iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 Issue: V Month of publication: May 2024

DOI: https://doi.org/10.22214/ijraset.2024.61522

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Enhancing Breast Cancer Classification in Histopathological Images through Federated Learning Framework

Hari Prasad Lavudya[1], Akhil Parsa[2], Uday Shanker Bollaveni[3], Suyash Agrawal[4]
[1, 2, 3]UG Students, [4]Assistant Professor, Department of Computer Science and Engineering(Internet of Things), Guru Nanak Institutions Technical Campus, Hyderabad, Telangana, India

Abstract: The proposed study will address the increasing mortality rate of breast cancer in women by implementing an automated disease diagnosis system. The system combines federated learning and deep learning to increase efficiency and address the challenge of securely sharing sensitive medical images. This process includes image capture, encryption using an image encryption method, secure data storage using the Federated Learning Flower framework, and disease classification using a deep learning neural network model. The proposed system achieves high performance in terms of precision, recall, precision, and F-measure, as demonstrated through simulation analysis using the Break His database. The results show promise in automated breast cancer diagnosis with improved safety and efficiency.
Keywords: Federated Learning Framework (FLF), Convolutional Neural Network (CNN), Deep Neural Network (DNN), Deep belief Network (DBN), Recurrent Neural Network (RNN)

I. INTRODUCTION

In recent years, artificial intelligence (AI) technology has further developed and can be used mainly for intelligent medical detection processes. The effectiveness of intelligent medical recognition is based on large amounts of high-quality data obtained from learning models. Nevertheless, data confidentiality and patient privacy are challenging issues for secure access to medical information. Therefore, ensuring security and privacy are two main concerns when using AI for intelligent medical detection. Currently, due to the improved efficiency of federated learning in data security, its growth is increasing day by day. This federated learning helps data owners to train techniques locally and capture specific parameters instead of combining data directly. To enhance the security of the provided medical data, the Federated Learning Flower (FLF) framework was used in the proposed study. In recent decades, the incidence of breast cancer in women has increased and, due to its severity, mortality has also increased. Due to its diverse morphological features, breast cancer is considered one of the major heterogeneous diseases. Therefore, the proposed study used breast cancer medical data to ensure safety through federated learning. To enhance the security of the provided medical data, the Federated Learning Flower (FLF) framework was used in the proposed study. In recent decades, the incidence of breast cancer in women has increased and, due to its severity, mortality has also increased. Due to its diverse morphological features, breast cancer is considered one of the major heterogeneous diseases. Therefore, the proposed study leveraged breast cancer medical data to ensure safety through federated learning. Among the different types of cancer, breast cancer is considered to be the most common cancer among women worldwide. Early detection of tumors is critical to reducing mortality from breast cancer. Currently, several techniques have been proposed for early detection of breast cancer. However, early detection of breast cancer remains a challenge for medical professionals. Depending on the histopathological dataset of breast cancer, advanced detection techniques have been developed to perform the cancer detection process. Existing research has developed both machine learning and deep learning models to classify breast cancer disease. However, machine learning methods yield lower classification results compared to deep learning models. Deep learning models that have existed to classify breast cancer include Convolutional Neural Networks (CNN), Deep Neural Networks (DNN), Deep Believe Networks (DBN), ResNet DenseNet, and Recurrent Neural Networks (RNN). Convolutional networks are widely used for cancer classification tasks in the Break His database. Furthermore, the use of histopathological data in experimental scenarios is highly complex. Sufficiently improved quality data is more important for successful training of learning models. Medical datasets pose several challenges due to insufficient and unevenly distributed data. Therefore, the conflict between data fusion and sufficient data protection requirements is an important obstacle for the development of intelligent health systems.

In this case, federated learning, a technique that combines the knowledge gained from data rather than the data itself, is suitable for development. Intelligent medical recognition systems are very suitable. Recently, the need for early detection of breast cancer has increased around the world, leading to new research approaches in several hospitals and laboratories. Significant advances in deep learning models play a key role in detecting breast cancer from provided input images. Additionally, the concept of federated learning provides an interesting mechanism for adding certainty to a given input sample. Therefore, authors are motivated to design robust FLF frameworks along with deep learning models to provide security and make efficient decisions in classification.

