IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⧖ Peer-reviewed & Refereed journal ⧖ Vol. 13, Issue 4, April 2024

DOI: 10.17148/IJARCCE.2024.134103

Applying Federated Learning For Breast Cancer Prediction

Shashank S[1], Shravan BB[2], Siddharth M Kalkur[3], Sriram M[4]

Student, Dept. of Computer Science Engineering, Dayanand Sagar University, Karnataka, India[1-4]

Abstract: This project aims Breast cancer remains a significant global health challenge, necessitating advancements in predictive modeling to enable early detection and personalized treatment. Traditional centralized machine learning approaches often face privacy concerns, especially when dealing with sensitive medical data. Federated Learning (FL) emerges as a promising solution, allowing model training across decentralized devices without sharing raw data. This project explores the implementation of Federated Learning for Breast Cancer Prediction, aiming to improve both privacy and prediction accuracy. The federated learning framework involves collaboration among multiple healthcare institutions, each possessing a subset of breast cancer patient data. The model is trained locally on each institution's data, and explored for decentralized model training on distributed data sources, preserving privacy and security. The trained federated model is evaluated and saved for deployment. Further research in this area holds promise for revolutionizing healthcare delivery worldwide.

Keywords: Breast cancer prediction, Artificial intelligence, Deep learning, Centralized learning, Decentralized learning, Image-based analysis, Machine learning, Healthcare applications.

I. INTRODUCTION

This project endeavors to develop a formidable challenge in the realm of global healthcare, with its prevalence and impact extending across diverse demographics. The pursuit of accurate predictive models for early detection and personalized treatment has been a cornerstone of breast cancer research. However, the inherent sensitivity and privacy concerns associated with medical data have prompted a paradigm shift in the way we approach predictive modeling. Enter Federated Learning (FL), an innovative and decentralized approach that holds the promise of significantly advancing breast cancer prediction while safeguarding the privacy of individual patient information.

Federated Learning emerges as a beacon of hope in the era of privacy- preserving machine learning. Initially conceived in response to privacy challenges in mobile applications, FL has found fertile ground in healthcare, promising a revolutionary way to harnessthe collective intelligence of distributed datasets without compromising individual privacy. This introduction delves into the conceptual underpinnings of Federated Learning, exploring its principles, advantages, and potential applications in the intricate landscape of breast cancer prediction. Breast cancer is a complex and heterogeneous disease, with various factors influencing its occurrence, progression, and response to treatment. Traditional machine learning approaches, often centralized, have shown promise in predictive modeling using comprehensive datasets. However, the centralized nature of these models raises ethical and privacy concerns, particularly when dealing with sensitive medical information. The need for large, diverse datasets collides with the imperative to protect patient privacy, necessitating the exploration of alternative methodologies.

II. LITERATURE SURVEY

In [1] they introduce Prop-FFL, a novel federated learning method for histopathology image classification, aiming to enhance model fairness among participating hospitals. Unlike existing approaches that optimize average aggregated loss, Prop-FFL employs a unique objective function to reduce performance disparities among hospitals, resulting in a more equitable model performance. The method is validated on histopathology and general datasets, showcasing promising outcomes in terms of speed, accuracy, and fairness. The paper acknowledges limitations in the FedAVG scenario and suggests future work for adapting local trainingto suit PropFFL.

[2] Explores Federated Learning (FL) as a transformative area of AI and machine learning. The FL model encompasses six key stages: problem identification, client selection, local computation, and central server aggregation. These stages ensure data privacy and efficient communication through techniques such as secure aggregation, noise addition, lossy compression, and update clipping.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024

DOI: 10.17148/IJARCCE.2024.134103

[3] About a privacy-preserving technique, initially proposed by Google in 2016, for training models on real data without compromising privacy. It finds applications in smart cities by facilitating data sharing and fusion for urban development. In the medical field, FL ensures secure disease detection model development while preserving patient privacy

[4] This study explores the potential of the VGG-16 architecture, a Convolutional Neural Network (CNN) model, for accurate brain tumor detection through deep learning. Utilizing a dataset consisting of 1655 brain MRI images with tumors and 1598 images without tumors, the VGG-16 model was fine-tuned and trained on this data. Initial training achieved an accuracy of 91%, which was improved to 94% after hyperparameter optimization

III. SCOPE AND METHODOLOGY

Aim of the project

Addresses the inherent challenges in healthcare data sharing, by fostering a collaborative ecosystem, federated learning can enhance model robustness against biases inherent in individual datasets, contributing to more equitable and accurate predictions across diverse patient populations.

