International Journal of Medical Toxicology & Legal Medicine Volume 27, No. 4S, 2024
e-ISSN: 0974-4614
p-ISSN: 0972-0448
https://doi.org/10.47059/ijmtlm/V27I4S/085

FedCervical: Shedding Light on Cervical Cancer Detection with Federated Learning and Explainable AI

Chinmay Kadam[1], Kyle Dsouza[2], Ayush Sharma[3], Prof. Prachi Patil[4], Prof. Monali Shetty[5]

[1,2,3]Students, Department of Computer Engineering, Fr. Conceicao Rodrigues College of Engineering, Mumbai
[4,5]Assistant Professor, Department of Computer Engineering, Fr. Conceicao Rodrigues College of Engineering, Mumbai
Email: chinmaykad172@gmail.com[1], kyle.dsouza.official@gmail.com[2], crce.9294.cs@gmail.com[3], prachi@fragnel.edu.in[4], monali_shetty@fragnel.edu.in[5]
*Corresponding Author

Received: 18.09.2024 Revised: 17.10.2024 Accepted: 14.11.2024

ABSTRACT
Cervical cancer, posing as a global health challenge, demands precise detection methods. This research brings in an innovative approach to identify cervical cancer, utilizing advanced machine learning on pap smear data. This research utilizes Federated Learning (FL) and Explainable AI (XAI) to improve cervical cancer detection accuracy while ensuring patient's privacy. After studying the SIPaKMeD dataset, we have developed a FL model capable of classifying cervical cell images into five distinct categories. The decentralized nature of FL ensures robust data privacy, allowing multiple institutions to collaboratively train a model without sharing sensitive patient information. The FedAvg Model implemented using the Flower framework achieves a 95.12% accuracy and paves a way for further usage of such technologies in the field of medical sciences.
Acknowledging the sensitivity of health information, Federated Learning (FL) emerges as a crucial tool, ensuring privacy in data handling. Pap smears, integral to screening, occasionally pose challenges in early detection. FL, through decentralized models, optimally manages extensive datasets, enhancing diagnostic accuracy. In an era of prioritizing data privacy, FL becomes a safeguard, allowing collaborative model training without compromising individual confidentiality. Beyond security, FL democratizes cervical cancer diagnostics, offering patients nuanced insights into their results. This research signifies a substantial step in advancing cervical cancer detection, marrying optimal model performance with stringent privacy protection, resonating with evolving healthcare paradigms and individual empowerment in the diagnostic process. Also, the amalgamation of Explainable AI (LIME) with FL leads to a solution which can be used by onco- pathologists around to world to successfully diagnose cancerous cells.

Keywords: Cervical Cancer, Machine Learning, Federated Learning, Explainable Artificial Intelligence, Pap Smear, Diagnostic Methodology, Privacy Protection, Healthcare Decentralization, Data Security, Medical Informatics.

1. INTRODUCTION

1.1 Background

With 604,000 new instances of cervical carcinoma in 2020, It is one of the most common cancers in women around the world, along with breast, colon, and lung cancer [1]. However, if the patient is diagnosed at the early stage of lesions or even before, the incidence of the disease may be considerably reduced. Cervical screening has been based on Papanicolaou's experiments for over 60 years. In the 1940s, George Papanicolaou, a Greek physician, employed the Papanicolaou test, also called the Pap smear test. It includes cervical exfoliation cells that can be examined under a microscope to track cancer or its antecedents.
Roughly 90 per cent of the 342,000 cervical cancer-related deaths took place in low- and middle-income nations. South- East Asia, Central America, and sub-Saharan Africa (SSA) have the greatest incidence and fatality rates of cervical cancer. Disparities in access to immunization, screening, and treatment services are linked to regional variations in the incidence of cervical cancer.
The novel method presented in this paper makes use of advanced computer vision algorithms to assist onco-pathologists in identifying cancerous cells in Pap smear test images. Using federated learning techniques, this methodology aimed to create a single global model capable of identifying benign cells. This paradigm has been

carefully selected in an inter- hospital collaborative effort to protect patient privacy and confidentiality. This research endeavors to enhance the field of diagnostic pathology by combining the power of computer vision technologies with federated learning paradigms, thereby increasing the precision and efficacy of cervical cytology malignancy diagnosis.