II. EXISTING SYSTEM

In existing systems, various studies have been attempted to develop mechanisms to detect breast cancer based on a given input sample. However, these approaches have limitations, including inefficiencies, challenges in securely sharing sensitive medical images, and difficulties in early detection due to the heterogeneity of breast cancer. Research includes a variety of techniques, including: B. Deep learning models (e.g.)

A. Existing System Disadvantages

1) They may face limitations in terms of performance, processing time, or privacy concerns.
2) Data privacy and security issues.

III. PROPOSED SYSTEM

The proposed system introduces an automated disease diagnosis system.InceptionV3 is a convolutional neural network (CNN) architecture developed by Google Research as part of the Inception project. It is designed for image classification and object detection tasks. InceptionV3 represents the third iteration of the Inception architecture, following Inception and InceptionV2. It was introduced in 2015 and is widely used in various applications such as image recognition, object detection, and image segmentation.

A. Proposed System Advantages

1) Efficient use of computational resources.
2) Better utilization of information.
3) InceptionV3 incorporates various regularization techniques such as batch normalization and dropout, preventing overfitting during training and contributing to the model's robustness and generalization ability

IV. METHODOLOGIES

The system represents a fusion of federated learning and deep learning techniques and is strategically designed to improve diagnostic efficiency while effectively managing the secure sharing of sensitive medical images. The workflow includes several key phases. It starts with the capture of relevant medical images, followed by encryption to ensure data security using an image encryption method. The securely encrypted images are stored using the Federated Learning Flower framework, allowing collaborative model training across distributed datasets without compromising privacy. The core of the system lies in the disease classification phase, where a deep learning neural network model carefully trained on encrypted medical images is used. Through thorough simulation analysis using the BreakHis database, the proposed system exhibits excellent performance metrics such as precision, recall, accuracy, and F-measure. These results highlight the system's potential to revolutionize automated breast cancer diagnosis, promising both to enhance safety measures and improve diagnostic efficiency in healthcare.

A. Dataset

The dataset used in the proposed study is the BreakHis database, a widely recognized and commonly used dataset in breast cancer research. The BreakHis database contains high-resolution histopathological images of breast tissue samples, carefully classified into four different classes based on the presence or absence of benign or malignant tumors and the histological subtypes of the tumors. This dataset has a diverse representation of breast tissue pathology and contains a variety of histological patterns and features commonly encountered in clinical practice.

B. Data Splitting

The data set may be split into multiple subsets for different purposes: B. For training, validation, and testing of the automated disease diagnosis system. Typically, the dataset is split into three major subsets: training, validation, and testing. The training set is used to train the machine learning model so that it can learn patterns and relationships in the data.

C. Data Augmentation

Data augmentation is an important technique used in machine learning and deep learning tasks,including image classification tasks such as automated disease diagnosis. It involves artificially increasing the size and diversity of the training data set by applying various transformations to the original data samples. These transformations typically include geometric transformations such as rotation, scaling, translation, and mirroring, as well as color and contrast adjustments, noise reduction, and cropping.

D. Model Building

InceptionV3 is a convolutional neural network (CNN) architecture developed by a Google research team as part of the Inception model series. It is primarily designed for image classification tasks and has gained popularity due to its efficiency and effectiveness in processing large image datasets. Inception V3 builds on the principles of previous Inception models, but introduces several key innovations to improve performance and efficiency.

E. Image Uploading

Image uploading is a fundamental process in many applications, especially in image processing, computer vision, and multimedia content. Image uploading basically refers to transferring a digital image file from a local device or storage location to a server, cloud storage, or application interface. This process allows users to share, store, analyze, and edit images within a variety of digital platforms and systems.

F. Prediction

Prediction is a fundamental concept in machine learning and statistical modeling and refers to the process of estimating or predicting the outcome of a particular event or phenomenon based on available data and learned patterns. In the context of machine learning models, prediction involves using a trained model to make educated guesses and decisions about unseen or future data samples. \N.The first module where users can sign up and log in is this one.

