131 Privacy Preserved Federated Learning for Diabetic Retinopathy Detection

Abstract

Diabetic Retinopathy (DR), a complication of diabetes, could result in different stages of vision loss. Early detection of this disease is critically important due to the high chances of blindness. Retinal images are basically taken and stored by the medical institutions, so the data is not always accessible due to the privacy of patients. Early detection of DR in a privacy preserving approach has led us to explore the concept of Federated Learning with its two popular method, called Federated Averaging (FedAVG) and Federated Prox (FedProx). Using four publicly available datasets containing different number of samples, we have developed our non-FL approach based on AlexNet architecture. FedAVG and FedProx also use the same machine learning model for their training. Collected results from our three models demonstrated that FedProx outperforms FedAVG in all cases.

Introduction

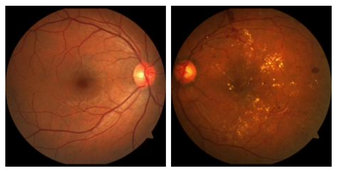
The number of people with diabetes have been increasing exponentially in the recent years. It is even predicted to rise more than 552 million by 2030. Diabetic Retinopathy (DR) is a complication of diabetes caused by blood sugar. Early detection of DR could extremely reduce the number of patients who lose vision due to this disease. Applying machine learning techniques has been a solution to better assist healthcare clinicians. However, accessing such data containing patients' information has been always a bottleneck for researchers. To develop an ML model for detecting DR, Federated Learning (FL) enables training without accessing the data. It basically distributes a global model to the datacenters/clients and receives the updated local model (which is trained on the data) from all the clients. The process could be iterated until the global model presents good results.

Research Question(s)

1. What is the role of federated learning in detection of Diabetic Retinopathy?
2. Explore how federated learning could train without seeing the data by preserving privacy.
3. Explore two popular FL techniques (FedAVG and FedProx) and provide directions for future adoption.

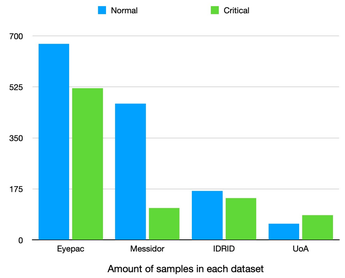
Materials and Methods

- The proposed research considers four main datasets containing retinal images to detect Diabetic Retinopathy. The datasets are including Eyepac, Messidor, IDRID, and the last one is collected by the University of Auckland (UoA) [1-4].
- Feature selection and preprocessing
- Using a non-FL model (AlexNet architecture) and two FL models (FedAVG and FedProx [5,6])

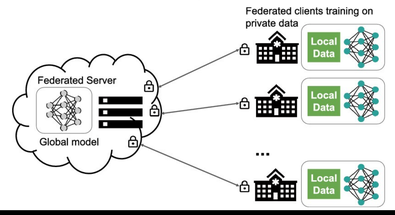


Normal (class "0")

Critical (class "4")

