Detecting Colon Diseases using Privacy Preserving Federated Learning

*Note: Sub-titles are not captured in Xplore and should not be used

1st Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

2nd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

3rd Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

4th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

5th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

6th Given Name Surname
dept. name of organization (of Aff.)
name of organization (of Aff.)
City, Country
email address or ORCID

***Abstract*—In Gastroenterology, accurately classifying diseases using endoscopic images is a considerable challenge. Conventional gastrointestinal disease detection is time consuming, although accurate. To prevent and control further development of the disease, which may turn to cancer in future, faster detection is necessary to start treatment as soon as possible. This research offers a method for aiding medical diagnosis procedures and diagnosing problems of the gastrointestinal system based on the categorization of characteristics extracted from endoscopic images using a federated learning model. In machine learning approaches, privacy has always been a concern as user data is exposed to machine learning models and central servers, which might cause a security threat in future. Here comes the role of federated learning which collects data from user and trains a local model in the user device. Then the training information is sent to the machine learning model to further train on this data and improve accuracy. This encrypted data approach solves the issue of privacy. Another issue is use of excessive power from user device which is solved by training local models of user device when the device is idle or charging. In this study, we propose a 95.63% accurate federated learning method for detecting gastrointestinal illnesses from wireless capsule endoscopy (WCE) images of the colon.**

***Index Terms*—federated learning, colon cancer, machine learning**

I. Introduction

The gastrointestinal (GI) tract, often known as the digestive system, is susceptible to a variety of disorders, including polyps, ulcer, colorectal cancer, etc. [1]. Common symptoms include discomfort or pain, belly distension, appetite loss, nausea and vomiting, abdominal pain, and weariness. Some GI illnesses frequently result in GI cancer, which is the second most prevalent cancer worldwide [2]. Muco-submucosal polyps, which are the result of chronic prolapse of the mucosa in the intestine, are a prevalent illness of the gastro-intestinal tract. [3]. Early-stage polyps are frequently asymptomatic, but as they grow, they can obstruct the entrance to the small intestine. Symptoms of polyps may include blood in the stool, resulting in anemia, stomach discomfort when touched, and nausea. On endoscopic imaging, these lesions show as polypoid masses and carry an elevated risk of malignancy. Esophagitis is a frequent gastrointestinal illness characterized by inflammation of the tube linking the esophagus and stomach. Esophagitis is characterized by swallowing difficulty, chest pain, heartburn, and food becoming lodged in the esophagus [4]. Typically, endoscopy reveals rings of diseased tissue. Ulcerative colitis, an inflammatory bowel disease, is a common disorder that causes inflammation of the GI tract, abdominal pain, diarrhea, tiredness, and bloody stool. These GI illnesses have symptoms that frequently overlap, making diagnosis challenging. Initial detection of these disorders may result in a cure or the prevention of deadly malignancy. Although visual evaluation of endoscopic pictures provides a first diagnosis, it is frequently time-consuming and subjective [5]. Furthermore, radiologist shortcomings and other human factors can frequently result in false positive or even false negative diagnoses, which can be harmful to the patient [6]. Thus, a computer-assisted diagnostic would be advantageous for early detection with high accuracy. In this paper, endoscopic pictures of patients with gastrointestinal illnesses are classified. For the categorization job, we adopted two distinct strategies. For classification, we employed federated learning learning to detect colon diseases. Our data set on gastrointestinal illnesses consists of four classes:

- Healthy control, or normal class
- Ulcerative colitis

Identify applicable funding agency here. If none, delete this.

- Polyps and
- Esophagitis

II. Methodology

In this paper, we have proposed a novel framework to detect gastrointestinal diseases from wireless capsule endoscopy (WCE) curated images with federated learning model.

A. Dataset Description

Our dataset was obtained from Kaggle. The dataset included WCE pictures of the digestive tract (GI). This dataset originally included 720 x 576-pixel images representing four classes: normal, ulcerative colitis, polyps, and esophagitis. Our machine learning algorithms were employed to classify this dataset into the aforementioned four categories.

B. Dataset Preprocessing

We have divided the dataset into three sets: training set, validation set, and test set, in order to train and evaluate our models based on a variety of quantitative performance evaluation measures. Scikit package was applied to divide the dataset into training, test, and validation sets. The dataset was divided into training and test sets in an 8:2 ratio. Additionally, the training set was divided into a training set and a validation set in a 9:1 ratio. Our photos have been downsized to 224x224x3. In addition, we have labeled the classes with numerical values for the classification models. The pixel values of images have been adjusted by dividing them by 255. To circumvent data limitations, we've enhanced our dataset images with the ImageDataGenerator function of the Keras library.

C. Training and Testing Model

We have developed a convolutional neural network model to classify colon images. Five convolutional blocks, each consisting of a convolutional layer plus a max pooling layer, were utilized to build the model. The kernel size of each convolutional layer was 3x3, the activation function was ReLu, and the stride was 1. The initial convolutional layer's input shape was (224,224,3). All the maximum pooling layers had a 2x2 pool size. To extract features, the filter numbers for first through fifth convolutional layers were 32, 32, 64, 64, and 128 consecutively.