V. REQUIREMENTS

These are the requirements for doing the project. Without using these tools & software's we can't do the project. So we have two requirements to do the project. They are

A. Hardware Requirements

The hardware requirements should be a comprehensive and uniform definition of the entire system since they might form the foundation of a contract for the system's implementation. Software engineers use them as the foundation for their system designs. It focuses on the functionality of the system rather than the best way to use it.

1) PROCESSOR : DUAL CORE 2 DUOS.
2) RAM : 2GB DD RAM
3) HARD DISK : 250 GB

B. Software Requirements

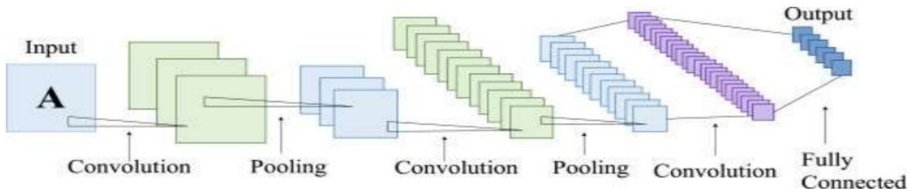
The software requirements document is the specification of the system. It should include both a definition and a specification of requirements. It is a set of what the system should do rather than how it should do it. The software requirements provide a basis for creating the software requirements specification. It is useful in estimating cost, planning team activities, performing tasks and tracking the teams and tracking the team's progress throughout the development activity.

1) Operating System: Windows 7/8/10
2) Platform: Spyder3
3) Programming Language: Python
4) Front End: Spyder3

VI. SYSTEM ARCHITECTURE

Based on the image you sent, it appears to be a convolutional neural network (CNN) architecture for image classification. CNNs are a type of deep learning algorithm commonly used in image recognition and classification tasks. Here's a breakdown of the general components visible in the image:

1) *Input Layer:* This layer receives the pre-processed image data as input. In breast cancer detection, this would likely be a mammogram image.
2) *Convolutional Layers:* These layers are the core of the CNN architecture. They apply filters (kernels) to the input image, extracting features like edges, shapes, and textures. Multiple convolutional layers can be stacked to learn increasingly complex features.
3) *Pooling Layers:* These layers down sample the output from the convolutional layers, reducing the dimensionality of the data and making the network more computationally efficient. Common pooling techniques include max pooling and average pooling.
4) *Fully Connected Layers:* These layers connect all the neurons from the previous layer to every neuron in the current layer. They take the features extracted by the convolutional layers and learn higher-level patterns for classification.
5) *Output Layer:* This layer represents the final output of the network. In breast cancer detection, it would typically have two neurons, corresponding to the probabilities of the input image being benign or malignant.The text labels in the image, "Convolution," "Pooling," "Fully Connected," further reinforce the presence of these CNN components.



VII. LITERATURE SURVEY

Recent advances in communication technology and the medical Internet of Things have transformed smart healthcare through artificial intelligence (AI). Traditionally, AI technologies have required centralized data collection and processing, but the highly scalable nature of modern healthcare networks and growing privacy concerns mean that this may not be possible in realistic healthcare scenarios. there is. As a new distributed and collaborative AI paradigm, Federated Learning (FL) is particularly attractive for intelligent healthcare because it coordinates multiple customers (e.g., hospitals) to provide AI training without sharing raw data. Therefore, we provide a comprehensive survey on the use of FL in smart healthcare. First, we introduce recent advances in FL, motivations and requirements for its use in smart healthcare. Next, we discuss the latest FL designs for smart healthcare, ranging from resource-aware FL, secure and privacy-aware FL, incentive FL, and personalized FL. We then provide an updated overview of Florida's emerging applications in key health care areas, including health data management, remote health monitoring, medical imaging, and COVID-19 detection. Several current Florida-based smart healthcare projects are analyzed and key findings from the research are highlighted. Finally, we discuss interesting research questions and possible directions for future FL research in the smart healthcare field.

