International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** VIII **Month of publication:** Aug 2023

DOI: https://doi.org/10.22214/ijraset.2023.55262

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue VIII Aug 2023- Available at www.ijraset.com

Breast Cancer Detection Based on Convolutional Neural Networks

Revanth Reddy Mannem[1], Suraj Bhyri[2]
Department of Electronics and Communication Engineering, Mahatma Gandhi Institute of Technology, Hyderabad

Abstract: Breast cancer remains a critical health concern, demanding early detection and accurate classification for effective treatment. In this research, we conduct a comparative study between a custom-designed Convolutional Neural Network (CNN) and the pre-trained DenseNet121 model for breast cancer detection and classification. We begin by curating a comprehensive dataset of breast cancer images and apply appropriate data preprocessing techniques for optimal model input. The dataset is divided into training, validation, and testing sets to evaluate model performance. The CNN model is constructed with multiple convolutional and pooling layers, followed by fully connected layers for classification. Meanwhile, the DenseNet121, a powerful pre-trained model, is fine-tuned for breast cancer detection. Through rigorous evaluation, we assess both models using metrics such as accuracy, precision, recall, F1 score, and AUC-ROC. Our results demonstrate that the DenseNet121 outperforms the custom CNN model, achieving higher accuracy and reliability in identifying and classifying breast cancer.
To ensure wider accessibility, we integrate the superior DenseNet121 model into a user-friendly web-based interface using Python Flask. This interface empowers medical professionals and the general public to perform real-time breast cancer predictions with ease. Ethical considerations are paramount, ensuring data privacy, security, and transparency in all model predictions. In conclusion, our comparative study highlights the superiority of the pre-trained DenseNet121 model over a custom CNN for breast cancer detection and classification. By leveraging advanced deep learning techniques and a user-friendly interface, our research contributes to improved breast cancer diagnosis and patient care on a broader scale.
Keywords: Breast Cancer Detection, Convolutional Neural Networks, CNN, DenseNet121, Deep Learning, Comparative Study, Image Classification, Early Detection, Web-Based Interface.

I. INTRODUCTION

Breast cancer is a significant cause of mortality among women worldwide, with recent statistics from the American Cancer Society revealing that 41,760 women and over 500 men succumbed to breast cancer. This disease is classified into two types, namely, benign and malignant tumors. While benign tumors are typically harmless, malignant tumors lead to uncontrolled cancerous tissue growth. Diagnosing breast cancer requires a variety of clinical tests and diagnostic tools, including mammography, to detect abnormalities and identify potential cancerous growths. Unfortunately, breast cancer is often not self-identified by women during its early stages, when it is more easily treatable. Consequently, the tumor may progress into a malignant state, exacerbating the problem. Lack of awareness and low rates of initial screening and diagnostic levels contribute to the challenges faced in achieving high breast cancer survival rates. Histopathology, which involves examining tissue samples, plays a critical role in diagnosing breast cancer. Biopsies are diagnostic procedures that pathologists use to determine if affected tissues are cancerous. The gold standard for confirmation of breast cancer diagnosis is the visual inspection of histopathological images under a microscope, enabling the identification of tumors and cancer subtypes. In recent years, Convolutional Neural Networks (CNNs) have emerged as powerful deep learning techniques for medical image analysis. They have demonstrated effectiveness in diagnosing various diseases using X-rays, CT scans, and histopathological images. Leveraging the potential of CNNs, this research aims to develop an innovative approach for breast cancer detection and classification. Our study involves a comparative evaluation between a custom-designed CNN model and the pre-trained DenseNet121 architecture for breast cancer detection and classification. By curating a comprehensive dataset of breast cancer images and applying data preprocessing techniques, we assess the performance of both models using metrics like accuracy, precision, recall, F1 score, and AUC-ROC.