III. Result Analysis

We have evaluated our models using a variety of quantitative indicators. To measure the performance of our models, we tested them on a set of data they had never seen before. The parameters employed to assess their performance are accuracy, precision, recall, and f1 score.

IV. Conclusion

In this paper, we have used federated learning to detect three gastrointestinal diseases: ulcerative colitis, polyps, and esophagitis, along with healthy colon images. The model scored an accuracy of 95.63%. We have faced resource utility and data limitation in conducting our work. We resolved data limitation by adopting augmentation approach. We aim to resolve our infrastructural limitation in near future. In future we plan to work on larger range of gastrointestinal diseases with improved accuracy of federated learning based approach. We aim to further ease the diagnosis with elastography technique using ultrasound.

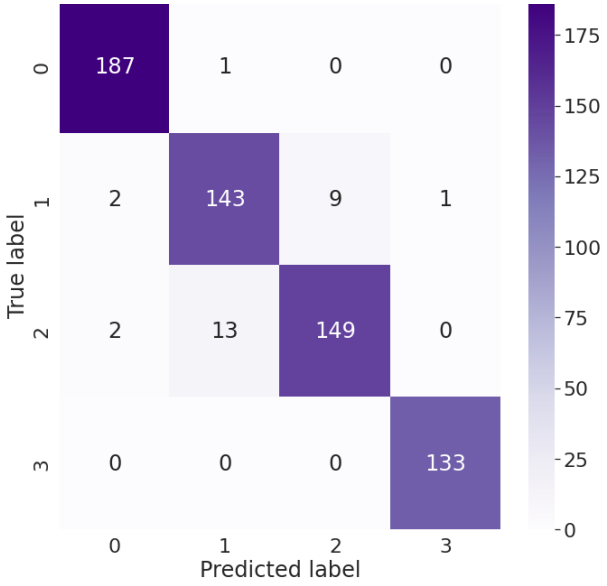

Fig. 1. Confusion Matrix

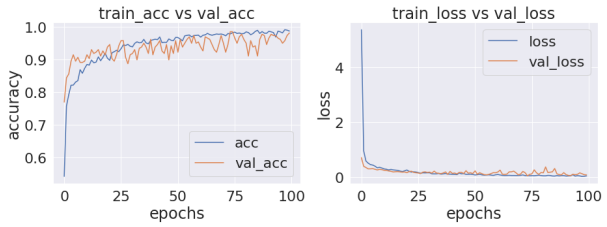

Fig. 2. Confusion Matrix

References

[1] M. B. Haile, A. O. Salau, B. Enyew, and A. J. Belay, "Detection and classification of gastrointestinal disease using convolutional neural network and svm," Jun 2022. [Online]. Available: http://dx.doi.org/10.1080/23311916.2022.2084878

[2] R. Sitarz, M. Skierucha, J. Mielko, J. Offerhaus, R. Maciejewski, and W. Polkowski, "Gastric cancer: epidemiology, prevention, classification, and treatment," p. 239–248, Feb 2018. [Online]. Available: http://dx.doi.org/10.2147/CMAR.S149619

[3] C. L. Tan, "Muco-submucosal elongated polyps of the gastrointestinal tract: A case series and a review of the literature," p. 1845, 2013. [Online]. Available: http://dx.doi.org/10.3748/wjg.v19.i11.1845

[4] P. Moayyedi, J. Santana, M. Khan, C. Preston, and C. Donnellan, "Medical treatments in the short term management of reflux oesophagitis," Apr 2007. [Online]. Available: http://dx.doi.org/10.1002/14651858.CD003244.pub2

[5] D. Jha, S. Ali, S. Hicks, V. Thambawita, H. Borgli, P. H. Smedsrud, T. de Lange, K. Pogorelov, X. Wang, P. Harzig, M.-T. Tran, W. Meng, T.-H. Hoang, D. Dias, T. H. Ko, T. Agrawal, O. Ostroukhova, Z. Khan, M. Atif Tahir, Y. Liu, Y. Chang, M. Kirkerød, D. Johansen, M. Lux, H. D. Johansen, M. A. Riegler, and P. Halvorsen, "A comprehensive analysis of classification methods in gastrointestinal endoscopy imaging,"

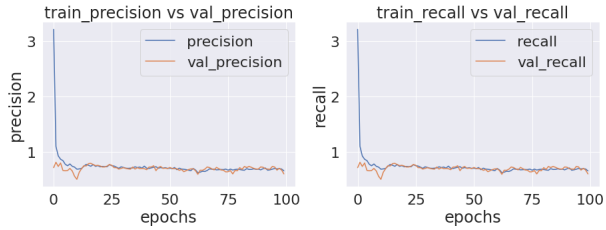

Fig. 3. Confusion Matrix

Medical Image Analysis, vol. 70, p. 102007, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361841521000530

[6] Öztürk and U. Özkaya, "Residual lstm layered cnn for classification of gastrointestinal tract diseases," p. 103638, Jan 2021. [Online]. Available: http://dx.doi.org/10.1016/j.jbi.2020.103638