Background and Objectives: Regular screening with mammography is recommended for early detection of breast cancer. Regular testing results in datasets with mostly negative samples. Limited representativeness of positive cases can be a challenge in training computer-aided diagnosis (CAD) systems. To mitigate this issue, collecting data from multiple institutions is a possible solution. Recently, federated learning has emerged as an effective tool for collaborative learning. In this setting, the local model performs calculations on private data to update the global model. The order and frequency of local updates affectsthe final global model. We investigate the role of the order in which samples are presented locally to the optimizer in the context of federated adversarial learning to improve multi-site breast cancer classification. Methodology: Define a new memory-aware curriculum learning method for federated settings. Our goal is to improve the consistency of local models and eliminate inconsistent predictions. H. To punish forgotten samples. Our curriculum controls the order of training samples and prioritizes samples that are forgotten after global model deployment. Our approach is combined with unsupervised domain adaptation to handle domain migration while preserving privacy.

Artificial intelligence (AI) technology is rapidly evolving. Many applications currently use AI to diagnose breast cancer. However, most of the new research to date has only been conducted in converged learning (CL) environments, which comes with the risk of data breaches. Furthermore, the use of AI technology to accurately identify and localize lesions and predict tumors is expected to increase the chances of patient survival. To solve these problems, we developed a federated learning (FL) feature that extracts features from the participating environment instead of CL features.

Novel contributions of this study include (i) the application of transfer learning to extract data features from regions of interest (ROIs) in images; This is intended to enable careful preprocessing and data enrichment for data training purposes. (ii) use synthetic minority oversampling technique (SMOTE) for data processing; This aims to classify data more uniformly and improve disease diagnostic prediction performance. (iii) Application of his FeAvg-CNN + MobileNet in FL framework to ensure customer privacy and personal security. (iv) We present experimental results from various deep learning, transfer learning, and FL models on balanced and unbalanced mammography datasets and demonstrate that our solution outperforms other approaches used in AI applications in healthcare. also yields much higher classification performance, showing that it is feasible.

VIII. EXISTING ALGORITHM

A. Convolutional neural Network (CNN)

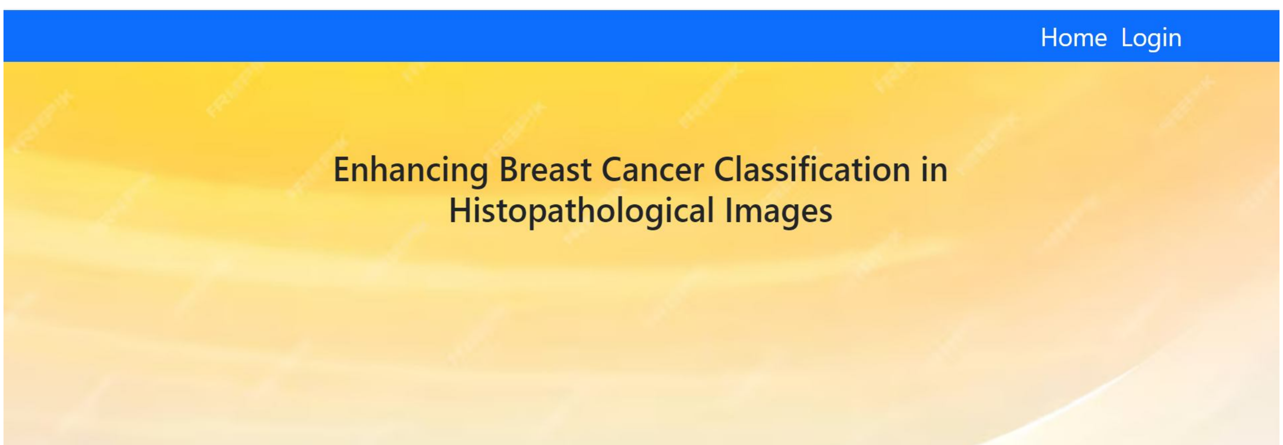
Some common algorithms mentioned in existing breast cancer classification studies include convolutional neural network (CNN), transfer learning methods, and various optimization algorithms for fine-tuning hyper parameters. Each study may have employed different algorithms based on its specific approach and requirements.