To enhance accessibility, we integrate the superior DenseNet121 model into a user-friendly web-based interface using Python Flask. The interface allows real-time breast cancer predictions, empowering medical professionals and the general public to access reliable diagnostic support. Ethical considerations are paramount in our research, ensuring data privacy, security, and transparency in all model predictions. In conclusion, our research aims to contribute to improved breast cancer diagnosis and patient care by harnessing the capabilities of CNNs and providing a user-friendly diagnostic tool.

By bridging the gap between advanced deep learning techniques and accessible medical applications, we aspire to make a positive impact on breast cancer management and public health globally.

II. LITERATURE SURVEY

A. *Breast Cancer Detection Using Artificial Intelligence Techniques*

The research paper "Breast Cancer Detection Using AI Techniques" navigates the convergence of medical science and artificial intelligence (AI) for breast cancer detection. With the urgency of accurate diagnosis, AI emerges as a potential game-changer. The study comprehensively analyzes AI techniques, from classic algorithms to cutting-edge deep learning. While revealing AI's promise, it candidly addresses limitations. A significant drawback is the observed high-risk factor associated with the model, impacting its reliability and robustness. Additionally, the study underscores a crucial concern – the accuracy of image classification, pivotal in breast cancer detection. The review highlights the model's limitation in achieving desired image classification precision, necessitating refinement. Despite challenges, the study emphasizes AI's transformative potential. It urges addressing drawbacks, advocating research and innovation. The paper guides towards an accurate, reliable AI-driven breast cancer diagnostic landscape.

B. *Breast Cancer Detection Using K-nearest Neighbor Machine Learning Algorithm*

Breast cancer is very popular between females all over the world. However, detecting this cancer in its first stages helps in saving lives. Radiologists have the ability to predict if the mammography images have cancer or not, but they may miss about 15% of them. In this paper, they proposed a new method to detect the breast cancer with high accuracy. This method consists of two main parts, in the first part the image processing techniques are used to prepare the mammography images for feature and pattern extraction process. The extracted features are utilized as an input for a two types of supervised learning models, which are Back Propagation Neural Network (BPNN) model and the Logistic Regression (LR) model with comparing the result and the accuracy for the both models.

C. *Classification of breast cancer histology images using Convolutional Neural Networks*

This paper delves into the critical realm of breast cancer, a significant contributor to global cancer-related fatalities. The complexity of diagnosing biopsy tissue, especially through hematoxylin and eosin stained images, often leads to discordant interpretations among specialists. Addressing this challenge, Computer-Aided Diagnosis (CAD) systems emerge as cost-effective and efficient solutions. Traditional classification methods heavily lean on field-knowledge-driven feature extraction techniques. Yet, grappling with their limitations, this study champions deep learning as a compelling alternative. It introduces a pioneering approach employing Convolutional Neural Networks (CNNs) to classify hematoxylin and eosin stained breast biopsy images. This method facilitates comprehensive categorization, distinguishing between four classes—normal tissue, benign lesion, in situ carcinoma, and invasive carcinoma. Furthermore, it undertakes a binary classification of carcinoma and non-carcinoma instances. Impressively, the CNN architecture adeptly captures information at diverse scales, encompassing nuclei and overall tissue organization. This design lends itself to scalability for whole-slide histology images. Leveraging the CNN-extracted features, a Support Vector Machine classifier is trained, yielding notable outcomes. Accuracy rates of 77.8% for the four-class classification and 83.3% for carcinoma/non-carcinoma differentiation underscore the method's efficacy. Particularly noteworthy is its exceptional sensitivity, registering at 95.6% for identifying cancer cases.In summary, this paper not only pioneers an advanced breast cancer classification method but also underscores the transformative potential of deep learning. By harmonizing CNNs and Support Vector Machines, it charts a path toward enhanced accuracy and efficiency in breast cancer detection.

