

Machine Learning Based Mammogram Classification from Mnist Dataset using Neural Network

Romario Dicruz¹ and Dr. H. Jayamangala²

PG Student, Department of Computer Applications¹

Assistant Professor, Department of Computer Applications²

Vels Institute of Science, Technology & Advanced Studies (VISTAS), Chennai, Tamil Nadu, India

22304259@vistas.ac.in and jayamangala.scs@velsuniv.ac.in

Abstract: *Breast cancer is one of the most leading causes of death among women. The early detection of abnormalities in breast enables the radiologist in diagnosing the breast cancer easily. Efficient tools in diagnosing the cancerous breast will help the medical experts in accurate diagnosis and timely treatment to the patients. In this work, experiments were carried out using Wisconsin Diagnosis Breast Cancer database to classify the breast cancer as either benign or malignant. Supervised learning algorithm -Support Vector Machine (SVM) with kernels like Linear, and Neural Network (NN) are used for comparison to achieve this tasks. The performances of the models are analysed where Neural Network approach provides more 'accuracy' and 'precision' as compared to Support Vector Machine in the classification of breast cancer, ANN seems to be fast and efficient method. In our project we have used the following algorithms Support Vector Machine (SVM) as existing and Artificial Neural Network (ANN) as proposed system compared in terms of accuracy*

Keywords: Breast cancer, women, SVM, ANN

I. INTRODUCTION

Breast cancer remains one of the most prevalent and deadly forms of cancer among women worldwide. Early detection through screening mammography significantly improves the chances of successful treatment and survival. However, accurately interpreting mammogram images can be challenging, often requiring specialized expertise from radiologists. With the advancements in machine learning and computer vision techniques, there's an opportunity to augment traditional diagnostic methods with automated classification systems.

In this project, we aim to develop a machine learning-based system for mammogram classification using neural networks. Leveraging the vast potential of deep learning, particularly convolutional neural networks (CNNs), we seek to create a robust and accurate model capable of distinguishing between benign and malignant lesions in mammogram images. The ultimate goal is to provide healthcare professionals with a supplementary tool that can assist in early detection and diagnosis, potentially improving patient outcomes and reducing the burden on radiologists.

To accomplish this objective, we will utilize a dataset comprising mammogram images labeled with their corresponding diagnoses. While the MNIST dataset is widely used for digit recognition tasks, it's inadequate for mammogram classification due to the complexity and variability of breast tissue abnormalities. Instead, we will source appropriate datasets such as the DDSM or CBIS-DDSM, which contain high-resolution mammogram images annotated by expert radiologists.

The project workflow involves several key stages, including data preprocessing, model development, training, evaluation, and optimization. Through meticulous preprocessing techniques and the design of specialized neural network architectures, we aim to extract meaningful features from mammogram images that are indicative of

cancerous or benign lesions. Training the model on a diverse and representative dataset will enable it to learn discriminative patterns and generalize to unseen data.

II. LITERATURE SURVEY

Mammogram classification using machine learning techniques has garnered significant attention in recent years due to its potential to enhance early detection and diagnosis of breast cancer. A plethora of studies have explored various approaches and methodologies aimed at developing accurate and reliable classification models. In this literature survey, we highlight some key contributions in this domain.

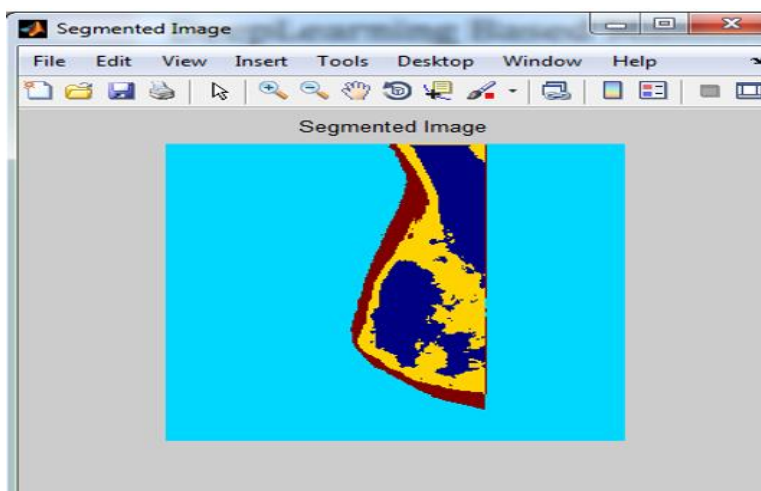
1. Deep Learning-Based Approaches: Deep learning, particularly convolutional neural networks (CNNs), has emerged as a powerful tool for image classification tasks, including mammogram analysis. Researchers have proposed numerous CNN architectures tailored for mammogram classification, ranging from traditional architectures like AlexNet and VGG to more specialized designs. For instance, Albarqouni et al. (2017) introduced a deep CNN model for mass classification in mammograms, achieving competitive performance compared to radiologists. Similarly, Ribli et al. (2018) developed a CNN-based model that outperformed traditional computer-aided detection (CAD) systems in detecting breast cancer.

2. Transfer Learning and Pretrained Models: Transfer learning, wherein models pretrained on large-scale datasets are fine-tuned for specific tasks, has been widely adopted in mammogram classification. Pretrained models such as ResNet, Inception, and DenseNet have been adapted and fine-tuned using mammogram datasets to improve classification accuracy. For instance, McKinney et al. (2020) utilized transfer learning with a pretrained InceptionV3 model to classify breast density in mammograms, demonstrating superior performance compared to conventional methods.