1.2 Motivation

The increasing prevalence of cervical cancer worldwide highlights the need for advanced detection techniques that make use of the recent explosion in technology. Applications of computer vision and machine learning have been popular in the medical field; nevertheless, the growing volume of training data presents computing difficulties. As data complexity increases, intelligent solutions are needed to maximize processing effectiveness. Patient confidentiality is inherently in danger from the traditional method of centralizing data for classification since it leaves room for potential exploitation and unwanted access. To address these issues, this research project recommends the use of Federated Learning, a decentralized paradigm that facilitates cooperative model training over dispersed datasets and avoids the exposure of private patient data. The pressing need to address the challenges faced by both technicians and doctors in diagnostics and with the aim of empowering them, motivates us in finding a solution. The amalgamation of machine learning insights from Pap smear tests with Federated Learning is what this studied is focused upon.

The rest of the paper goes as follows, Section 2 talks about the review of the related work. Section 3 details about our proposed methodology and the results are presented in Section 4. Section 5 highlights the limitations in our solution while section 6 concludes the paper with directions for future work.

2. LITERATURE REVIEW

In [2], a ground-breaking Federated learning algorithm which utilizes data from the Danish Gynecological Cancer Database (DGCD) and other European cancer registries is given to us by the authors. The analysis of the patient's data was achieved with the use of a novel decentralized approach, which eliminated the need to share private sensitive patient data

Table 1: Summary of Research Papers on Federated Learning and Cervical Cancer Detection.

Paper Title	Findings	Research Summary	Citations
Using federated learning to identify women with early-stage cervical cancer at low risk of lymph node metastasis	This paper utilizes DGCD/SQRGC model and SQRGC/NCR model to get moderate diagnostic ac- curacy, with AUCs ranging between 0.78–0.80	Only discusses theory behind FL and delves into what exactly a model detects	[2]
Flower: A Friendly Federated Learning Research Framework	Researchers can federate existing ML models and transition from large-scale simulation to execution on heterogeneous devices	Discusses reducing the disparity between FL research and real-world FL systems	[3]
Computer-Assisted Screening for Cervical Cancer Using Digital Image Processing of Pap Smear Images	This paper discusses computer-assisted screening for cervical cancer	Accuracies found are not significant	[4]
Hybrid model for detection of cervical cancer using causal analysis and machine learning techniques	This research utilizes SVM, random forest, and regression models to predict cervical cancer risk and type with an accuracy of 68.4%, 79.8%, and 81.2%	Future research elaborates the need for integrating emerging technologies for faster, more efficient computation	[5]
A model for predicting cervical cancer using machine learning algorithms	This paper utilizes decision tree, logistic regression, support vector machine, K-nearest neighbors algorithms	Although this paper deals with ML, it does not provide an interface healthcare professionals	[6]
A Federated Learning Framework for Breast Cancer Histopathological Image Classification	This research shows a federation application for cancer classification	Discusses potential implications of using a highly balanced dataset	[7]

The study's conclusions suggest that 3,962 women as a group met the well-defined selection criteria that they described in their academic paper. The most common type of cancer to present at diagnosis was squamous cell carcinoma (64%), followed by adenocarcinoma (32%), and Aden squamous carcinoma (4%). 46 years old was the average age upon diagnosis (±12 standard deviation). Remarkably, 467 people, or 12% of the study sample, had pN1 presentations. These minute nuances are fully documented inside the boundaries of their scholarly discussion. In addition, by consolidating the results of five classifiers—LD (linear discriminant), SVM (support vector machine), KNN (k-nearest neighbor), boosted trees, and bagged trees — [2] improves the final outputs. The SIPaKMeD and Herlev datasets were used to demonstrate the efficiency of the suggested framework. The SIPaKMeD dataset gave a 98.27% accuracy in two-class classification and 94.0% accuracy in five-class classification, based on experimental results. In two-class and five-class situations, comparison of the suggested technique with five classifiers led to much better results.

In [4] the study suggested methodological approach is based on segmenting pre-processed images using K-means clustering, which is applied to the Herlev dataset. The segmented nuclei are used to extract various shape properties. A Random Forest Classifier is then used to classify the nuclei according to these shape features. The Random Forest Classifier's performance is then compared to that of other classifiers in a thorough examination, and each classifier's results on the Herlev dataset are methodically investigated. This method adds to the body of knowledge regarding automated cervical cancer screening by providing a clear and consistent methodology and evaluating its performance against other classification methods.

In [8], the authors delve into the challenges of analyzing the Pap smear images by the onco-pathologist and conclude that there are certain restrictions like a time-consuming process, staining image, poor contract images and wrong interpretation by a human which makes this process unreliable. Pathologists can benefit greatly from an automated system that divides and classifies cervical cancer to accurately predict the cancer's stages and choose the best course of treatment. The principal aim of the methodology suggested in this work is to reduce false-negative results through accurate segmentation and classification. The method's effectiveness is especially remarkable in situations where there are several cells arranged in overlapping cellular structures. According to extensive survey evaluations, Support Vector Machine (SVM) is the only algorithm used for classification due to its superior performance. This study adds to the body of knowledge in the field by empowering us with an improved methodology aimed at enhancing the accuracy of cervical cancer diagnosis by automated segmentation and classification.

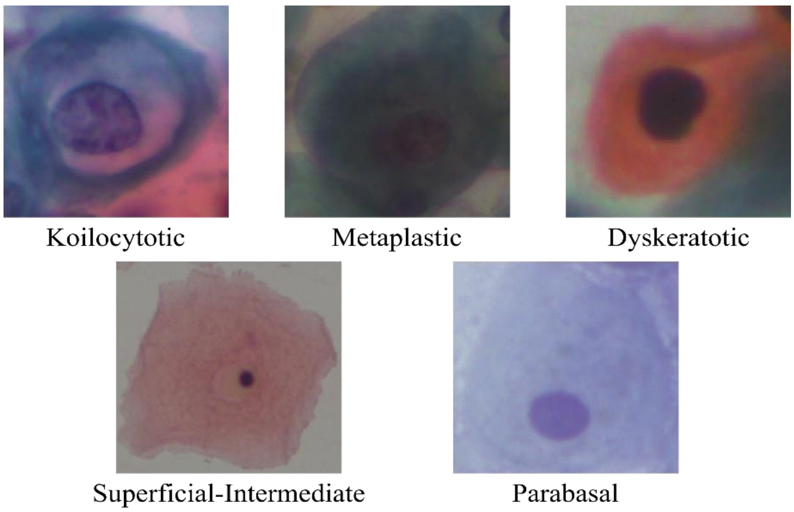

Figure 1: The five classes of nuclei categorized

The dataset employed in this study is categorized into three distinct classes: 1) the Normal class, comprising cells with normal characteristics, 2) CIN1 class, denoting cells exhibiting Cervical Intraepithelial Neoplasia Grade 1, and 3) CIN2/CIN3 class, encompassing cells indicative of Cervical Intraepithelial Neoplasia Grades 2 and 3. Post feature extraction, the dataset undergoes a three-fold classification, with the first class representing normal cells, the second class assigned to CIN1 cells, and the third class dedicated to CIN2/CIN3 cells. The experimental outcomes reveal a commendable accuracy rate of 96%, with a specificity of 100% and a sensitivity of 95.6%.

3. METHODOLOGY

3.1 Dataset

The dataset utilized for this research is sourced from the SIPaKMeD Database [9], a critical repository for the investigation of cervical cancer cell detection. The SIPaKMeD Database consists of 4049 meticulously curated

images of isolated cells which have been edited carefully from 966 cluster cell images extracted from Pap smear slides. A CDD camera which has been adapted to an optical microscope was used to capture these images, guaranteeing high resolution and an all-inclusive analysis of the dataset.
The cell images within the dataset are categorized into five distinctive classes, namely Dyskeratotic, Koilocytotic, Meta- plastic, Parabasal, and Superficial-Intermediate, representing key categories in the spectrum of cervical cell abnormalities as depicted in Figure 1. Each class encapsulates crucial information critical for understanding the differences of cervical cancer progression and leading to accurate diagnostic capabilities.
The categories of the dataset and class labels provide a refined perspective on cell abnormalities, giving us a detailed analysis of cervical cancer at various stages. This comprehensive dataset, captured through advanced imaging techniques, serves as the linchpin for this research, allowing for a focused exploration of cervical cancer cell detection methodologies and insights into the intricacies of the disease.

3.2 System Architecture

Our system architecture seamlessly progresses through a streamlined sequence of steps. Commencing with the input of a cervical cancer cell image from the SIPaKMeD dataset, the system initiates preprocessing to ensure standardized and optimized data for analysis. Subsequently, the Federated Learning (FL) framework, preserving data privacy, distributes model updates to local devices for individualized training. This decentralized model training incorporates insights from diverse datasets without compromising sensitive health information. Following local training, model aggregation forms a global model, refined through secure mechanisms. The final model, now collectively informed, classifies cervical cancer cells with enhanced accuracy. The decision of including Explainable Artificial Intelligence (XAI) techniques further helps the model's decision-making process, providing insights into combination of features. Healthcare professionals are provided the final output consisting of the model's prediction and XAI-generated images which lead to informed and effective diagnostic and decision-making experience.

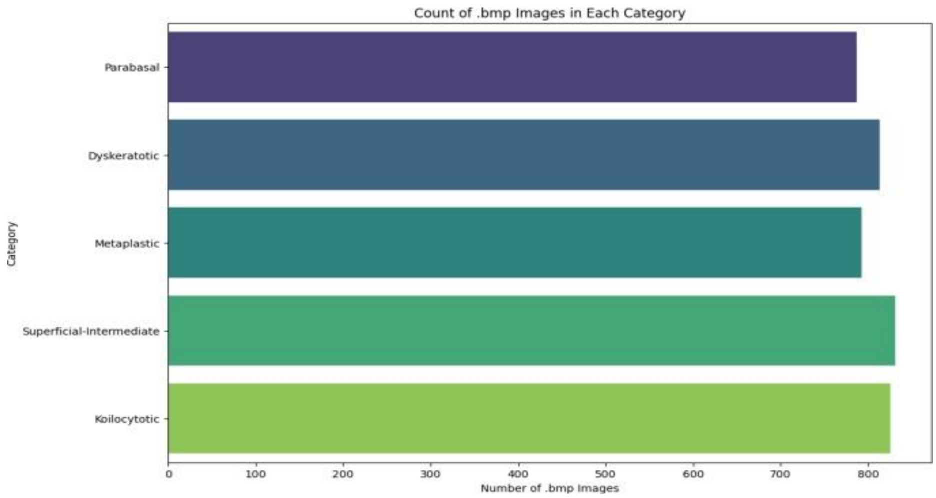

Figure 2: The five classes of nuclei categorized

3.3 Preprocessing

The dataset exhibited a nearly equal distribution of images across the five classes, as shown in Figure [2]. However, to maintain optimal model impartiality, a deliberate decision was made to undertake under-sampling, limiting the dataset to 700 images per class. This strategy aimed at mitigating any potential bias that might arise from imbalances in class representation.
Following the under-sampling process, the dataset underwent meticulous preprocessing. A process made up of transformations, including conversion to gray scale and normalizing dimensions to 224x224 pixels was applied on all the images which were loaded from the dataset folder. A well-structured dataset made up of features vectors from flattened images was created to conduct in depth analysis and make up a robust model. The overall effectiveness in detection of cervical cancer was enhanced by equipping the model to identify subtle patterns which are associated with various classes by careful preprocessing and training over a balanced and refined dataset.

3.4 Model Selection

A rigorous exploration of machine learning models to identify which models work best in categorizing the images into the five classes was conducted. Given the intricate nature of the features present in the dataset, emphasis was given to the usage of deep learning techniques. Specifically, VGG16, InceptionNet, Resnet50 and

MobileNet techniques were used. VGG16 is distinguished by its consistent usage of 3x3 convolutional filters and max-pooling layers. It has 16 weight layers, consisting of 13 convolutional and 3 fully connected layers. Thanks to its parameter-sharing method and deep, homogenous structure, it performs exceptionally well in picture classification tasks. InceptionNet, also known as GoogleNet, is a deep convolutional neural network architecture developed by Google. Notable for its inception modules, which use filters of multiple sizes within the same layer, InceptionNet aims to capture diverse scale features efficiently. Convolutional neural network (CNN) architecture ResNet-50 is well-known for its depth and effectiveness in deep neural network training. InResNet-50, the ”50” stands for depth, meaning that there are 50 weight layers. For image classification tasks, ResNet-50 has been widely used and is renowned for producing cutting-edge results. MobileNet is a convolutional neural network architecture that is lightweight and optimized for effective deployment on devices with limited resources, like smartphones. With depth wise separable convolutions, MobileNet, which was unveiled by Google in 2017, lowers computational complexity without sacrificing performance. The dataset was split into training and testing set for the training phase which lasted 3 epochs for each model. At the end of each model, the training and validation metrics were noted. An in-depth analysis of each model's performance allowing insights into their generalization abilities was made available

3.5 Federated Learning

In the field of machine learning, federated learning is a sophisticated paradigm in which several decentralized nodes cooperatively participate in the training of a global model. This joint effort results in the construction of a model that represents the mean knowledge extracted from each node's training process. One important aspect of federated learning is data privacy preservation, which is an important consideration when working with sensitive datasets like medical data. In contrast to traditional methods, federated learning reduces the need for raw data to be transferred between devices, eliminating the inherent hazards related to data exposure. Federated Learning has a dual benefit: it provides increased privacy protection as well as significant computational re-source savings. After the necessary preprocessing steps, the SipakMed dataset appears to be quite large—roughly 9 gigabytes. For individual hospitals, training on datasets of this size is computationally demanding and impracticable. The implementation of federated learning strategically avoids these obstacles by enabling effective training procedures and the creation of a single, powerful global model. This revolutionary method protects data privacy while optimizing computing efficiency. The complex dynamics of federated learning, illustrated in the Fig. 4, describe how model training is coordinated globally and explain how local agents communicate with and update the global model automatically. Next, there was a need to choose a framework which will be used to implement the intended federated learning architecture. For this, firstly it was needed to understand the needs of the ecosystem where this architecture will be implemented i.e. hospitals and have a comparison between the various FL frameworks available and choose the best amongst them.

3.5.1 Federated Learning Framework

Several frameworks are available which streamline the process of implementation and deployment of Federated Learning. They vary in terms of their features, performance**,** and their shortcomings. Out of all of them, the two major frameworks under consideration were Flower and TensorFlow.

- **Tensor Flow Federated:** The team and Google created this open-source framework which allows developers to create distributed ML models while making use of the data originating from various sources in different forms and provide a high-level API (Application Programming Interface)
- **Flower:** On virtue of its adaptable and extendable architecture, Flower comes forward as a great choice for creation FL systems. TensorFlow, PyTorch, and Keras are some of the systems with which Flower is compatible with. Its intuitive interface simplifies setup and deployment. Support for several machine learning frameworks highlights its adaptability and lets users make the most of their favorite libraries.

3.5.2 Comparisons between different Federated learning frameworks

Several frameworks are available which streamline the process of implementation and deployment of Federated Learning. Flower and TensorFlow were the two major frameworks under consideration.

- **Tensor Flow Federated:** TFF does not have a user-friendly interface which Flower has, providing an edge to the later which requires less programming compared to the former.
- **Flower:** Flower stands out as an extremely flexible and extensible architecture, on the grounds that it is compatible with frameworks like TensorFlow, PyTorch, and Keras. Its intuitive interface is what simplifies the setup and deployment. Support for several machine learning frameworks highlights its adaptability and lets users make the most of their favorite libraries.

3.6 Flower

Federated Learning requires a platform which would be able to facilitate communication between the clients and server, sending the models back and forth, train and evaluating local data and sending the updated models.

Flower frameworks leverage a distributed computing paradigm that is made up of global and local computation components, two essential elements. Coordination of the federated learning process amongst the involved clients is done by the server-side global computation. Parameter update aggregation, configuration, and client selection, under the direction of a predetermined strategy, come under this. Local computation comprises real training tasks that utilize each client's distinct data, as well as other pertinent tasks that occur on the client side.
The client architecture is very simple and is defined by the way it receives communication from the server. The messages received by the clients are reacted upon by it by invoking user provided training and assessment features . The algorithm being utilized in the Flower framework for this application can be summarized as depicted in Algorithm 1.

Algorithm 1 Federated Learning Algorithm

procedure FederatedLearning(D, n, E, η)
  Divide the dataset D into n disjoint subsets $D_1, D_2, \ldots, D_n$
  Initialize a global model θ_0
  for $t = 1, 2, \ldots, E$ **do**
    $\theta_t \leftarrow \theta_{t-1}$
    for $i = 1, 2, \ldots, n$ **in parallel do**
      $\theta_i \leftarrow$ ClientUpdate(θ_t, D_i, η)
    end for
    $\theta_t \leftarrow$ FedAvg($\theta_1, \theta_2, \ldots, \theta_n$)
  end for
  return θ_E
end procedure

3.7 XAI

Concerns about the reliability and authenticity of conclusions drawn from Machine Learning (ML) models have gained prominence, especially as this technology becomes more and more integrated into important real-world decision- making processes. Due to the covert activities of the hidden layers inside these models, a major difficulty is that consumers are unaware of how these models arrive at results. Explainable Artificial Intelligence (XAI) is used to address this opacity in decision-making processes. XAI algorithms play a crucial role in bringing transparency into the ML models' decision- making processes. This openness, made possible by XAI, creates a strong basis for legitimacy and trust regarding the model's results, allaying worries and boosting confidence in the application of ML technologies for situations involving complex decision-making [11].
The two central models of explanation provided by the XAI models are model-specific and model-agnostic methods. Explanations tailored to the unique features of the model are provided by certain techniques. It is crucial to understand that, although explanations derived from models are inextricably related to the specifics of the model, not all processes created for specific models rely on the properties of the model itself [12].
The two XAI models which were considered were LIME and Grad-CAM. The decision of model selection was based heavily upon the comprehensive study based upon [13]. TA model-agnostic approach was chosen for the application after weighing the pros and cons of each of the two XAI approaches. The integration of XAI into the solution not only enhances transparency and interpret-ability but also fosters trust and acceptance among healthcare professionals. Corroboration of model predictions, refinement of diagnostic decisions and tailoring treatment plants to individual patient needs can be performed by clinicians by leveraging XAI insights. Ultimately, XAI empowers clinicians with actionable insights and facilitates collaborative decision-making, thereby advancing the efficacy and reliability of our solution in clinical practice as show in Figure 3.

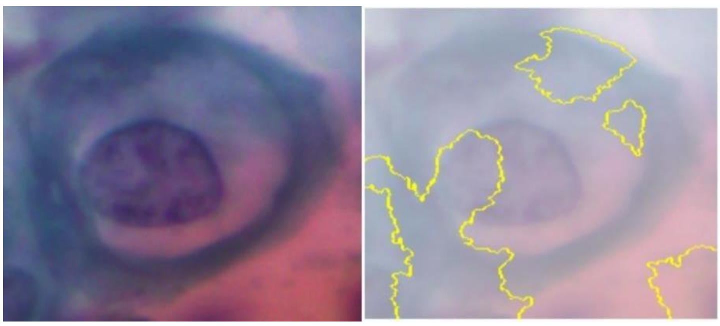

Figure 3: Federated Learning Solution Block Diagram

3.8 Federated Learning Setup

The usage of the Flower framework for the emulation of a federation environment was the goal of this research. Flower gave us with a seamless way to emulate the federated setting where training data remained decentralized across multiple sources. Our local datasets were kept private, while Flower facilitated the aggregation of model updates from each party to collaboratively train a global model. This federated approach enabled us to benefit from combined data while adhering to data privacy constraints. The Flower simulation demonstrated the capacity of federated learning to build robust machine learning models without direct sharing of sensitive medical data across institutions.

3.8.1 Flower Client initialization

We initialize the Flower Clients, each having their own train and test loaders along with a copy of the model on which they each must train. We implemented a custom FlowerClient class inheriting from NumPyClient to manage training and evaluation in a federated learning setup. The class initializes with the model, client ID, and data loaders. The get parameters method retrieves model weights. The fit method updates weights, trains for one epoch, and returns updated weights. The evaluate method computes loss and accuracy on validation data, returning loss, data length, and accuracy dictionary. This FlowerClient facilitates federated learning by enabling clients to train locally and exchange updates with a central server.

3.8.2 Flower Server implementation along with strategy declaration

We set the number of clients to 10 using NUM CLIENTS = 10. We then defined a FedAvg strategy using fl.server.strategy.Fe with specific parameters like fraction fit = 1.0 (all available clients participate in training), fraction evaluate = 0.5 (50% of clients participate in evaluation), min fit clients = 10 (minimum 10 clients for training), min evaluate clients = 5 (minimum 5 clients for evaluation), min available clients = 10 (minimum 10 clients should be avail- able), and evaluate metrics aggregation fn = weighted average (weighted average for aggregating evaluation metrics). We also checked if the device type is” cuda” (GPU), and if so, we set client resources to allocate one GPU per client. Finally, we started the federated learning simulation using fl.simulation.start simulation, passing the client function (client fn), number of clients (NUM CLIENTS), server configuration fl.server.ServerConfig with 5 rounds), the defined FedAvg strategy, and the client resources (if GPU is available).

4. Result

Table 2 illustrates the time to train (TTT) for each model during the training process along with the accuracy achieved during the validation phase of the model evaluation. This tabular representation makes it easier for us to understand and have a comparative analysis among the models.

After the experimentation, it was learnt that these models resulted in a maximum testing accuracy of 78% for VGG16 and a TTT value of around 1 hour 23 minutes,77% and 45 minutes for Inception v3. Resnet50 took 30 minutes to train and demonstrated an accuracy of 41 per cent during the validation phase. MobileNet achieved the highest accuracy of 91.8%, taking 40 minutes to train itself. We chose MobileNet as the primary model after conducting thorough comparisons with the rest of the models as yielded a testing accuracy of 94% which was proof enough of it being better at extracting underlying features. Traditional machine learning systems for cervical cancer detection often face significant flaws, particularly concerning privacy and communication overhead. In conventional centralized ML approaches, sensitive patient data must be aggregated into a central repository for model training. This aggregation raises serious privacy concerns as it increases the risk of unauthorized access and data breaches. Moreover, the communication overhead associated with sending very large volumes of sensitive medical data to a central server can be substantial, leading to latency issues and potential network congestion.

Table 2: Time To train for each model and their accuracies

Model	Time To Train (TTT)	Accuracy Achieved (%)
VGG16	1 hour 23 minutes	78
Inception V3	45 minutes	77
Resnet50	30 minutes	41
Custom	56 minutes	88.8
MobileNet	40 minutes	91.8
Flower	1 hour	95.12

The implementation of our cervical cancer detection solution has yielded remarkable results, particularly in the domain of federated learning. While the MobileNet machine learning model exhibited an impressive validation accuracy of 91%, the true breakthrough emerged from our exploration of federated learning techniques. By leveraging the Flower frameworks simulator, we successfully simulated a federated environment where training data remained decentralized across 20 healthcare institutions (clients). This innovative approach facilitated collaborative model training without the need to directly share sensitive patient data, addressing critical privacy and security concerns in 31 the healthcare sector. The success of our federated learning implementation, achieving an accuracy of 95.12%, represents a significant milestone in the realm of cervical cancer detection. This incredible performance surpasses that of the centralized MobileNet model, underscoring the efficacy of our approach.

By decentralizing training data across 20 healthcare institutions through the Flower framework's simulator, we have effectively addressed critical privacy and security concerns inherent in traditional centralized approaches. The enhancement of data privacy and cooperation among institution is promoted using this collaborative model training paradigm, while also leveraging various datasets for improving model robustness.

Limitations

The availability of a single database could limit the findings of this research along with limiting the number of clients which could be created while the process of Federation is underway. Nevertheless, Federated Learning can still be used in such situations as the global model keeps learning from each new additional image it processes, getting better with each round. This data (Pap smear slides) is sensitive data protected by laws like GDPR and HIPAA and should be handled with the proper care that it demands.

Also, if some participants (hospitals) have more data or higher-quality data, the model may become biased towards that hospital's contribution which could lead to poor generalization of the model. Latency could also be an issue for hospital with a limited bandwidth capacity.

XAI is a new and emerging technology which is untested in medical field. Medical practitioners may face difficulty in converting the insights provided by XAI into clinically meaningful interpretations. There is also the factor of trustworthiness, where the explanation provided could be wrong or misleading.

Conclusion and Future Scope

In summary, FedCervical has showcased the effectiveness of FL in cervical cancer detection, achieving superior accuracy while safeguarding patient privacy. The success of our federated learning implementation, achieving an accuracy of 95.12%, represents a significant milestone in the realm of cervical cancer detection. By decentralizing training data across various clients through the Flower framework's simulator, we have effectively addressed critical privacy and security concerns inherent in traditional centralized approaches. The enhancement of data privacy and cooperation among institution is promoted using this collaborative model training paradigm, while also leveraging various datasets for improving model robustness.

Moving forward, our results provide a blueprint for implementing federated learning in clinical practice, paving the way for further research and innovation in privacy-preserving healthcare solutions. By leveraging FedAvg and XAI, this project not only advances cervical cancer detection but also sets a precedent for ethical and transparent machine learning practices in healthcare.

Incorporating additional modalities such as clinical data, genomic information, and patient demographics to develop a comprehensive and holistic cervical cancer detection system. Integrating multimodal data sources can enhance the robustness and generalization of the model, giving more precise and personalized diagnostic predictions. Further, the conduction of an extensive study to evaluate the performance and efficiency of the detection solution in real-world healthcare settings is required.

Collaborating with healthcare institutions to deploy the model in clinical practice and assess its impact on patient outcomes, healthcare workflows, and diagnostic accuracy.

Lastly, investigating the use of advanced XAI techniques such as LIME to provide more informative and interpretable explanations for model predictions. In essence, the future of cervical cancer detection holds promising opportunities for advancements in multimodal data integration, and the adoption of advanced XAI techniques. Real-world deployment, ethical considerations, and collaboration with the healthcare community are paramount for ensuring the successful translation of research findings into impactful clinical solutions.

References

1. S. Zhang, H. Xu, L. Zhang, Y. Qiao, Cervical cancer: Epidemiology, risk factors and screening, Chinese Journal of Cancer Research 32 (6)(2020) 720.
2. H. H. Wenzel, A. N. Hardie, A. Moncada-Torres, C. K. Høgdall, R. L. Bekkers, H. Falconer, P. T. Jensen, H. W. Nijman, M. A. van der Aa,F. Martin, et al., A federated approach to identify women with early-stage cervical cancer at low risk of lymph node metastases, European Journal of Cancer 185 (2023) 61–68.
3. D. Beutel, T. Topal, A. Mathur, X. Qiu, T. Parcollet, N. Lane, Flower: A friendly federated learning research framework, arXiv preprint arXiv:2007.14390doi:10.48550/arXiv.2007.14390.
4. K. P. Win, Y. Kitjaidure, K. Hamamoto, T. Myo Aung, Computer-assisted screening for cervical cancer using digital image processing of pap smear images, Applied Sciences 10 (5) (2020) 1800.11

5. U. K. Lilhore, M. Poongodi, A. Kaur, S. Simaiya, A. D. Algarni, H. Elmannai, V. Vijayakumar, G. B. Tunze, M. Hamdi, Hybrid model for detection f cervical cancer using causal analysis and machine learning techniques, Computational and Mathematical Methods in Medicine 2022 (1) (2022) 4688327.arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1155/2022/4688327, doi:https://doi.org/10.1155/2022/4688327.URLhttps://onlinelibrary.wiley.com/doi/abs/10.1155/2022/4688327URL https://onlinelibrary.wiley.com/doi/abs/10.1155/2022/4688327
6. N. Al Mudawi, A. Alazeb, A model for predicting cervical cancer using machine learning algorithms, Sensors 22 (11). doi:10.3390/s22114132.URL https://www.mdpi.com/1424-8220/22/11/4132
7. L. Li, N. Xie, S. Yuan, A federated learning framework for breast cancer histopathological image classification, Electronics 11 (22 doi:10.3390/electronics11223767. URL https://www.mdpi.com/2079-9292/11/22/3767
8. H. Bandyopadhyay, M. Nasipuri, Segmentation of pap smear images for cervical cancer detection, in: 2020 IEEE Calcutta Conference (CALCON), IEEE, 2020, pp. 30–33.
9. M. E. Plissiti, P. Dimitrakopoulos, G. Sfikas, C. Nikou, O. Krikoni, A. Charchanti, Sipakmed: A new dataset for feature and image- based classification of normal and pathological cervical cells in pap smear images, in: 2018 25th IEEE International Conference on Image Processing (ICIP), 2018, pp. 3144–3148. doi:10.1109/ICIP.2018.8451588.
10. R. Hamsath Mohammed Khan, A comprehensive study on federated learning frameworks: Assessing performance, scalability, and benchmarking with deep learning model (2023).
11. A. Singhal, K. Agrawal, A. Quezada, A. Agui˜naga, S. Jim´enez, S. Yadav, Explainable artificial intelligence (xai) model for cancer image classification, Computer Modeling in Engineering Sciences 141 (2024) 1–10. doi:10.32604/cmes.2024.051363.
12. B. H. van der Velden, H. J. Kuijf, K. G. Gilhuijs, M. A. Viergever, Explainable artificial intelligence (xai) in deep learning-based medical image analysis, Medical Image Analysis 79 (2022) 102470. doi: https://doi.org/10.1016/j.media.2022.102470.
13. URL https://www.sciencedirect.com/science/article/pii/S1361841522001177
14. M. Arya, N. Mittal, G. Singh, Cervical cancer detection using segmentation on pap smear images, in: Proceedings of the International Conference on Informatics and Analytics, 2016, pp. 1–5.