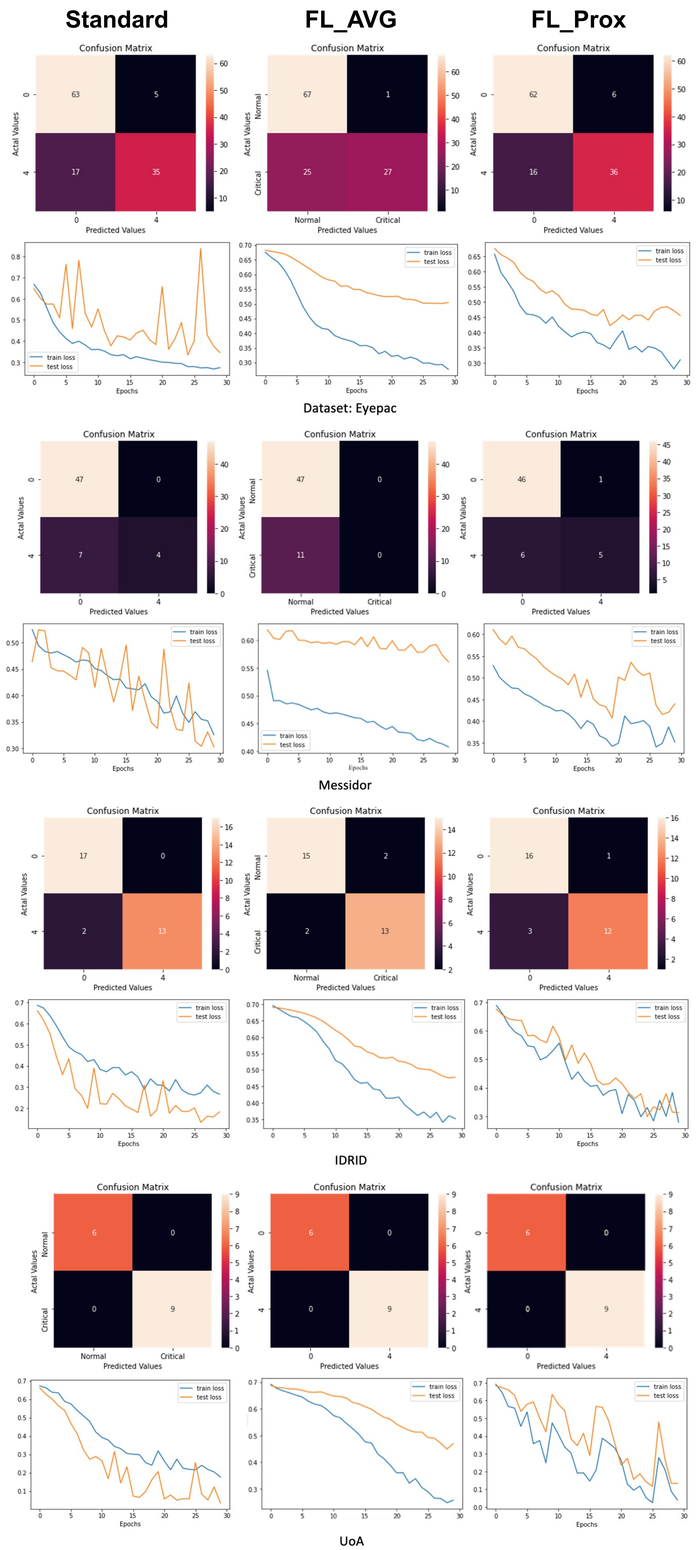


Amount of samples in each dataset



Concept of Federated Learning

Results

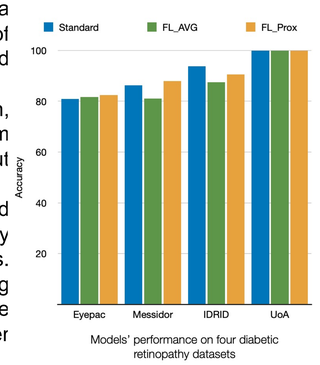


Conclusions

Federated Learning could be deployed as a privacy preserving approach in detection of DR. It enhances the privacy of data compared to the traditional machine learning.

Federated learning uses an averaging function, which averages the weights collected from each client to build a global model without accessing the data.

The basis behind the two deployed federated learning approaches are the same. The only difference is in their aggregation strategies. FedProx achieves better results when facing heterogenous data. As we can see in the figure, FedProx has demonstrated better results.



Models' performance on four diabetic retinopathy datasets

Acknowledgments

Special thanks to Dr. Karakaya and Dr. Pouriyeh for their expert advice and support through this research.

The results will be submitted to a suitable journal.

Contact Information

Author: Mohammad Nasajpour, mnasajp1@students.Kennesaw,edu
Supervisor: Mahmut Karakaya, mkarakay@kennesaw.edu
Supervisor: Seyedamin Pouriyeh, spouriye@kennesaw.edu

References

[1] https://www.eyepacs.com/data-analysis
[2] https://www.adcis.net/en/third-party/messidor/
[3] Prasanna Porwal, Samiksha Pachade, Ravi Kamble, Manesh Kokare, Girish Deshmukh, Vivek Saha srabuddhe, Fabrice Meriaudeau, April 24, 2018, "Indian Diabetic Retinopathy Image Dataset (IDRiD)", IEEE Dataport, doi: https://dx.doi.org/10.21227/H25W98.
[4] Abdulla, Waleed, and Renoh Johnson Chalakkal. "UoA-DR Database Info." (2018).
[5] McMahan, Brendan, et al. "Communication-efficient learning of deep networks from decentralized data." *Artificial intelligence and statistics*. PMLR, 2017.
[6] Li, Tian, et al. "Federated optimization in heterogeneous networks." *arXiv preprint arXiv:1812.0612* (2018).
[7] Roth, Holger R., et al. "Federated learning for breast density classification: A real-world implementation." *Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning*. Springer, Cham, 2020. 181-191.

KENNESAW STATE UNIVERSITY
COLLEGE OF COMPUTING AND SOFTWARE ENGINEERING

Author: Mohammad Nasajpour
Advisors(s): Dr. Mahmut Karakaya, Dr. Seyedamin Pouriyeh