The decentralized nature of this approach not only upholds data privacy but also encourages the inclusion of data from various healthcare systems, Furthermore, the scalability of federated learning allows for the integration of emerging data sources. As the healthcare community increasingly embraces collaborative efforts, federated learning stands poised to revolutionize breast cancer prediction, ultimately leading to improved diagnostic accuracy and patient outcomes.

Existing system

The existing paradigms of breast cancer prediction rely on centralizing medical data, posing a significant challenge as it jeopardizes patient privacy and impedes collaborative efforts among institutions. The crux of the issue lies in the compromise between enhancing prediction accuracy and safeguarding sensitive medical information.

The current methodologies not only hinder secure data sharing but also impede effective cross-institutional research endeavors. To address this multifaceted problem, the proposed project aims to leverage Federated Learning, an innovative and decentralized technique.

Proposed system

By employing Federated Learning, the project seeks to facilitate collaborative model training without the necessity of sharing raw medical data.

This approach holds the potential to revolutionize breast cancer prediction by simultaneously advancing accuracy, respecting the paramount importance of data privacy, and fostering a conducive environment for collaborative research across diverse healthcare institutions. The goal is to establish a robust and privacypreserving framework that transcends the limitations of conventional methods.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024
DOI: 10.17148/IJARCCE.2024.134103

