Distributed Learning in Healthcare: Application of Federated Learning to Skin Cancer Diagnosis

Yeke Zhang[a]
Automation, Northeastern University, Shenyang, China

Keywords: Federated Learning, Skin Cancer Diagnosis, Health System.

Abstract: Skin cancer is one of the deadliest diseases, but some of its types can be treated and cured if diagnosed in early stages. Machine learning (ML) is important for accurate skin cancer detection. Using ML to train a module is definitely helpful to image recognition and cancer prediction. However, patients' relevant data is private and sensitive. It is illegal for module trainers to transport raw data directly. To solve this problem, federated learning (FL) provides a grand new approach to construct an accurate but private skin cancer detection system. This paper introduces the theory how FL is applied to skin cancer diagnosis and reviews the development of FL and skin cancer detection in recent years. The paper mainly focuses on certain outstanding applications, especially those have been proven effective and better than traditional method. Besides, through discussion on limitations and challenges of FL in this field, the paper explores the future direction of research. The aim is to highlight the potential of FL in the skin cancer diagnosis and application to future health system.

1 INTRODUCTION

Machine Learning (ML) has become a very important method in many fields over the past decade (Liu, 2021; Liu, 2023; Qiu, 2020; Qiu, 2022). Based on large-scale training databases, machine learning algorithms learn from the data and are able to classify, make prediction and so on. With the rapid development of ML, a lot of neural networks are proposed such as AlexNet (Yuan, 2016), VGG (Tammina, 2019), GoogLeNet (Yu, 2022). They perform well in vision tasks. And other ML algorithm are also used into production. For example, Support Vector Machine (SVM) in rice prediction (Liakos, 2018). Machine learning is applied into various realms and seen outstanding result.

Medical image classification plays a significant role in cancer detection and risk prediction. In recent years, machine learning has been applied into this field and make a good contribution. It improves the process and help doctors to recognize whether it is cancer through learning past cancer images and data. By this way, ML increases speed and accuracy of diagnosis. For instance, melanoma is an extremely deadly skin cancer that may cause 10 million fatalities in 2020 according to global statistics. However, due to combination of ML and medicine, melanoma diagnostic accuracy increased from 50% to 75% (Yaqoob, 2023).

However, traditional ML algorithms have a common feature, they all update and transform through a central node, which is usually the administrator's computer. It is available to finish the machine learning task, but it may undermine the patient's right to privacy. It also may create a data breach (Hameed, 2021). The medical data contain patients' biological information and family privacy, their leak may cause harmful influence to patients. In the decades, privacy laws are getting stricter, ML algorithm that is applied into the medical field should consider more about privacy and traditional ML algorithm is not enough in this regard.

In this case, federated learning can be considered since it mainly focuses on solving privacy problem. It provides a platform that allows clients to train model on their own computer and only give a completed model about their own data. During this way, users' information avoids leaking and the central administrator can collect all necessary information to train and update model. Using federated learning

[a] https://orcid.org/0009-0006-4885-6889

Zhang, Y.
Distributed Learning in Healthcare: Application of Federated Learning to Skin Cancer Diagnosis.
DOI: 10.5220/0012967700004508

In *Proceedings of the 1st International Conference on Engineering Management, Information Technology and Intelligence (EMITI 2024)*, pages 675-678
ISBN: 978-989-758-713-9

algorithm, model trainer can train model based on limited information so that they need not to worry about privacy law because the raw information is not sent (Chowdhury, 2021).

Many researches have been done to develop such methods. For example, Xu et al. adopted federated learning to achieve fairness in dermatological disease diagnosis (Xu, 2022). Wicaksana et al. combined federated learning with image classification to better work (Jeffry, 2022). What's more, one research develops a possibility to cooperate between different hospitals through federated learning (Hosseini, 10).

The author will focus on follow points: 1) The development of cancer prediction application of federated learning. 2) The theory how federated learning are adopted into cancer detection. 3) The limitation and challenge of this technical.