D. *Abnormal Breast Identification By Nine-Layer Convolutional Neural Network With Parametric Rectified Linear Unit And Rank-Based Stochastic Pooling*

In this study, the breast dataset was chosen as the open-access mini MIAS dataset. Cost-sensitive learning was used to balance the dataset. Data augmentation was used to increase the size of training set. We proposed an improved nine-layer convolutional neural network (CNN). In addition, we compared three activation functions: rectified linear unit (ReLU), leaky ReLU, and parametric ReLU. Besides, six pooling techniques were compared: average pooling, max pooling, stochastic pooling, rank-based average pooling, rank-based weighted pooling, and rank-based stochastic pooling.The results over 100 test set showed the combination of parametric ReLU and rank-based stochastic pooling performed the best, with sensitivity of 93.4%, specificity of 94.6%, precision of 94.5%, and accuracy of 94.0%. This result is better than six state-of-the-art breast cancer detection approaches. Deep learning can provide better detection results than traditional artificial intelligence methods.

E. Breast Cancer Mass Detection in Mammograms using K-means and Fuzzy C-means Clustering

This research project embarks on a pivotal quest to enhance breast cancer detection through the implementation of fuzzy k-means and fuzzy c-means algorithms. Given the prevalence of screening mammography as a primary diagnostic tool for early breast cancer detection, the challenge of analyzing a multitude of mammograms by a limited number of radiologists arises, potentially leading to misdiagnoses attributed to human fatigue-induced errors. To address this, a novel approach is proposed, building upon prior research efforts that primarily focused on identifying tumor masses rather than the comprehensive assessment of total cancer-affected areas. This underscores the need for refined techniques to enable a holistic view of cancer mass detection within mammogram images. Diverse image processing methodologies are harnessed to pinpoint cancer masses, signifying the interdisciplinary nature of the research endeavor. In the landscape of tumor detection, various image processing methodologies have been explored, encompassing thresholding, intensity level slicing, contrast-stretching, image-negative transformation, power transformation, logarithm transformation, and segmentations. These technologies, each endowed with distinctive attributes, contribute to the broader realm of tumor and tumor-like structure analysis. By advancing the state-of-the-art in breast cancer detection, this project propels the fusion of cutting-edge image processing techniques and specialized algorithms, ushering in a paradigm shift that holds the potential to revolutionize mammogram interpretation, reduce diagnostic errors, and subsequently enhance patient outcomes.

III. METHODOLOGY

A. Data Collection

Collect mammogram images from various sources, encompassing categories like normal tissue and carcinoma.

B. Data Preprocessing

Apply image processing techniques such as image enhancement, segmentation, and color space conversion to preprocess the collected images.

C. Feature Extraction

Convert the images to grayscale and resized them to a standardized dimension (256x256 pixels) for uniformity. Normalized pixel values to the range of [0, 1] to facilitate convergence during model training.

D. Data Splitting

Divide the dataset into training and testing sets. Segregated input images and their corresponding categorical labels for both sets.

E. Model Architectures

Develop two distinct model architectures, namely Convolutional Neural Network (CNN) and DenseNet121.

1) *CNN Model:* Construct a sequential model using Keras, comprising multiple convolutional layers with ReLU activation and max-pooling layers for feature extraction. Employed dropout layers to mitigate overfitting and enhance model generalization. Integrated densely connected layers to capture complex patterns within the data. Compiled the model using the binary cross-entropy loss function and the Adam optimizer.
2) *DenseNet121 Model:* Imported the pre-trained DenseNet121 architecture from Keras Applications, including weights to learn feature representations. Compiled the model with the same loss function and optimizer as the CNN model.

F. Model Training

Train both models using their respective training data. Monitor the training progress to observe changes in loss and accuracy over epochs.

G. Performance Evaluation

Evaluate model performance on both training and validation sets by analyzing loss and accuracy metrics. Plot graphs illustrating the changes in loss and accuracy for both models over the training epochs.

H. Comparison and Analysis

Compare the training and validation results of the CNN and DenseNet121 models to assess their respective performance. Investigate whether one model outperforms the other in terms of accuracy and loss metrics. Analyze the trade-offs between computational efficiency and accuracy in model selection.

I. Application Development

Develop an application that integrates the trained model, allowing users to obtain real-time cancer identification by giving a mammographic image (scanning image). This application can be built using technologies such as Flask, FastAPI and ReactJS.