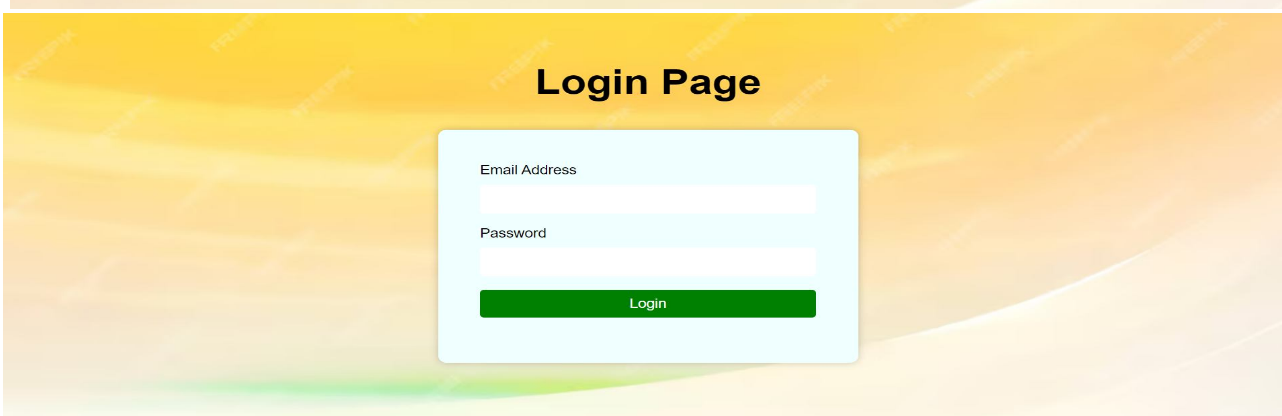
IX. PROPOSED ALGORITHM

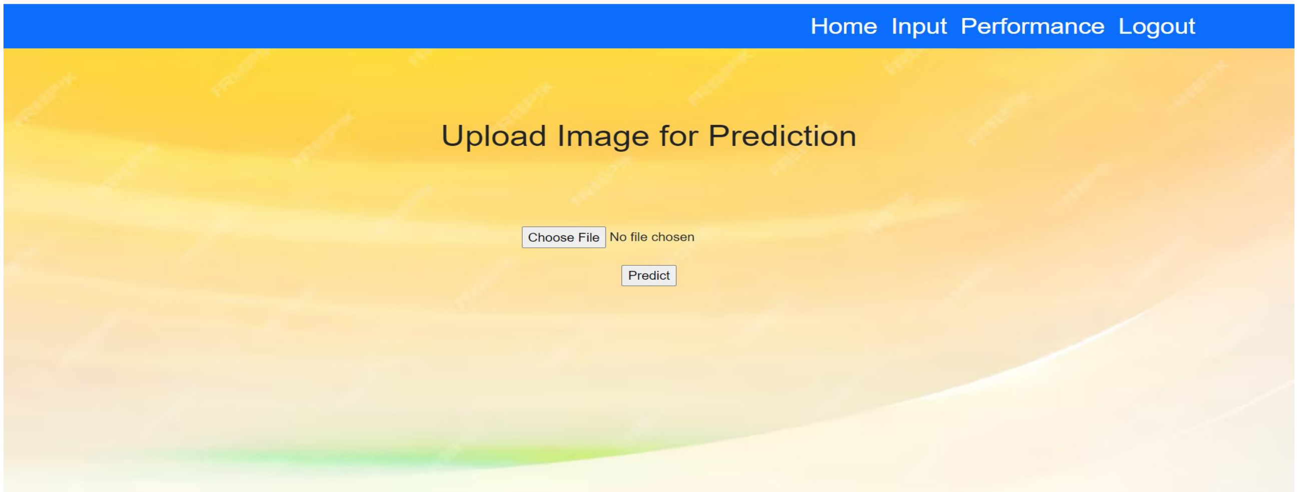
A. Inception V3

InceptionV3 features a deep network architecture composed of multiple layers of convolutional and pooling operations. It incorporates several key innovations to improve efficiency and performance compared to its predecessors. The hallmark of the Inception architecture is its use of "Inception modules," which are carefully designed blocks that allow the network to capture features at different spatial scales efficiently. These modules consist of multiple parallel convolutional branches with different filter sizes and pooling operations.

