (2019), <https://api.semanticscholar.org/CorpusID:233176472>