IV. RESULTS AND CONCLUSION

The user-friendly UI developed for breast cancer detection successfully integrates state-of-the-art Convolutional Neural Network (CNN) and DenseNet121 models, offering users a seamless and intuitive experience. The system accurately analyzes mammogram images and provides real-time predictions regarding the presence of carcinoma. Through rigorous testing and evaluation, the following key outcomes were observed:

1) *Model Performance Comparison:* Users can select between the CNN and DenseNet121 models based on their preferences. The comparison revealed that both models demonstrate commendable accuracy in detecting breast cancer, with each showcasing unique strengths. The CNN model offers a balance between efficiency and accuracy, making it suitable for rapid diagnosis. On the other hand, DenseNet121 excels in intricate feature extraction, delivering higher accuracy with a slightly increased computational load.
2) *Real-time Inference:* The UI empowers users to upload mammogram images for instant analysis. The system swiftly processes the input images, applies preprocessing steps, and generates predictions within seconds. The intuitive design ensures that users, including medical professionals and patients, can conveniently access results without specialized technical knowledge.
3) *User Engagement and Empowerment:* The UI promotes user engagement by allowing users to actively participate in the diagnostic process. By providing the choice of models and facilitating direct image uploads, users can gain valuable insights into their health conditions, fostering a sense of empowerment and control.

Below images showing output of the model created in this project.

