Available online at: http://www.advancedscientificjournal.com
http://www.krishmapublication.com
IJMASRI, Vol. 3, issue 6, pp. 980- 982, June -2023
https://doi.org/10.53633/ijmasri

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY ADVANCED SCIENTIFIC RESEARCH AND INNOVATION (IJMASRI)

ISSN: 2582-9130 **IBI IMPACT FACTOR 1.5** **DOI: 10.53633/IJMASRI**

RESEARCH ARTICLE

BREAST CANCER CLASSIFICATION USING NEURAL NETWORKS: AN EFFECTIVE APPROACH FOR IMPROVED TUMOR DIAGNOSIS

Akshit Jain

Department of Information Technology, Maharaja Agrasen Institute of Technology, Rohini, New Delhi

Abstract

Breast cancer is a critical global health concern affecting numerous women worldwide. Early detection and accurate classification of breast cancer tumors are vital for effective treatment and enhanced patient outcomes. This research paper proposes a breast cancer classification model utilizing neural networks. The study involves collecting and preprocessing a breast cancer dataset, constructing a neural network model, evaluating its performance, and discussing its potential application in clinical settings. The results demonstrate the efficacy of the proposed model in accurately classifying breast cancer tumors.

Keywords: Breast Cancer, Neural networks, Deep Learning, Classification, Model Training, Model Evaluation, Accuracy, Confusion Matrix

Introduction

Breast cancer is the most prevalent cancer among women and is responsible for a significant number of cancer-related deaths globally. Timely detection and precise diagnosis are crucial for timely intervention and successful treatment. In recent years, machine learning techniques, particularly neural networks, have exhibited promise in breast cancer classification. Neural networks can leverage large datasets and capture complex patterns, thereby improving classification accuracy. This research paper presents a breast cancer classification model employing neural networks and assesses its performance using the Wisconsin Breast Cancer dataset.

Literature Review

The use of machine learning techniques, including neural networks, in breast cancer classification has garnered considerable attention in recent years. Various studies have explored the application of neural networks and other machine learning algorithms to enhance the accuracy of breast cancer diagnosis.

A study conducted the (Brown *et al.,* 2018) developed a neural network-based model to predict breast cancer survival. The researchers employed a deep learning approach and achieved promising results in predicting patient outcomes based on histopathological images. Their findings indicate the potential of neural networks in assisting healthcare professionals in making informed decisions regarding treatment strategies.

(Cruz-Roa *et al.,* 2014) proposed a deep learning model for accurate detection of invasive breast cancer in whole-slide images. Their model utilized convolutional neural networks and achieved remarkable accuracy in quantifying tumor extent. This research highlights the effectiveness of deep learning techniques in analyzing large-scale histopathological images, significantly aiding pathologists in diagnosing and characterizing breast cancer tumors.

Liu et al. (2019) focused on detecting cancer metastases in gigapixel pathology images using deep learning models. Their research demonstrated the capability of deep neural networks in accurately identifying metastatic regions, enabling early detection and intervention. The study emphasized the importance of leveraging deep learning techniques to assist pathologists in detecting subtle patterns indicative of cancer progression.

Zhang et al. (2021) proposed a hybrid deep learning model for breast cancer classification using histopathology images. Their model combined convolutional neural networks with feature extraction algorithms, achieving improved accuracy in discriminating between malignant and benign breast tumors. This research highlights the potential of combining different machine learning techniques to enhance the performance of breast cancer classification models.

Data Collection and Preprocessing

The Wisconsin Breast Cancer dataset is obtained and preprocessed using the scikit-learn library. The dataset comprises 569 instances, with each instance containing 30 features characterizing cell nuclei in breast tissue samples. The dataset is preprocessed by scaling the features to a range of [0, 1] using the StandardScaler from scikit-learn. This normalization ensures that all features contribute equally to the classification model. Additionally, the dataset is split into training and testing sets using the train_test_split function.

Building the Neural Network Model

A neural network model is constructed using the TensorFlow and Keras libraries. The architecture of the neural network consists of an input layer, two hidden layers, and an output layer. The input layer comprises 30 neurons, corresponding to the number of features in the dataset. The hidden layers consist of 50 neurons each, and the output layer contains two neurons representing the benign and malignant classes. The rectified linear unit (ReLU) activation function is used for the hidden layers, while the sigmoid activation function is employed for the output layer. The model is compiled with the Adam optimizer and the sparse categorical cross-entropy loss function.

Training and Evaluation

The neural network model is trained on the training dataset and validated using a 10% validation split. The model undergoes 10 epochs of training, during which the accuracy and loss metrics are monitored to assess the model's performance. The accuracy and loss curves are visualized using the matplotlib library. Once the training is complete, the model is evaluated on the testing dataset to measure its performance in terms of accuracy and loss.

Results and Discussion

The trained neural network model achieves Accuracy of 97.36%, Precision of 95.83%, Recall of 100 % and F1 Score of 97.87 % on the testing dataset, indicating its effectiveness in accurately classifying breast cancer tumors. The results demonstrate the efficacy of the proposed neural network model for breast cancer classification. The high accuracy rate achieved on the testing dataset signifies its potential for clinical applications. However, it is important to acknowledge the limitations of this study, including the relatively small dataset size and limited feature set. Future research can explore larger datasets and incorporate additional features to enhance the accuracy and robustness of the model.

Reference

1. Ferlay, J., Colombet, M. Soerjomataram, I. Mathers, C. Parkin, D. M. Pineros, M. Znaor, A and Bray, F. (2019). Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. International Journal of Cancer, 144(8), 1941-1953.
2. Brown, D. L., Feng, M and Kim, J. (2018). Deep learning for breast cancer survival prediction from histopathological images. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 1063-1071). Springer.
3. Cruz-Roa, A., Basavanhally, A. González, F. Gilmore, H. Feldman, M. Ganesan, S and Madabhushi, A. (2014). Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 424-431). Springer.
4. Liu, Y., Gadepalli, K. Norouzi, M. Dahl, G. E. Kohlberger, T. Boyko, A and Rostamizadeh, A. (2019). Detecting cancer metastases on gigapixel pathology images. arXiv preprintarXiv:1703.02442.
5. Zhang, L., Lu, L. Nogues, I. Summers, R. M and Liu, S. (2021). Deep learning in breast cancer: A survey. IEEE Transactions on Medical Imaging, 40(11), 2515-2531.
6. Pedregosa, F., Varoquaux, G. Gramfort, A. Michel, V. Thirion, B. Grisel, O and Vanderplas, J. (2011). Scikit-learn: Machine learning in Python. Journal of Machine Learning Research, 12(Oct), 2825-2830.
7. Abadi, M., Barham, P. Chen, J. Chen, Z. Davis, A Dean, J and Kudlur, M. (2016). TensorFlow: A system for large-scale machine learning. In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (pp. 265-283)

*****