X. RESULTS







XI. CONCLUSION

In conclusion, this study presents a novel approach for securely classifying breast cancer using a federated learning framework (FLF) combined with a deep learning model, specifically the InceptionV3 model. Recognizing the critical need for securing medical data to ensure accurate disease diagnosis, our approach integrates efficient FLF mechanisms with robust encryption techniques to safeguard patient privacy and prevent unauthorized access.

REFERENCES

[1] R. F. Mansour, A. E. Amraoui, I. Nouaouri, V. G. Díaz, D. Gupta, and S. Kumar, ''Artificial intelligence and Internet of Things enabled disease diagnosis model for smart healthcare systems,'' IEEE Access, vol. 9, pp. 45137–45146, 2021.

[2] S. Tian, W. Yang, J. M. L. Grange, P. Wang, W. Huang, and Z. Ye, ''Smart healthcare: Making medical care more intelligent,'' Global Health J., vol. 3, no. 3, pp. 62–65, Sep. 2019.

[3] S. Ahmed, M. F. Hossain, M. S. Kaiser, M. B. T. Noor, M. Mahmud, and C. Chakraborty, ''Artificial intelligence and machine learning for ensuring security in smart cities,'' in Data Driven Mining, Learning and Analytics for Secured Smart Cities. Cham, Switzerland: Springer, 2021, pp. 23–47.

[4] T. M. Ghazal, ''Internet of Things with artificial intelligence for health care security,'' Arabian J. Sci. Eng., Aug. 2021.

[5] D. C. Nguyen, Q.-V. Pham, P. N. Pathirana, M. Ding, A. Seneviratne, Z. Lin, O. Dobre, and W.-J. Hwang, ''Federated learning for smart healthcare: A survey,'' ACM Comput. Surveys, vol. 55, no. 3, pp. 1–37, Mar. 2023.

[6] J. C. Jiang, B. Kantarci, S. Oktug, and T. Soyata, ''Federated learning in smart city sensing: Challenges and opportunities,'' Sensors, vol. 20, no. 21, p. 6230, Oct. 2020.

