e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

BREAST CANCER CLASSIFICATION WITH A SIMPLE NEURAL NETWORK (NN) AND DEEP LEARNING

Satishkant Ram*1

*1UG, Department Of Information Technology, B.K Birla College Kalyan, (Empowered Autonomous Status), India.

ABSTRACT

Breast cancer is indeed a significant and widespread health concern that affects a large number of people, particularly women. This research paper presents a simple neural network-based approach for the classification of breast cancer tumors into benign and malignant categories using dataset analysis. The proposed neural network architecture enhances the diagnostic process by utilizing a publicly available dataset of breast cancer preprocessing. Techniques are applied to prepare the data for the neural network model. The approach presented in this paper offers a solution for healthcare theorists and researchers as it combines the accessibility of a straightforward neural network with the potential to make a substantial impact in the fight against breast cancer however at that time deep learning performed well in various tasks of computer vision and pattern recognition, but it still has some challenges. One of the main challenges is the lack of training data to address this challenge and optimize performance. We have utilized a transfer learning technique which is where the deep learning models.

Keywords: Neural Network And Deep Learning, Breast Cancer Classification And Analysis.

I. INTRODUCTION

Breast cancer is a significant global health concern, affecting around 1.7 million women annually and ranking as a leading cause of cancer-related deaths. In the United States alone, there were approximately 268,600 new cases of invasive breast cancer and 62,930 new cases of in-situ breast cancer diagnosed in 2019, resulting in approximately 41,760 deaths from breast cancer. Early diagnosis is crucial for improving survival rates and reducing the financial and health burdens associated with breast cancer.

Breast tumors are generally categorized as malignant (cancerous) or benign (non-cancerous) based on various cell features, with subtypes and distinct properties within each category. Benign tumors typically pose no significant health risks but can represent variations in breast tissue structure. Malignant tumors are further classified as in-situ or invasive, depending on whether cancer cells are confined to ducts or lobules within the breast or have spread beyond these structures.

Diagnostic methods for breast cancer include self-assessment (palpation), mammography, ultrasonic imaging, and needle biopsy, with the latter being a reliable technique for determining the likelihood of malignant tissue growth. Pathologists play a critical role in assessing breast tissue microscopically, aided by the Hematoxylin and Eosin (H&E) staining process, which enhances the visualization of tissue elements and structures. The pathologists' analysis involves evaluating cell distribution, architectural patterns, tissue staining variations, cell shape regularities, density, and nuclei organization. The need for precise breast cancer diagnosis, given its high costs and morbidity rates, has spurred research into developing more accurate diagnostic models. This research paper will likely explore such models and their potential for improving breast cancer diagnosis and patient outcomes. Another effective approach employed by pathologists is biopsy, where a tissue sample from the affected area of the breast is collected and microscopically analyzed to detect and classify tumors. Biopsy is regarded as a proficient and precise method for breast cancer detection and classification. Automated and accurate methods for histopathological image analysis are essential for breast cancer diagnosis and early detection. They can reduce the workload on pathologists and improve diagnostic precision. The proper screening of histopathological images is vital for advancing breast cancer classification systems and ultimately enhancing early detection and diagnosis.

II. LITERATURE REVIEW

Breast cancer classification using neural network (NN) models is an important application of artificial intelligence and machine learning in healthcare. Neural networks, specifically deep learning models, have shown promising results in accurately diagnosing and classifying breast cancer from medical imaging data, such as mammograms and biopsy images. Here's an overview of the process.

The author addresses the limitations of traditional breast cancer classification techniques, which are primarily reliant on handcrafted features and performance that is highly dependent on the selected features. Additionally, these methods exhibit high sensitivity to variations in size and intricate shapes. They also talk about the challenges of pattern recognition of breast cancer classification. One of the main challenges is the lack of training data. To address this challenge and optimize the performance.[1]

According to Authors Yi-Sheng Sun and Zhao Zhao, Early detection of breast cancer is a crucial strategy for its prevention. In certain developed nations, the 5-year relative survival rate for breast cancer patients exceeds 80%.[2]

Lifestyle plays an important role in breast cancer. Consuming alcohol can lead to an elevation in the levels of estrogen-related hormones in the bloodstream, thereby activating estrogen receptor pathways. A meta-analysis, encompassing 53 epidemiological studies, revealed that the risk of breast cancer can increase by 32% with a daily alcohol intake of 35-44 grams, and for each additional 10 grams of alcohol per day, there is a 7.1% increase in the relative risk.[3]

Introduced a breast cancer image classification approach employing ensembles of deep learning models. They utilized pre-trained VGG16 and VGG19 architectures to train four distinct models, achieving an impressive overall accuracy of 95.29% across various breast cancer image classes [4]. Meanwhile, Gupta et al. conducted breast cancer image classification, distinguishing between malignant and non-malignant tumors. They employed a range of supervised machine learning algorithms, including k-nearest Neighborhood, Logistic Regression, Decision Tree, Random Forest, and SVM. By incorporating Adam Gradient Descent Learning models, they achieved a notable accuracy rate of 98.24%[5]

Common machine learning libraries, such as TensorFlow and Keras, or deep learning frameworks like PyTorch, can be used to implement and train these neural network models for breast cancer classification.