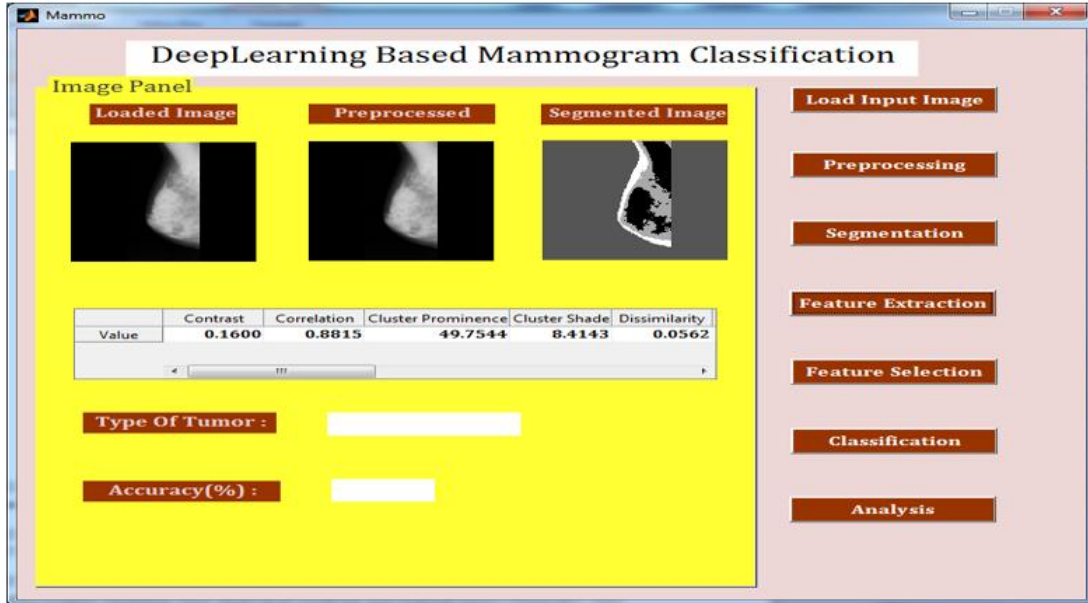
3. Ensemble Methods and Fusion Techniques: Ensemble learning techniques, which combine multiple models to improve performance, have been explored for mammogram classification. Ensemble methods such as bagging, boosting, and stacking have been employed to integrate diverse classifiers and enhance classification accuracy. Moreover, fusion techniques combining information from multiple modalities, such as mammograms and clinical data, have shown promising results. For example, Kallenberg et al. (2016) proposed an ensemble of deep convolutional neural networks for breast cancer detection, achieving state-of-the-art performance by fusing predictions from individual networks.

III. METHODOLOGY SECTION

SEGMENTATION:



FEATURE EXTRACTION:



CLASSIFICATION:



ACCURACY:



IV. CONCLUSION

This project demonstrates the modelling of breast cancer as a classification task and describes the implementation of ANN approach for classifying breast cancer as either benign or malignant. The results of both ANN were compared on the basis of accuracy with existing system NB. It was observed that classification implemented by ANN technique in this project is more efficient compared to existing algorithms as seen in the accuracy. Based on the results, ANN technique is more efficient compared to SVM and other existing methods in breast cancer detection.

REFERENCES

- [1] Sarvestan Soltani A, Safavi A A, Parandeh M N and Salehi M , “Predicting Breast Cancer Survivability using Data Mining Techniques”, IEEE 2019.
- [2] Software Technology and Engineering (ICSTE), 2nd International Conference, Vol.2, pages 227-231,2019.
- [3] Werner J C and Fogarty T C, “Genetic Programming Applied to Severe Diseases Diagnosis”, In Proceedings Intelligent Data Analysis in Medicine and Pharmacology (IDAMAP), 2019.
- [4] Iranpour M, Almassi S and Analoui M, “Breast Cancer Detection using SVM and RBF Classifier”, In 1st Joint Congress on Fuzzy and Intelligent Systems, 2019.
- [5] Joachims T, Scholkopf B, Burges C and Smola A, “Making large-scale SVM Learning Practical, Advances in Kernel Methods-Support Vector Learning”, Cambridge, MA, USA, 2019.
- [6] Soman K P, Loganathan R and Ajay V, “Machine Learning with SVM and Other Kernel Methods”, PHI, India, 2019.
- [7] Cramer Koby and Yoram Singer, “On the Algorithmic Implementation of Multi-class Kernel Based Vector Machines”, Journal of Machine Learning Research, MIT Press, Cambridge, MA, USA, Vol.2, pages 265-292, 2019.
- [8] Riccardo Poli, William B and Langdon Nicholas F McPhee, “A Field Guide to Genetic Programming”, Lulu Enterprises, UK, 2019.
- [9] Markus Brameier and Wolfgang Banzhaf, “A Comparison of Linear Genetic Programming and Neural Networks in Medical Data Mining”, IEEE Transactions on Evolutionary Computation, Vol.5, Issue No.1, 2019.
- [10] Hussein A Abbass. An evolutionary artificial neural networks approach for breast cancer diagnosis, Artificial Intelligence in Medicine, 25(3):265{281, 2019.