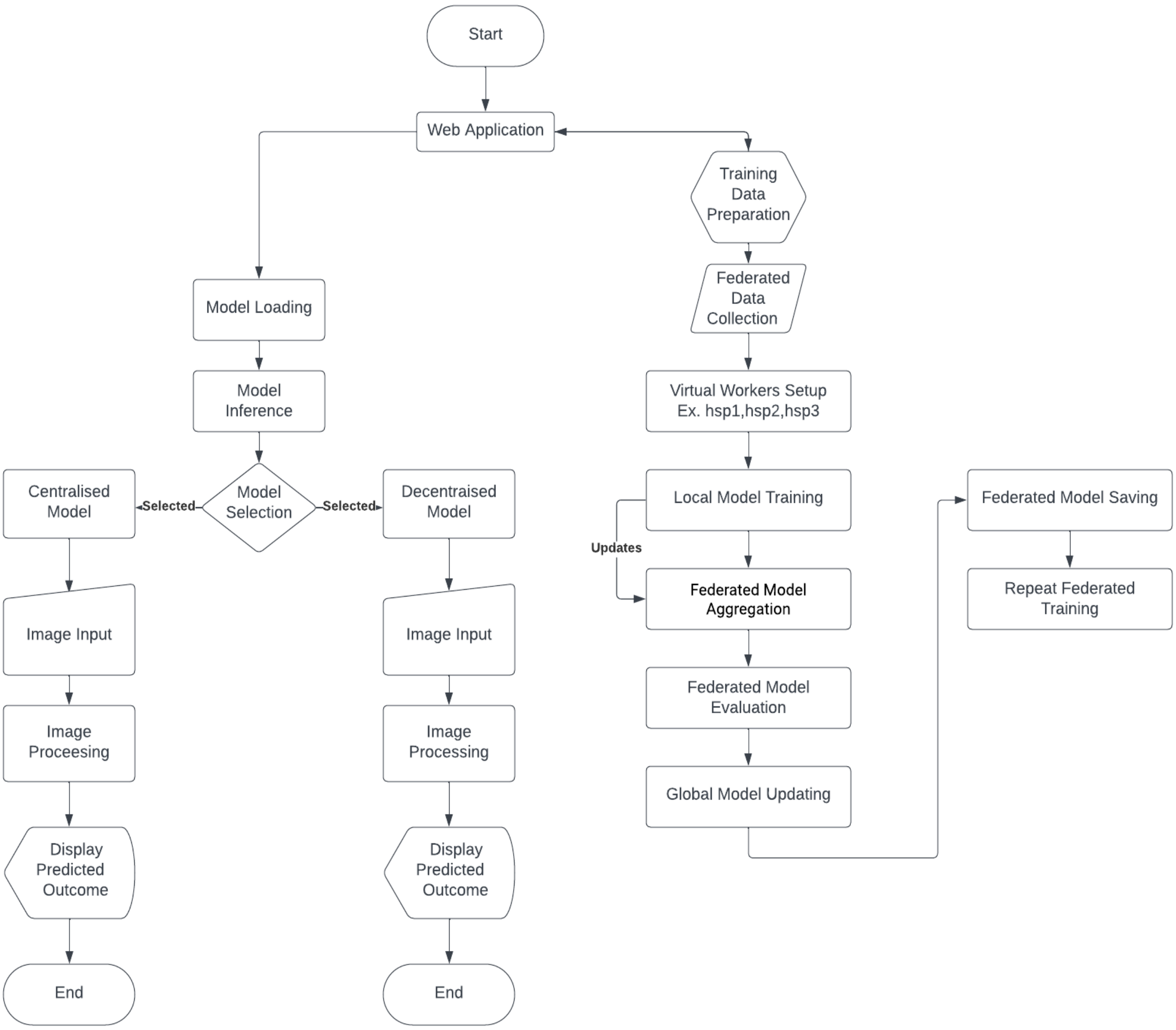

Fig 1 Proposed system

System Architecture

The architecture diagram depicts a streamlined process for automated leaf description generation. It begins with inputting leaf images, followed by pre-processing to standardize them.

Features such as area, color, and vein structure are extracted. A machine learning model analyzes these features to predict the leaf description, including geolocation, species, age, and disease status, enabling efficient plant monitoring and classification.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁑ Peer-reviewed & Refereed journal ⁑ Vol. 13, Issue 4, April 2024