2 METHOD

2.1 Introduction of Federated Learning

Federated learning is a process that is mainly based on the distributed learning. At first, the administrator distributes the old model to the nodes, which are the independent clients. Then, each client uses this model to process the data and calculates the bias, usually gradient. By neural network, the algorithm can aggregate the weights and bias. The next step is returning the processed data instead of clients' raw data. Finally, the administrator receives the processed data and update the model. It is a round of machine learning. After many rounds, the model will be more and more accurate and practical.

Federated learning also has many optimizations. Based on these optimization methods, optimized federated learning performs better on a certain target. For example, some researchers use offline calculation, which changes modulo power operations to modulo multiplication operations and speed up computing. If module constructors focus on sparse data, they can adopt sparse matrix calculation and sparse bar chart optimization. During data transportation, many encrypt algorithms can be applied to protect the data security or reduce the scale of transported data. With the development of federated learning, more and more new methods are invented to optimize federated learning to adapt to certain task.

2.2 Federated Learning-Based Skin Cancer Detection

Federated learning has demonstrated its significant impact in the field of skin cancer diagnostics through numerous successful applications. By integrating with a variety of techniques, federated learning enhances the diagnostic process for skin cancer.

Haggenm et al. develop a method for melanoma that is a kind of serious skin cancer diagnostics by decentralized federated learning. They pretrained models on ImageNet. Then they used the tree-structured Parzen estimator30 to choose the hyperparameters to maximize the area under the receiver operating characteristic curve (AUROC). During the process, they increased the learning rate at the former part and decreased the learning rate at the latter part. And the federated data come from five independent hospitals. All data are protected. Finally, researchers found that decentralized method performed better than traditional federated learning method. It had lower AUROC on the test dataset. Compared with traditional algorithm, it is a good alternative (Hekler, 2024).

Ain et al. also succeeded to develop a method to predict skin cancer. They combined many private hospitals to support their project and got data. Every private hospital is a client, so they train their own patients' data on their own computer using Support Vector Machine (SVM) and Convolutional Neural Networks (CNN). And they use certain algorithm to transform the data into weight that can be communicated among different hospitals. At the same time, nodes transport their data to the central server. At last, central sever can update the model to predict the skin cancer. The system also uses median filtering method, watershed algorithm, gray Level co-occurrence matrix (GLCM) feature technique and ABCD rule. Researchers used this kind of method to test dataset and got a high accuracy. It performs a certain level of robustness and fitness. By now, it is one of the most effective methods as a classifier or predictor (Ain, 2023).

Based on federated learning, a new method is developed. An asynchronous and weighted approach is used to help federated learning for skin lesion diagnosis. Yaqoob et al. adopted FedOpt approach to reduce communication overhead. They creatively develop an asynchronous technique to update shallow and deep lays at different rates. Using this method, doctors can better find skin diseases. What's more, communication costs are reduced so it is really economical.

FedPerl is another good example that federated learning is used to improve skin classification. This method improves its model through anonymous peers. By this kind of data transportation, patients' data will be safe. Bdair et al. firstly built communications, then focused on peer learning and peer anonymization. Finally, they got a good result. The test shows that their model has a better parameter and lower communication cost. This means a grand new way to classify skin disease and diagnose the most serious cancers (Bdair 2021).

Hashmani et al. developed a method to apply federated learning into intelligent dermoscopy device. Traditional algorithms are usually based on visual pattern recognition of morphological features, so this new method provides a general solution to different images. Researchers use federated learning to support intelligent dermoscopy device and attain better performances and universal application (Hashmani, 2021). With the help of this technique, doctors can better recognize different skin legion, which will definitely contribute to the skin cancer detection because skin cancer usually signals beneath the surface of skin.

3 DISCUSSIONS

Federated learning has achieved a great deal in the skin cancer field, it completes the task to train a practical model without transforming raw data directly and protects the clients' privacy. However, there are still some limitations, especially interpretability and poor performance compared to traditional visual methods.