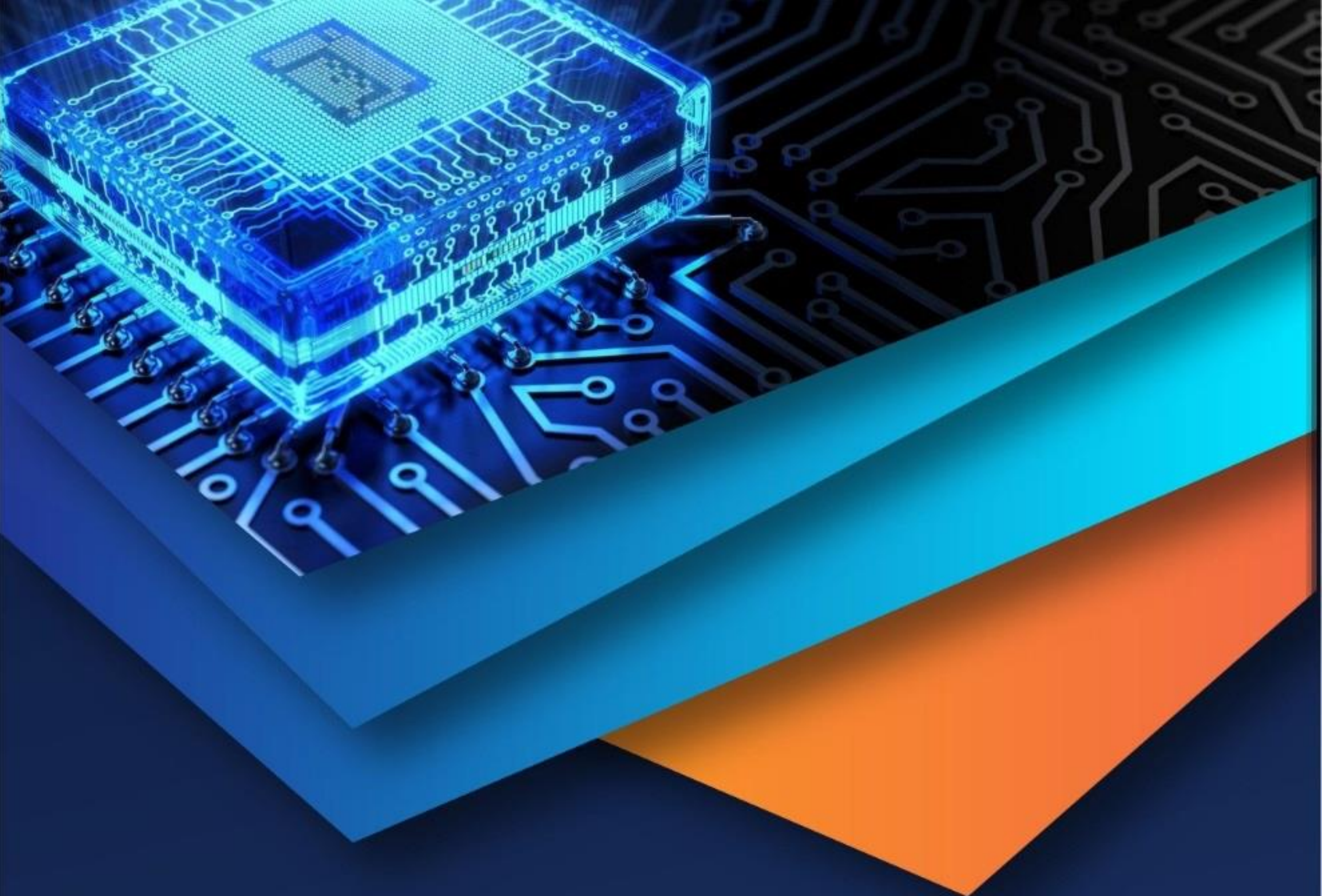
[7] M. A. Rahman, M. S. Hossain, M. S. Islam, N. A. Alrajeh, and G. Muhammad, ''Secure and provenance enhanced Internet of Health Things framework: A blockchain managed federated learning approach,'' IEEE Access, vol. 8, pp. 205071–205087, 2020.

[8] C. Zhang, Y. Xie, H. Bai, B. Yu, W. Li, and Y. Gao, ''A survey on federated learning,'' Knowl.-Based Syst., vol. 216, Mar. 2021, Art. no. 106775.

[9] A. G. Waks and P. E. Winer, ''Breast cancer treatment: A review,'' Jama, vol. 321, no. 3, pp. 288–300, 2019.

[10] M. F. Ullah, ''Breast cancer: Current perspectives on the disease status,'' Breast Cancer Metastasis and Drug Resistance. New York, NY, USA: Springer, 2019, pp. 51–64.