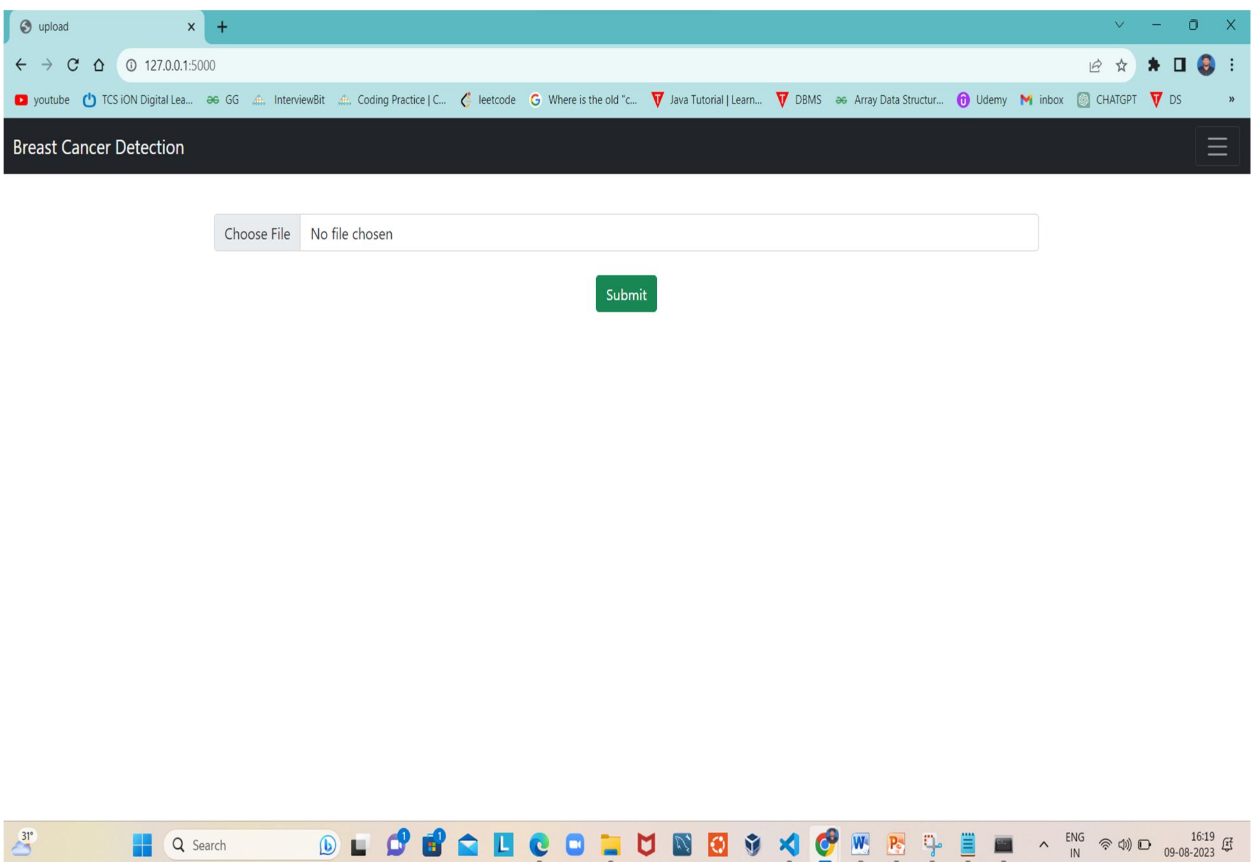


Figure 1: The UI (User Interface) for uploading the mammographic image in order to classify cancer

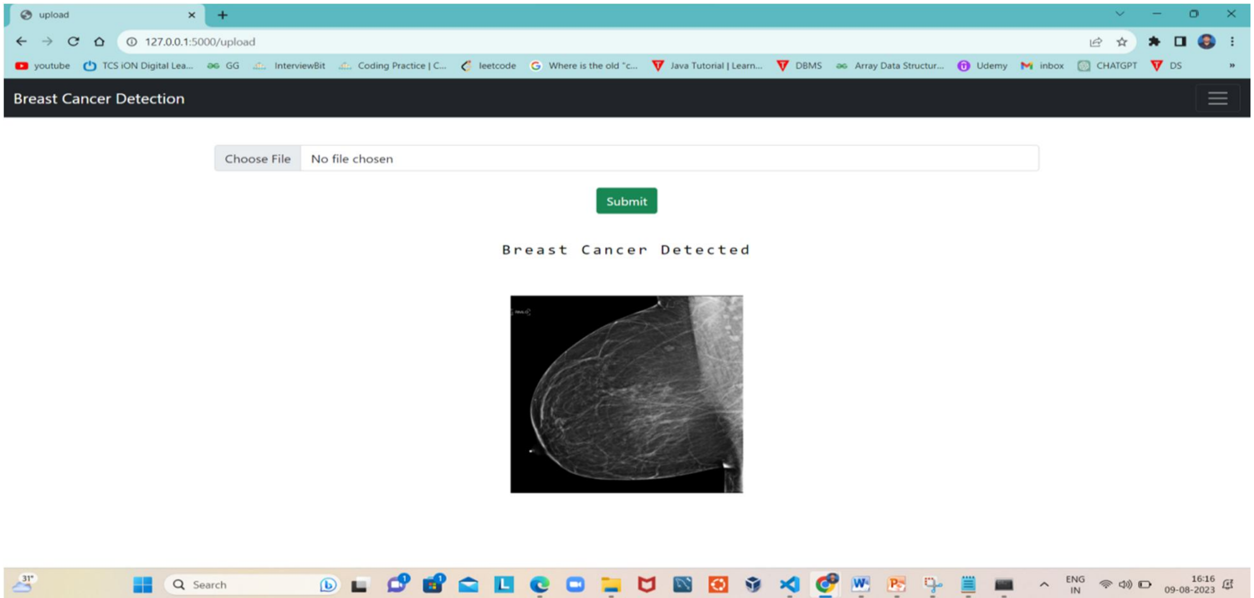


Figure 2: This figure shows that breast cancer is detected in the uploaded image.