It's important to note that more complex neural network architectures, like convolutional neural networks (CNNs) or recurrent neural networks (RNNs), can also be applied to breast cancer image analysis tasks where the input data includes histopathological images or medical imaging data.

The choice of the model and architecture should be based on the nature of the data and the specific goals of the breast cancer classification task. Researchers typically experiment with different architectures and configurations to determine which one performs best for their dataset and objectives.

III. METHODOLOGY

Dataset

The Breast Cancer Wisconsin (Diagnostic) Data Set is an extensively used dataset in the field of machine learning research. It comprises 569 instances of breast cancer tissues, each of which is characterized by 30 features. These features are derived from a digitized image of a fine needle aspiration (FNA) of a breast mass, and they describe the characteristics of the cell nuclei present in the image. The objective of the dataset is to predict whether the cancer is benign or malignant. The dataset is well-balanced, with 212 malignant cases and 357 benign cases, providing a comprehensive representation of both types of cases. . I have used Google Collaboratory where I have used the CNN tensorflow library and Keras.

PROPOSED MODEL

A. MLP Model

A multilayer perceptron (MLP) is a type of artificial neural network (ANN) that processes information in a feedforward manner. It consists of multiple layers of neurons arranged in a directed graph, with each layer fully connected to the next layer. The MLP is a supervised learning algorithm that is trained on labeled datasets. During the training phase, the weights of the connections between the neurons are adjusted so that the MLP can accurately predict the labels of new examples. The ultimate goal of the MLP is to provide accurate predictions for various applications.

A common and straightforward neural network model that can be used for breast cancer classification is a basic feedforward neural network. This type of neural network consists of an input layer, one or more hidden layers, and an output layer. The number of neurons in each layer, the activation functions, and the architecture can be customized based on the specific requirements of the task.

- **Input layer:** Neurons representing input features (e.g., tumor size, age, histological features).
- **Hidden Layer:** One or more hidden layers with neurons that perform computations and learn.
- **Output Layer:** A single neuron for binary classification (benign or malignant) or multiple neurons for multi-class classification

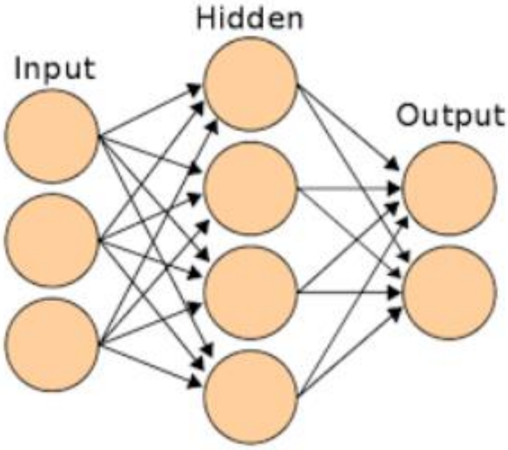

Figure 1. Multilayer Perceptron (MLP) Neural Networks

B. Logistic regression

Logistic regression is a statistical technique used to predict the probability of a binary outcome, such as 'yes' or 'no', 'true' or 'false', or '0' or '1'. This is done by fitting the available data to a logistic function - a sigmoid curve that models the probability of the event occurring.

IV. RESULTS AND DISCUSSION

An artificial neural network can be viewed as an extension of the logistic regression model, where additional layers of feature interactions are added. An artificial neural network can indeed be seen as an extension of the logistic regression model, where additional layers of feature interactions are incorporated. Logistic regression is a relatively simple linear model, and its accuracy may be limited when dealing with complex data. In contrast, artificial neural networks, with their ability to capture non-linear relationships through multiple layers, often achieve higher precision and minimize data loss in more intricate tasks.

The model was trained using a deep learning classification approach, with the data like Benign and Malignant. the model is trained with 50 epochs. A training-validation split to monitor the model's performance and prevent overfitting. The loss function, optimizer, and learning rate were selected through experimentation to achieve optimal results.