Interpretability: Due to privacy laws, different clients share different types of data. It means that individual clients can benefit from the collaborative training only if their data is compatible with that of other participating institutions (Roschewitz, 2021). For instance, different collected different patients' data because of privacy laws, but not all the data is suitable to be used to train model. Besides, federated learning may cause client data is not interoperable. The interpretability makes federated learning face many challenges in front of the skin cancer treatment and diagnosis. Although iFedAvg improved the phenomenon to a large extent, it still haves non-negligible false-positive rate.

Poor performance compared to ABCDE rule: ABCDE rule (Duarte, 2021) is a traditional rule to distinguish between benign and malignant lesions. ABCDE rule examines skin cancer according to color, lesion scale and so on. It has been proven trusty and fundamental. Although method based on federated learning is effective, it still shows more prediction mistakes than traditional ABCDE rule. The reason mainly is pattern recognition complexity for malignant lesion characteristics in medical imaging (Riaz, 2023). Because of this, the method is still thought of as a black-box method. In practice, doctors prefer to choose ABCDE rule to visually examine the skin cancer.

To optimize federated learning in skin cancer detection and treatment, follow aspects may be considered: increasing data balance, improving data interpretability and optimizing strategies in different stages of the disease. Above aspects directly linked to the application of the algorithm, breakthroughs in these areas would contribute significantly to development of combination of federated learning and skin cancer.

Although federated learning has many challenges, it has potential for researchers to explore. There is still a lot of room for optimization between algorithms and practical applications. In addition, the hardware situation and transmission mechanisms should be also improved to combine with federated learning algorithms well (Deng, 2019; Deng, 2023; Sugaya, 2019). Much research is ongoing to improve federated learning in learning rate, privacy protection, data heterogeneity and resource cost. Many more algorithms should be combined with federated learning to solve practical problems. For example, FedDecorr are created to solve dimensional collapse in the learning process (Shi, 2023). FedPerl was created to make skin cancer lesion classification more accurately. In short, further investigations are required to develop federated learning and skin cancer.

4 CONCLUSIONS

This paper overviews recent developments of federated learning for skin cancer, and discusses the applications of federated learning in kinds of skin cancer treatments or diagnosis. Federated learning has a successful application, which offers a considerable method to improve skin cancer treatment system and doesn't against the privacy laws. Many algorithms and machines use federated learning to improve their performance and protect clients' privacy. However, using federated learning to improve skin cancer prediction or treatment will also have some challenges such as more time spending and poor interpretability. Besides, prediction accuracy and model complexity also need to improve. So, aiming to have an adequate understating of application of federated learning for skin cancer, further research and investigation are needed to find out potential benefits and challenges.

REFERENCES

Ain, Q. U., Khan, M. A., Yaqoob, M. M., Khattak, U. F., Sajid, Z., Khan, M. I., & Al-Rasheed, A. 2023. Privacy-Aware Collaborative Learning for Skin Cancer Prediction. Diagnostics, 13(13), 2264.

Bdair, T., Navab, N., & Albarqouni, S. 2021. FedPerl: semi-supervised peer learning for skin lesion classification. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 336-346). Cham: Springer International Publishing.

Chowdhury, A., Kassem, H., Padoy, N., Umeton, R., & Karargyris, A. 2021. A review of medical federated learning: Applications in oncology and cancer research. In International MICCAI Brainlesion Workshop (pp. 3-24). Cham: Springer International Publishing.

Deng, X., Oda, S., Kawano, Y. 2023. Graphene-based midinfrared photodetector with bull's eye plasmonic antenna. Optical Engineering, 62(9), 097102-097102.

Duarte, A. F., Sousa-Pinto, B., Azevedo, L. F., Barros, A. M., Puig, S., Malvehy, J., ... & Correia, O. 2021. Clinical ABCDE rule for early melanoma detection. European Journal of Dermatology, 31(6), 771-778.

Hameed, S. S., Hassan, W. H., Latiff, L. A., & Ghabban, F. 2021. A systematic review of security and privacy issues in the internet of medical things; the role of machine learning approaches. PeerJ Computer Science, 7, e414.

Hashmani, Manzoor Ahmed, et al. 2021. An adaptive federated machine learning-based intelligent system for skin disease detection: A step toward an intelligent dermoscopy device. Applied Sciences, 11(5), 2145.

Hekler, A., Maron, R. C., Haggenmüller, S., Schmitt, M., Wies, C., Utikal, J. S., ... & Brinker, T. J. 2024. Using multiple real-world dermoscopic photographs of one lesion improves melanoma classification via deep learning. Journal of the American Academy of Dermatology.

Hosseini, S. M., Sikaroudi, M., Babaie, M., & Tizhoosh, H. R. 2023. Proportionally fair hospital collaborations in federated learning of histopathology images. IEEE transactions on medical imaging.

Liakos, K. G., Busato, P., Moshou, D., Pearson, S., & Bochtis, D. 2018. Machine learning in agriculture: A review. Sensors, 18(8), 2674.

Liu, Y., Liu, L., Yang, L., Hao, L. and Bao, Y. 2021. Measuring distance using ultra-wideband radio technology enhanced by extreme gradient boosting decision tree (XGBoost). Automation in Construction, 126, 103678.

Liu, Y. and Bao, Y. 2023. Intelligent monitoring of spatially-distributed cracks using distributed fiber optic sensors assisted by deep learning. Measurement, 220, 113418.

Qiu, Y., Wang, J., Jin, Z., Chen, H., Zhang, M., & Guo, L. 2022. Pose-guided matching based on deep learning for assessing quality of action on rehabilitation training. Biomedical Signal Processing and Control, 72, 103323.

Qiu, Y., Yang, Y., Lin, Z., Chen, P., Luo, Y., & Huang, W. 2020. Improved denoising autoencoder for maritime image denoising and semantic segmentation of USV. China Communications, 17(3), 46-57.

Riaz, S., Naeem, A., Malik, H., Naqvi, R. A., & Loh, W. K. 2023. Federated and Transfer Learning Methods for the Classification of Melanoma and Nonmelanoma Skin Cancers: A Prospective Study. Sensors, 23(20), 8457.

Roschewitz, David, et al. 2021. Ifedavg: Interpretable data-interoperability for federated learning. arXiv preprint arXiv:2107.06580.

Shi, Yujun, et al. 2023. Understanding and mitigating dimensional collapse in federated learning. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Sugaya, T., Deng, X. 2019. Resonant frequency tuning of terahertz plasmonic structures based on solid immersion method. 2019 44th International Conference on Infrared, Millimeter, and Terahertz Waves, 1-2.

Tammina, S. 2019. Transfer learning using vgg-16 with deep convolutional neural network for classifying images. International Journal of Scientific and Research Publications (IJSRP), 9(10), 143-150.

Wicaksana, J., Yan, Z., Yang, X., Liu, Y., Fan, L., & Cheng, K. T. 2022. Customized federated learning for multi-source decentralized medical image classification. IEEE Journal of Biomedical and Health Informatics, 26(11), 5596-5607.

Xu, G., Wu, Y., Hu, J., & Shi, Y. 2022. Achieving fairness in dermatological disease diagnosis through automatic weight adjusting federated learning and personalization. arXiv preprint arXiv:2208.11187.

Yaqoob, M. M., Alsulami, M., Khan, M. A., Alsadie, D., Saudagar, A. K. J., AlKhathami, M., & Khattak, U. F. 2023. Symmetry in privacy-based healthcare: a review of skin cancer detection and classification using federated learning. Symmetry, 15(7), 1369.

Yu, Z., Dong, Y., Cheng, J., Sun, M., & Su, F. 2022. Research on face recognition classification based on improved GoogleNet. Security and Communication Networks, 2022, 1-6.

Yuan, Z. W., & Zhang, J. 2016. Feature extraction and image retrieval based on AlexNet. In Eighth International Conference on Digital Image Processing (ICDIP 2016) (Vol. 10033, pp. 65-69). SPIE.

Zhao, F., Yu, F., Trull, T., & Shang, Y. 2023. A new method using LLMs for keypoints generation in qualitative data analysis. In 2023 IEEE Conference on Artificial Intelligence (CAI) (pp. 333-334). IEEE.