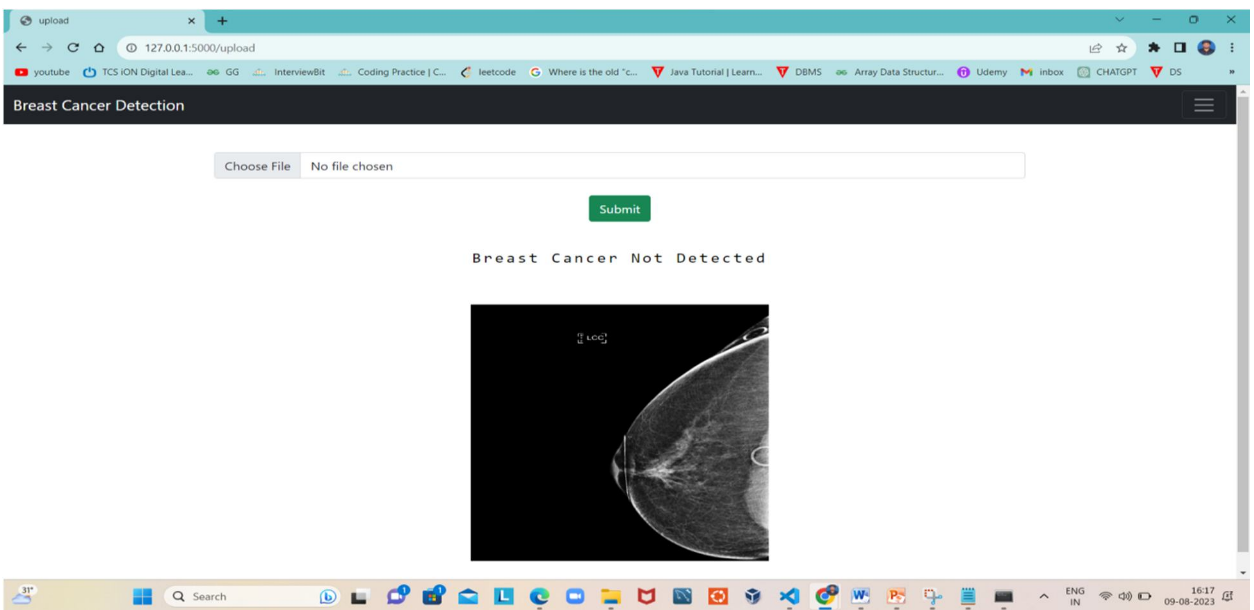


Figure 3: This figure shows that breast cancer is not detected in the uploaded image.

REFERENCES

[1] Ali Bou Nassif, Manar Abu Talib, Qassim Nasir, Yaman Afadar, Omar Elgendy, “Breast cancer detection using artificial intelligence techniques”in 2021.

[2] F.A. Spanhol, L.S.Oliveria, C. Petitjean, L. Heutte, “Breast cancer histopathological image classification using convolutional neural networks”, in: Proceedings of the International Joint Conference on Neural Networks, 2016-October, 2016, pp. 2560.

[3] Spanhol, F. A., Oliveira, L. S., Cavalin, P. R., Petitjean, C., &Heutte, L. (2017,Oct) “Deep features for breast cancer histopathological image classification”.In 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC) (pp. 1868-1873).