[11] A. Conti, A. Duggento, I. Indovina, M. Guerrisi, and N. Toschi, ‘‘Radiomics in breast cancer classification and prediction,’’ Seminars Cancer Biol., vol. 72, pp. 238–250, Jul. 2021. 43
[12] B. Sahu, S. Mohanty, and S. Rout, ‘‘A hybrid approach for breast cancer classification and diagnosis,’’ EAI Endorsed Trans. Scalable Inf. Syst., vol. 6, no. 20, pp. 1–8, 2019.
[13] F. K. Al-Thoubaity, ‘‘Molecular classification of breast cancer: A retrospective cohort study,’’ Ann. Med. Surgery, vol. 49, pp. 44–48, Jan. 2020.
[14] E. H. Houssein, M. M. Emam, A. A. Ali, and P. N. Suganthan, ‘‘Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review,’’ Exp. Syst. Appl., vol. 167, Apr. 2021, Art. no. 114161.
[15] D. A. Ragab, O. Attallah, M. Sharkas, J. Ren, and S. Marshall, ‘‘A framework for breast cancer classification using multi-DCNNs,’’ Comput. Biol. Med., vol. 131, Apr. 2021, Art. no. 104245.
[16] F. F. Ting, Y. J. Tan, and K. S. Sim, ‘‘Convolutional neural network improvement for breast cancer classification,’’ Exp. Syst. Appl., vol. 120, pp. 103–115, Apr. 2019.
[17] M. Lamba, G. Munjal, and Y. Gigras, ‘‘A hybrid gene selection model for molecular breast cancer classification using a deep neural network,’’ Int. J. Appl. Pattern Recognit., vol. 6, no. 3, pp. 195–216, 2021.
[18] A. M. Abdel-Zaher and A. M. Eldeib, ‘‘Breast cancer classification using deep belief networks,’’ Exp. Syst. Appl., vol. 46, pp. 139–144, Mar. 2016.
[19] Q. A. Al-Haija and A. Adebanjo, ‘‘Breast cancer diagnosis in histopathological images using ResNet-50 convolutional neural network,’’ in Proc. IEEE Int. IoT, Electron. Mechatronics Conf. (IEMTRONICS), Sep. 2020, pp. 1–7.
[20] H. Yao, X. Zhang, X. Zhou, and S. Liu, ‘‘Parallel structure deep neural network using CNN and RNN with an attention mechanism for breast cancer histology image classification,’’ Cancers, vol. 11, no. 12, p. 1901, Nov. 2019.
[21] L. Li, N. Xie, and S. Yuan, ‘‘A federated learning framework for breast cancer histopathological image classification,’’ Electronics, vol. 11, no. 22, p. 3767, Nov. 2022.
[22] B. S. Abunasser, M. R. J. Al-Hiealy, I. S. Zaqout, and S. S. Abu-Naser, ‘‘Breast cancer detection and classification using deep learning xception algorithm,’’ Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 7, pp. 1–7, 2022.
[23] K. Jabeen, M. A. Khan, M. Alhaisoni, U. Tariq, Y.-D. Zhang, A. Hamza, A. Mickus, and R. Damaševičius, ‘‘Breast cancer classification from ultrasound images using probability-based optimal deep learning feature fusion,’’ Sensors, vol. 22, no. 3, p. 807, Jan. 2022. 44
[24] A. Jiménez-Sánchez, M. Tardy, M. A. G. Ballester, D. Mateus, and G. Piella, ‘‘Memory aware curriculum federated learning for breast cancer classification,’’ Comput. Methods Programs Biomed., vol. 229, Feb. 2023, Art. no. 107318.
[25] R. O. Ogundokun, S. Misra, M. Douglas, R. Damaševičius, and R. Maskeliunas, ‘‘Medical Internet-of-Things based breast cancer diag- ¯ nosis using hyperparameter-optimized neural networks,’’ Future Internet, vol. 14, no. 5, p. 153, May 2022.
[26] Y. N. Tan, V. P. Tinh, P. D. Lam, N. H. Nam, and T. A. Khoa, ‘‘A transfer learning approach to breast cancer classification in a federated learning framework,’’ IEEE Access, vol. 11, pp. 27462–27476, 2023.
[27] P. Manikandan, U. Durga, and C. Ponnuraja, ‘‘An integrative machine learning framework for classifying SEER breast cancer,’’ Sci. Rep., vol. 13, no. 1, pp. 1–12. 2023.
[28] J O. Du Terrail et al., ‘‘Federated learning for predicting histological response to neoadjuvant chemotherapy in triple-negative breast cancer,’’ Nature Med., vol. 29, pp. 1–12, Jan. 2023.
[29] D. S. Laiphrakpam and M. S. Khumanthem, ‘‘Medical image encryption based on improved ElGamal encryption technique,’’ Optik, vol. 147, pp. 88–102, Oct. 2017.
[30] H. A. Khikani, N. Elazab, A. Elgarayhi, M. Elmogy, and M. Sallah, ‘‘Breast cancer classification based on histopathological images using a deep learning capsule network,’’ 2022, arXiv: 2208. 00594.



10.22214/IJRASET

INDEX COPERNICUS 45.98

IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)