In the case of the logistic regression model, after 50 epochs of training, the accuracy reaches 94.94%. However, when transitioning to an artificial neural network (ANN) model, the accuracy notably improves to 97.37%, demonstrating the enhanced performance of the ANN in comparison to the logistic regression model.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:10/October-2023 Impact Factor- 7.868 www.irjmets.com

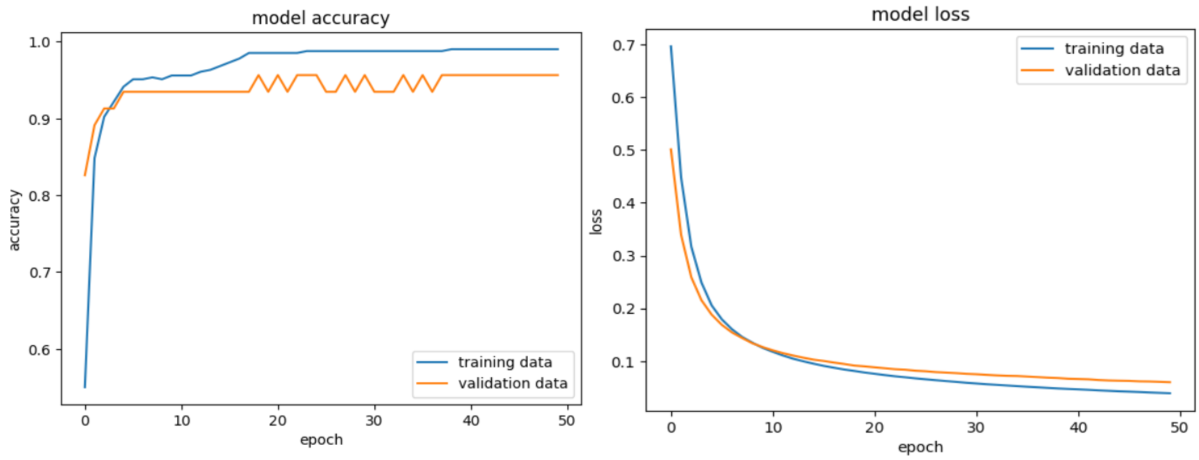

Figure 2. Accuracy and loss of ANN Model

V. CONCLUSION

Breast cancer is a significant and prevalent health concern, especially affecting women. This research paper introduces a straightforward neural network-based approach to classify breast cancer tumors as benign or malignant by analyzing a publicly available dataset. The proposed neural network architecture streamlines the diagnostic process and prepares the data for model training effectively. This approach presents a promising solution for healthcare practitioners and researchers, combining the simplicity of a neural network with the potential to have a meaningful impact in the battle against breast cancer. It's worth noting that while deep learning has shown remarkable performance in various computer vision and pattern recognition tasks, it still faces challenges, such as the scarcity of training data. To address this challenge and optimize performance, the paper employs transfer learning, leveraging pre-trained deep learning models. This research represents a significant step forward in the ongoing effort to improve breast cancer diagnosis and underscores the importance of adapting cutting-edge technology to address critical healthcare issues.

VI. REFERENCES

[1] Alzubaidi, L., Al-Shamma, O., Fadhel, M. A., Farhan, L., Zhang, J., & Duan, Y. (2020). Optimizing the performance of breast cancer classification by employing the same domain transfer learning from hybrid deep convolutional neural network model. Electronics, 9(3), 445.

[2] Sun, Y. S., Zhao, Z., Yang, Z. N., Xu, F., Lu, H. J., Zhu, Z. Y., ... & Zhu, H. P. (2017). Risk factors and preventions of breast cancer. International journal of biological sciences, 13(11), 1387.

[3] Rainey, L., Eriksson, M., Trinh, T., Czene, K., Broeders, M. J., van der Waal, D., & Hall, P. (2020). The impact of alcohol consumption and physical activity on breast cancer: the role of breast cancer risk. International journal of cancer, 147(4), 931-939.

[4] Hameed, Z., Zahia, S., Garcia-Zapirain, B., Javier Aguirre, J., & Maria Vanegas, A. (2020). Breast cancer histopathology image classification using an ensemble of deep learning models. Sensors, 20(16), 4373.,

[5] Ghrabat, M. J., Hussien, Z. A., Khalefa, M. S., Abduljabba, Z. A., Nyangaresi, V. O., Al Sibahee, M. A., & Abood, E. W. (2022). Fully automated model on breast cancer classification using deep learning classifiers. Indonesian Journal of Electrical Engineering and Computer Science, 28(1), 183-191.