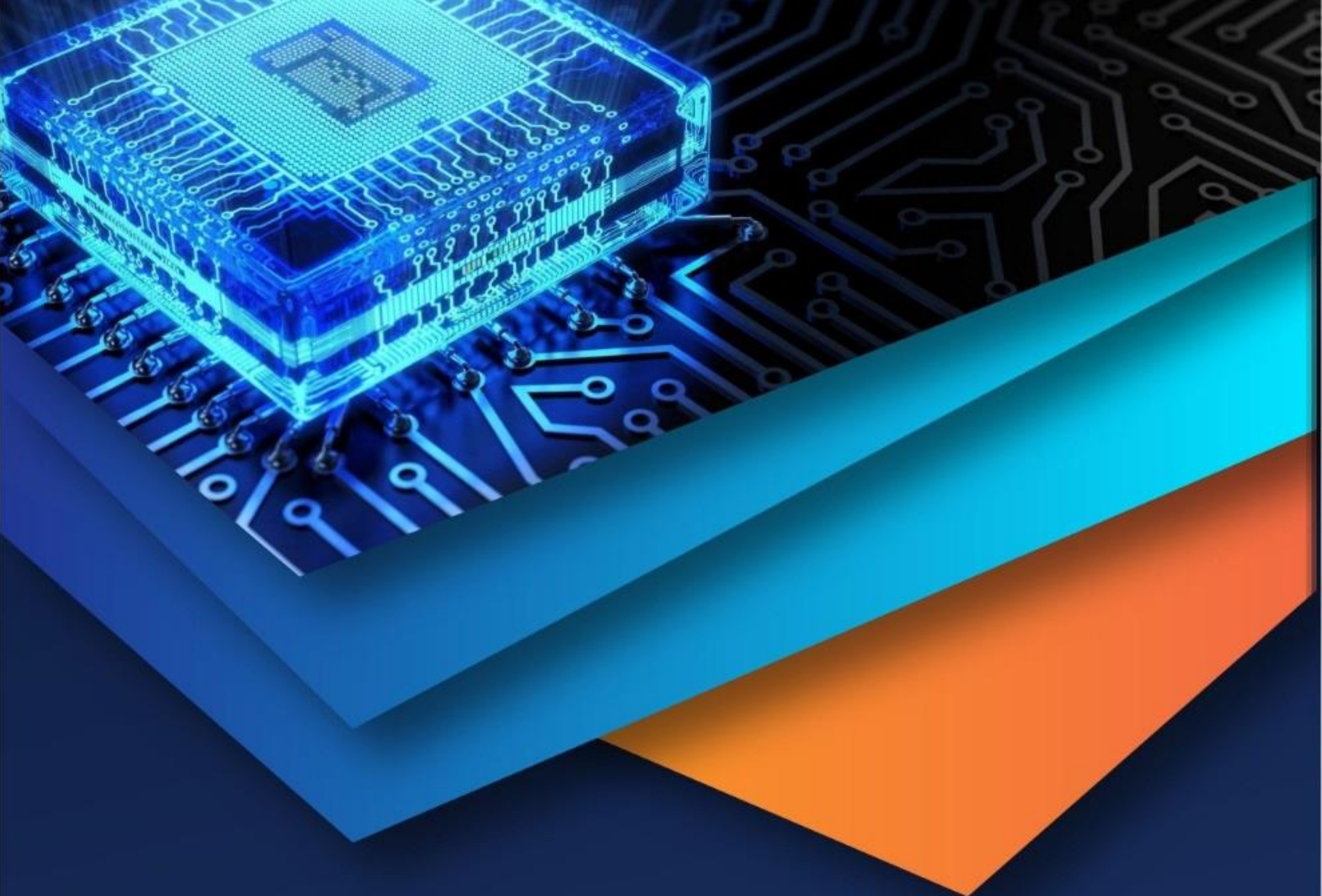
[4] José Rouco, Paulo Aguiar, Catarina Eloy, António Polónia, “Classification of breast cancer histology images using Convolutional Neural Networks” in 2019.

[5] Huang G., Liu Z., Der Maaten V., Weinberger K. Densely connected convolutional networks. 2016.

[6] Moh’d Rasoul Al-hadidi, Abdulsalam Alarabeyyat, Mohannad Alhanahnah, “Breast Cancer Detection using K-nearest Neighbor Machine Learning Algorithm” in 2016.

[7] Yu-Dong Zhang, Chichun Pan, Xianquing Chen, Fubin Wang, “Abnormal breast identification by nine-layer convolutional neural network with parametric rectified linear unit and rank-based stochastic pooling” in 2018.

[8] Nalini Singh, Ambarish G Mahapatra, Gurukalyan Kanungo, “Breast Cancer Mass Detection in Mammograms using K-means and Fuzzy C-means Clustering” in 2011.



10.22214/IJRASET

INDEX COPERNICUS 45.98

ISRA JIF IMPACT FACTOR: 7.129

TOGETHER WE REACH THE GOAL

IMPACT FACTOR: 7.429

ISI

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY