Endoscopic Image Classification using Federated Learning

Abstract

The emergence of IoT, AI, Machine Learning, and Deep Learning algorithms has completely changed the healthcare sector and given rise to data-driven medical applications. However, due to difficulties in achieving rigorous secu- rity, privacy, and quality of service standards, the adoption of AI-based medical applications is still low. With the processing of medical data at the network's edge, federated learn- ing has come to be seen as a possible solu- tion to these problems because it allows for the distributed training of sophisticated machine- learning models while maintaining the privacy and security concerns. In this paper, we have applied some federated learning based algorithms for endoscopic image classification. We used an open-source KVASIR dataset to conduct our research.

Keywords: Federated Learning, Endoscopy, Image Classification, KVASIR

1 Introduction

Gastrointestinal (GI) infections and cancer is a major public health concern because it is frequently identified at an advanced stage when treatment is difficult. Gastrointestinal cancer treatment includes surgery, targeted therapy and chemotherapy but the effectiveness of these treatments depends on the cancer stage at diagnosis [1]. Early detection and prevention through regular screening are crucial to improving outcomes and reducing the burden of gastrointestinal cancer.

The text states that the United States has the highest number of cases of gastric cancer and a large number of people suffer from bowel infections, with new cases reported every year [2][3]. Due to the rise of patients with GI problems, the demand for gastroenterologists is increasing Endoscopy is a diagnostic imaging technique that is particularly effective in identifying various GI tract abnormalities. Many studies have been conducted to develop automated classification algorithms for detecting and diagnosing GI tract illnesses using endoscopic pictures.

Artificial intelligence (AI) is used to detect abnormalities in medical images. The system performs several steps to analyze the images, including pre-processing (cleaning and enhancing the images), identifying important characteristics in the images, choosing the most relevant features, and classification (determining if the image shows a normal or abnormal result). This system can help with early detection of abnormalities, which can lead to earlier treatment and better outcomes for patients.

Research was also carried out in order to build automatic classification schemes that are capable of identifying and treating disorders in the gastrointestinal (GI) tract using gastrointestinal images [4]. By performing feature mining and selection, and classification on accessible medical pictures, these AI-based systems have been shown to be effective in the early diagnosis of problems. Nevertheless, owing to the existence of numerous noises such as annotations, black borders, and other abnormalities, pre-processing of endoscopic pictures might be a difficult task [5].

These noises can affect the accuracy of the classification model and lead to false positives or negatives, thereby compromising the system's reliability. Therefore, developing effective pre-processing techniques that can effectively remove these noises and enhance the quality of endoscopic images is crucial for improving the accuracy and reliability of automated classification systems for GI abnormalities.

This research paper aims to investigate the use of federated learning in improving the accuracy and interpretability of GI abnormality detection and classification. We propose a novel the framework that uses federated learning to train a Deep learning model on GI endoscopy images from multiple sources while preserving patient privacy. We also propose using XAI techniques to interpret the model's decision-making process and provide insights into the features used to classify GI abnormalities. Our suggested approach, we believe, has the chance of enhancing the accuracy and privacy of GI abnormality detection and classification while also providing interpretability, thus enabling clinicians to make more informed decisions and improving patient outcomes.

2 Related Works

The author here has proposed a diagnostic decision-support system that utilizes machine learning for the semi-automatic evaluation of in-vivo gastral images acquired via capsule endoscopy. To assess the utility of the proposed technique, a quantitative analysis was conducted. They have developed a pipeline for the combined categorization and interpretation of capsule endoscopy image frames. To provide bleeding diagnosis, convolutional neural networks have been used for training and verifying the image data collection. To increase the trustworthiness of the black box predictions, LIME-based explanation abilities have been added to the machine learning-based solution.

The feasibility of using variably private federated learning in the medical field for evaluating complicated histopathology pictures has been shown in this paper. The Cancer Genome Atlas (TCGA) dataset was used to model a distributed setting. The dataset has also been used to contrast the effectiveness of private, dispersed training with that of standard training. Differential privacy increases the degree of privacy security by imposing numerical constraints. They have shown the effectiveness of a method known as federated learning (FedAvg) using synthetic real-world data with independent and non-independent data distributions. Private federated learning is a viable option for distributed training on medical data because it achieves equivalent outcomes to traditional centralized training.

In recent years, deep learning has had astonishingly good effects on medical diagnosis. In this paper[3] a new approach is proposed for classifying endoscopic images that combine the Inception-Resnet-v2, VGG16, ResNet-50, MobileNetV2, and ResNet-152 models with a Grad-CAM model designed for explainable artificial intelligence. Additionally, they have applied a data augmentation technique to boost the effectiveness of medical images. They have also used a noise reduction technique to solve the overfitting issue while utilizing a short dataset. An open-source KVASIR dataset with more than 8,000 photos of the esophagus, stomach, and colon is included in the dataset, along with annotations that show the presence of different disorders such as ulcers, inflammation, and polyps has been used here. In this paper[3], they applied various widely used CNN models, including ResNet-152, VGG16, ResNet-18, DenseNet201, MobileNetv2 and for the classification of wireless endoscopic images. ResNet-152 performed the endoscopic image classification with the highest level of accuracy. Grad-Cam produced improved heat map visualization outcomes for explainable AI.

This paper [4] suggests a method for automatically classifying Wireless Capsule Endoscopic photos in order to find the bleedy images. A model named Bleedy Image Recognizer(BIR) which is a combination of MobileNet and a custom-defined CNN architecture model, has been used in this paper, which is a deep learning approach. Feature extraction is done using MobileNet, while CNN is used for training. The dataset includes 4550 normal pictures and 450 bleedy images from 33 patients, respectively. The parameters of accuracy, precision, recall, F1 score, and Cohen's kappa are used to evaluate the BIR's efficacy. The obtained values of 0.993, 1.000, 0.994, 0.997, and 0.995 for Cohen's kappa, F1 score, accuracy, precision, recall, and F1 score, respectively, demonstrate successful results. The BIR model is also analyzed using an additional dataset from Google and achieves accuracy of 0.978.

The researcher [5] proposes a framework to classify the endoscopic abnormal images. For classification in their paper, they used EWT and CNN networks. The KVASIR dataset is used and divided into two sections. The main motive of their work is to detect different types of diseases in the digestive system. The image recognition process is performed at two classification levels and the proposed CNN model is trained twice. A classification model's performance is assessed using the confusion matrix. Framework efficiency is assessed using a confusion matrix and various performance indicators.

This research suggests a brand-new algorithm for intestinal centerline extraction. The work in the paper includes- creating a federated learning framework and a data learning mechanism while guaranteeing the confidentiality of the data sets, introducing the blockchain mechanism to improve data interaction and eliminate the federated learning server's single point of failure, suggesting ways to improve the current centerline extraction framework and implement centerline extraction by concentrating on the edges and adding sources. To achieve complete segmentation of the intestine, their system develops a deep learning module and suggests a network structure strengthened by multiscale fusion. By having medical professionals name the intestine's area and the centerline based on endoscopic pictures, a total of 403 sets of intestinal CT data were acquired. In the future, XAi can be used to provide clear and understandable explanations for their judgments and actions, especially in high-stakes situations when the consequences of making an incorrect decision could be severe.

3 Methodology

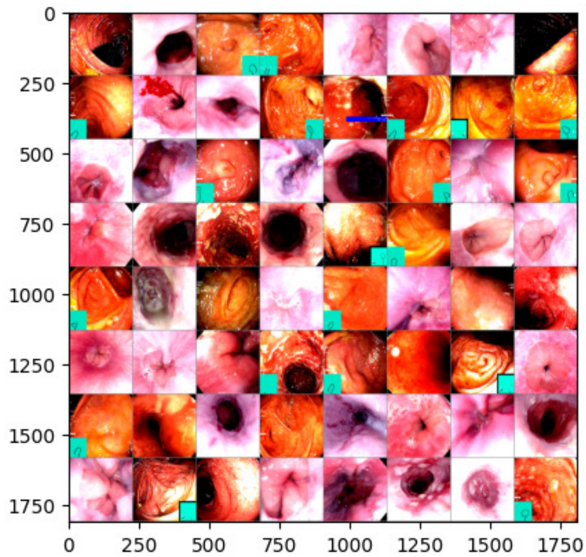

Fig. 1 KVASIR dataset sample images

3.1 Dataset

Table 1 Result after applying Federated CNN

Number Of Epoch	Average Clients Loss	Average Clients Accuracy(%)	Server Testing Accuracy(%)
1	0.9857	39.16	31.14
2	0.8128	48.52	43.72
3	0.7367	54.41	48.39
4	0.6848	58.43	51.56
5	0.6442	62.41	53.93
6	0.608	65.16	55.42
7	0.576	68.38	56.71
8	0.546	72.71	57.96
9	0.517	74.92	58.77
10	0.489	78.76	59.62

The 'KVASIR' Dataset [6] was made available as part of MediaEval's medical multimedia challenge. The collection is made up of 8,000 endoscopic pictures taken during an endoscopy operation from the GI tract. Medical specialists interpret and verify the photos, which are divided into eight classifications based on three anatomical landmarks, three pathological findings, and two more classes linked to the polyp removal procedure. We utilized four dataset classes in our paper: 'ulcerative-colitis', 'normal-cecum', 'normal-z-line', and 'esophagitis'. There were a total of 2000 images in these four classes. Here, Figure 1 shows the sample images of this dataset.

3.2 CNN

Convolutional neural networks are feed-forward neural networks that are commonly used to evaluate visual pictures by processing data in a grid-like fashion. A convolution neural network comprises several hidden layers that aid in image information extraction. The layers are the Convolution layer, ReLU layer, Pooling layer, and Fully connected layer[7]. Typically, the first layer removes fundamental characteristics like horizontal or diagonal edges. The result is sent to the next layer, which recognizes additional characteristics like corners or combinational edges. when we get further into the network, it can recognize greater complex characteristics like objects, faces, and so forth[8].

3.3 Federated Learning

Federated learning is a distributed method of training machine learning models that do not need the sharing of the data that underlies them. Algorithms are distributed to several data warehouses for local training. When trained, just the algorithm, not the data, returns to the central location. The updated predictions are then delivered to each individual dataset to keep and enhance[9]. Federated learning involves each client privately training a copy of the central model, which is represented by the model weights ω, and reporting its modifications back to the server for aggregation across clients without exposing local private data. Federated learning may be expressed as:

$$\min_{\omega \in R} f(\omega) \quad \text{with} \quad f(\omega) = \frac{1}{n} \sum_{i=1}^{n} fi(\omega)$$

4 Result

The Training loss vs. Number of epochs plot of basic CNN is shown in Figure 3. The figure shows that as the number of epochs increases, the training loss decreases. Where Fig. 2 shows the Training accuracy vs. the Number of Epochs plot of a non-federated CNN. It is a rising graph because accuracy rises with increasing epochs. From the graph, we can see that, after 10 epoch, the accuracy is 56%. Both the Training loss vs. Number of epoch plots and the Training accuracy vs. Number of epoch plots for Federated CNN is shown in Fig. 4. According to the graph for Training loss vs Number of epochs, the testing loss is greater than 2.0 at the start but decreases to about 1.0 after 10 epochs, which is quite good. In the Training Accuracy vs. Number of Epochs graph, the accuracy rises to nearly 60%.

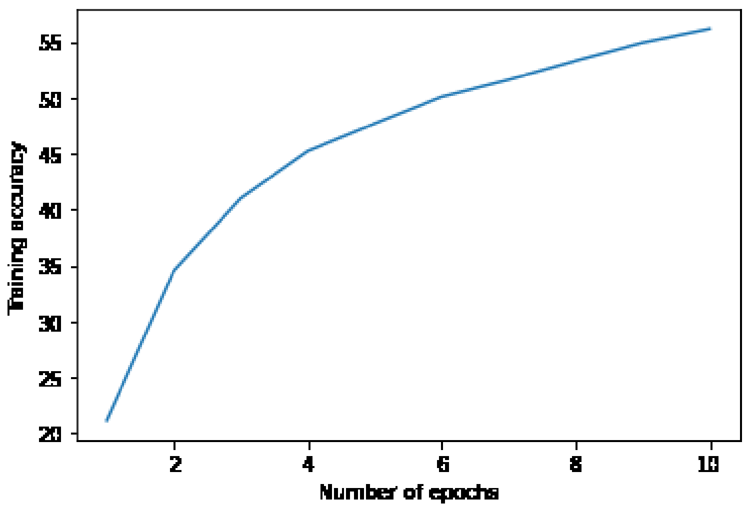

Fig. 2 Training accuracy vs. Number of Epochs plot of a non- federated CNN

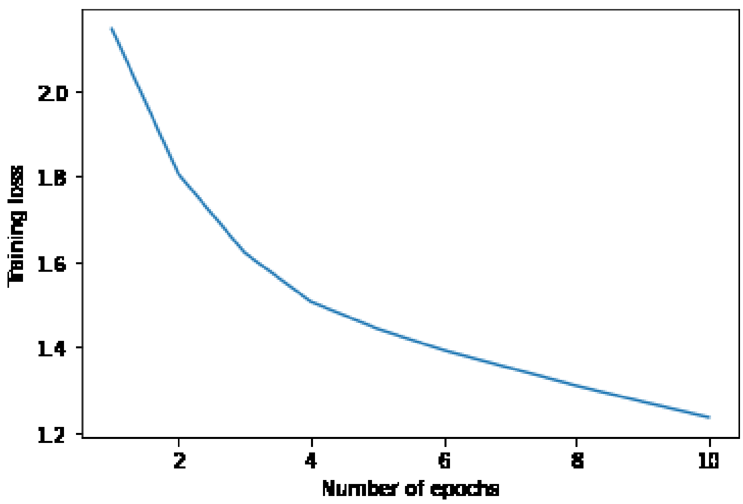

Fig. 3 Training loss vs Number of epoch plots of basic CNN.

Table 1 shows Epoch No., Average Clients Loss, Average Clients Accuracy, Server Testing Accuracy of the federated CNN model. From the table, we can see that after 1st epoch the Average Clients Loss is 0.98, and the Average Clients Accuracy: is 39.16% and the Server Testing Accuracy: is 31.14%. In the 2nd epoch, the Average Clients Loss becomes 0.81, and the Average Clients Accuracy becomes 48.52% and the Server Testing Accuracy becomes 43.72%. So after only one epoch, the Average Clients Loss reduces up to 0.17, and the Average Clients Accuracy increases by 9.36% and the Server Test- Fig. 4: Training loss vs Number of epoch plots and Training

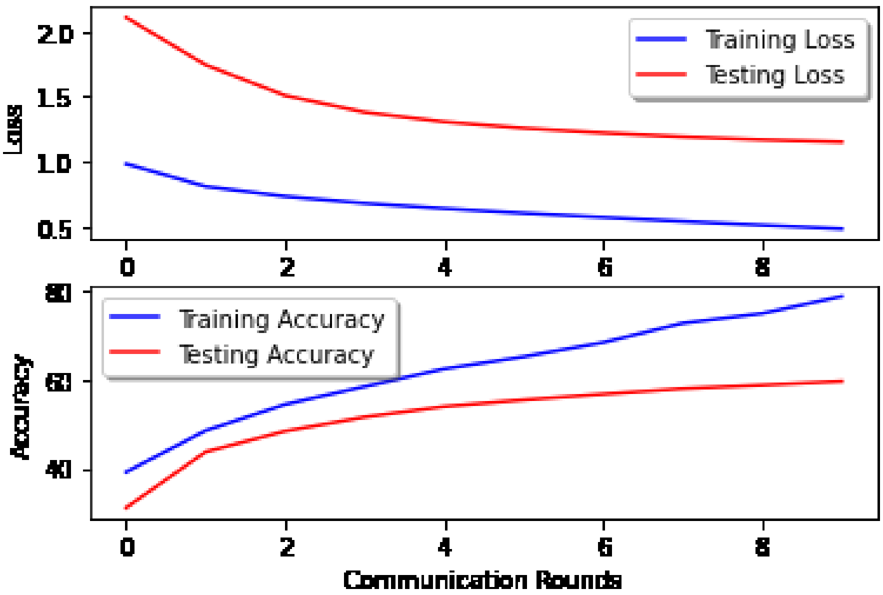

Fig. 4 Training loss vs Number of epoch plot and Training accuracy vs Number of epoch plots for Federated CNN

accuracy vs Number of epoch plots for Federated CNN. ing Accuracy increases by 12.58%. Overall these results are really impressive. However after the 10th epoch, the Aver- age Clients Loss becomes 0.48, Average Clients Accuracy becomes 78.76% and the Server Testing Accuracy becomes 59.62%. The data table shows that after the 10th epoch, the average server accuracy nearly doubles. This result is also superior to the basic CNN, which has a test accuracy of 56%. So, using this federated version of CNN not only provides us with better accuracy but also with data security and access.

5 Conclusion

As IoT, AI, machine learning, and deep learning algorithms proliferated and gave rise to data-driven medical applications, the healthcare sector experienced a radical transformation. Due to difficulties in achieving strong security, privacy, and service quality standards, however, the adoption of AI-based medical applications are still low. With the processing of medical data at the network's edge, federated learning has come to be seen as a possible answer to these problems because it allows for the distributed training of sophisticated machine-learned models while maintaining privacy and security concerns. In this paper, we have used federated learning-based algorithms for endoscopic image classification with an open-source KVASIR dataset.Many different tests, experiments, and adaptations have been left for the future due to a lack of time. Future work concerns the deeper analysis of particular mechanisms, new proposals to try different methods, or simply curiosity. In the future, XAI can be used to provide clear and understandable explanations for their judgments and actions, especially in high-stakes situations when the consequences of making an incorrect decision could be severe.

References

[1] Mohapatra, S., Pati, G.K., Mishra, M., Swarnkar, T.: Gastrointestinal abnormality detection and classification using empirical wavelet transform and deep convolutional neural network from endoscopic images. Ain Shams Engineering Journal **14**(4), 101942 (2023)

[2] Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A., Jemal, A.: Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians **68**(6), 394–424 (2018)

[3] Sharif, M., Attique Khan, M., Rashid, M., Yasmin, M., Afza, F., Tanik, U.J.: Deep cnn and geometric features-based gastrointestinal tract diseases detection and classification from wireless capsule endoscopy images. Journal of Experimental & Theoretical Artificial Intelligence **33**(4), 577–599 (2021)

[4] Khan, M.A., Kadry, S., Alhaisoni, M., Nam, Y., Zhang, Y., Rajinikanth, V., Sarfraz, M.S.: Computer-aided gastrointestinal diseases analysis from wireless capsule endoscopy: a framework of best features selection. IEEE Access **8**, 132850–132859 (2020)

[5] Cogan, T., Cogan, M., Tamil, L.: Mapgi: Accurate identification of anatomical landmarks and diseased tissue in gastrointestinal tract using deep learning. Computers in biology and medicine **111**, 103351 (2019)

[6] Kvasir Dataset — kaggle.com. https://www.kaggle.com/datasets/meetnagadia/kvasir-dataset. [Accessed 06-May-2023]

[7] Biswal, A.: Convolutional Neural Network tutorial [update]. Simplilearn (2023). https://www.simplilearn.com/tutorials/deep-learning-tutorial/convolutional-neural-network

[8] Mandal, M.: Introduction to convolutional neural networks (CNN) (2022). https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/

[9] What is federated learning? https://owkin.com/what-is-federated-learning