DOI: 10.17148/IJARCCE.2024.134103

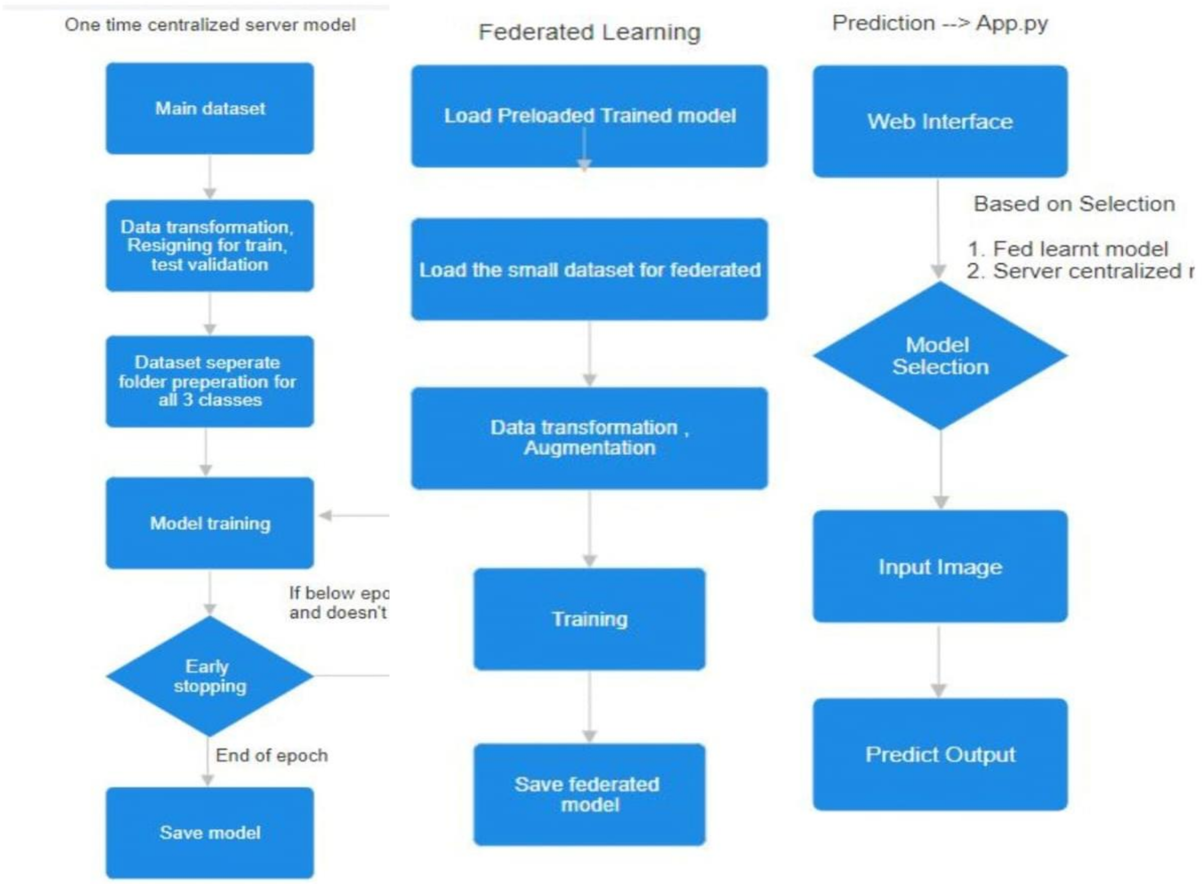

Fig 2. System Architecture

IV. RESULTS

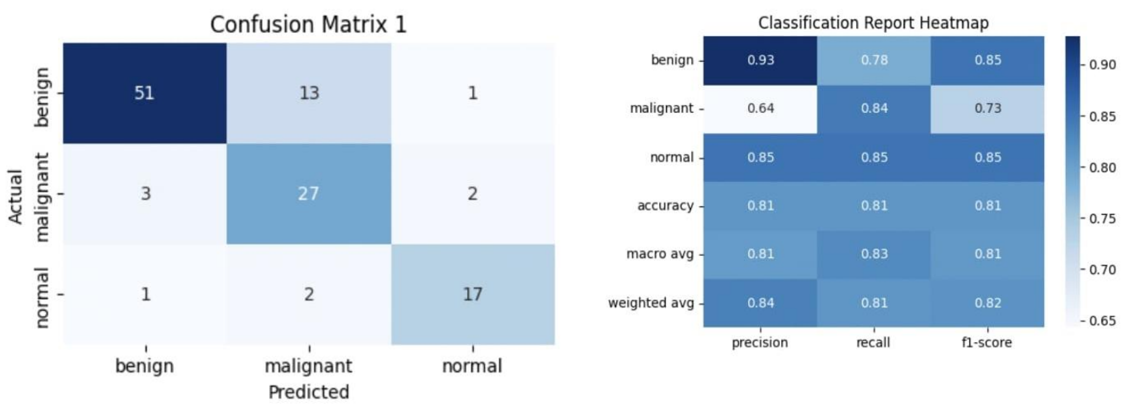

Fig 3 Confusion Matrix

 ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⩳ Peer-reviewed & Refereed journal ⩳ Vol. 13, Issue 4, April 2024

DOI: 10.17148/IJARCCE.2024.134103

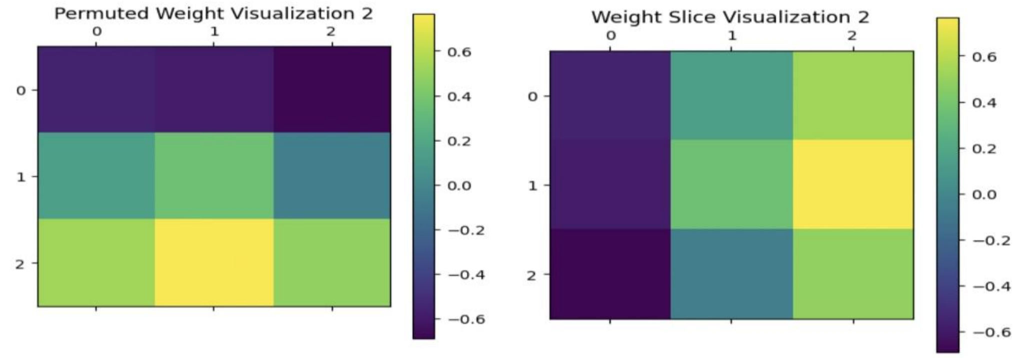

Fig 3.1 Visualization

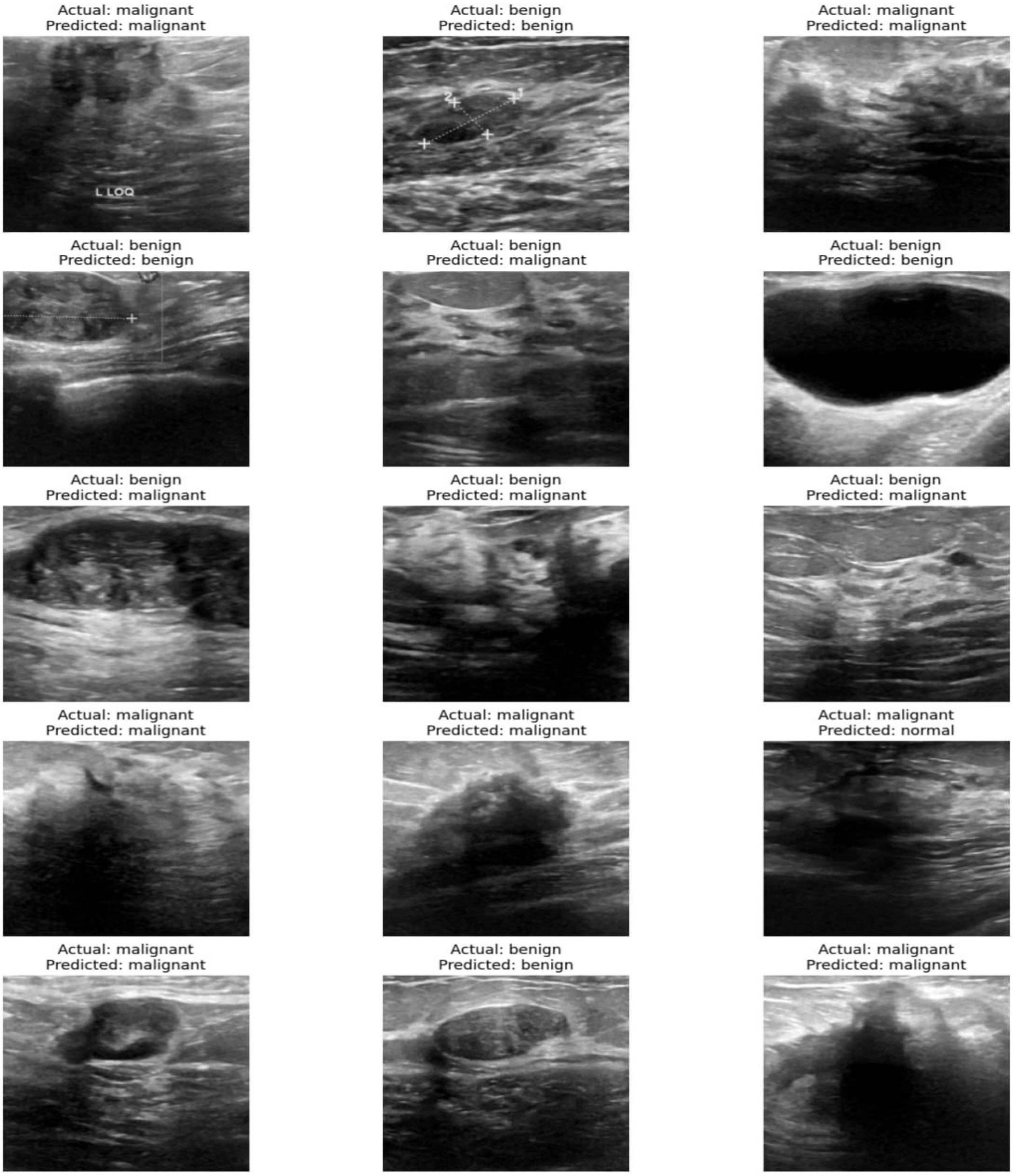

Fig 3.2 Test Results

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024
DOI: 10.17148/IJARCCE.2024.134103

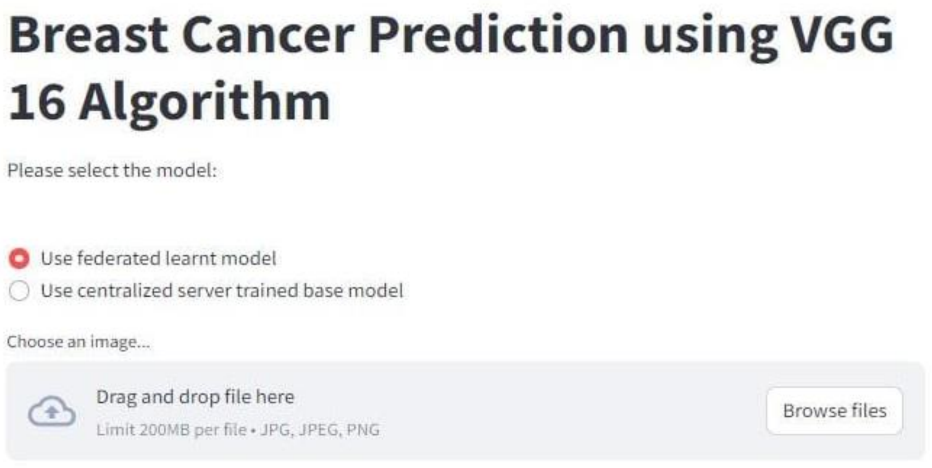

Fig 3.4 User Interface

V. CONCLUSION

In conclusion, the project "Harnessing AI for Precise Estimation of Medicinal Leaf Characteristics" signifies a significant advancement in the field of medicinal plant identification, achieved through the integration of state-of-the-art AI technologies such as machine learning and deep learning.

Utilizing sophisticated AI models, the project has demonstrated remarkable accuracy in discerning medicinal plant species based on their leaf characteristics, promising safer and more efficient utilization of herbal remedies. The user-friendly Streamlit interface ensures accessibility for researchers, botanists, and enthusiasts, facilitating broad adoption and understanding of these innovative techniques in medicinal plant identification.

REFERENCES

[1]. S. Hosseini, M. Sikaroudi, M. Babaie, H. Tizhoosh, and M. Hosseini, "Proportionally Fair Hospital Collaborations in Federated Learning of Histopathology Images," in 2023 IEEE Transactions on Medical Imaging.
[2]. S. Kurupathi and W. Maass, "Survey on Federated Learning towards Privacy Preserving AI,"
[3]. J. Wen, Y. Lan, J. Cai, W. Zhang, Z. Zhang, and Z. Cui, "A survey on federated learning: challenges and applications," Springer Nature, 2022.
[4]. L. Li, N. Xie, and S. Yuan, "A Federated Learning Framework for Breast Cancer Histopathological Image Classification,"
[5]. Atul Kathole and Savita Kumbhare, "Federated Learning aided Breast Cancer Detection with Intelligent Heuristic-based Deep Learning Framework," in Proceedings of the IEEE International Conference on Intelligent Systems and Control (ISCO), [2022]..
[6]. Kumar, P., & Sanjeev. "A Review Paper on Breast Cancer Detection Using Deep Learning." In Proceedings of the IOP,[2021].
[7]. Paulson, A.; Ravishankar, S. AI Based Indigenous Medicinal Plant Identification. In Proceedings of the 2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCT HPA), Cochin, India, 2–4 July 2020.
[8]. Prerepa Gayathri, Aiswarya Dhavileswarapu, Sufyan Ibrahim, Rahul Paul, Reena Gupta, "Exploring the Potential of VGG-16 Architecture for Accurate Brain Tumor Detection Using Deep Learning," [Online]. Available: Context_1. [Accessed: Insert Date].