IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

BREAST CANCER DETECTION USING AUTOENCODER

Ms. N. V. S. Suma Varsha[1], Ms. K. Jahnavi[1], Ms. N. Sai Lavanya[1], Mr. MD. Rayan[1]
[1]B. Tech Final Year Student Lendi Institute of Engineering and Technology (A),
Jonnada, Vizianagaram, Andhra Pradesh, India

Abstract: Breast cancer stands as the predominant ailment among females, with a prevalence of 2.1 million cases annually, leading to over 70,000 fatalities globally. Convolutional Neural Networks (CNNs) have emerged as pivotal tools for image classification tasks, demonstrating proficiency in analyzing breast cancer images. While conventional feature extraction methods have been widely employed, the utilization of CNNs for both feature extraction and classification has garnered attention. In pursuit of heightened classification accuracy and efficiency, Convolutional Autoencoders (CAEs) are integrated to extract intricate features and reduce dimensionality, subsequently augmenting CNN-based classification. This project aims to enhance classification accuracy by integrating feature extraction mechanisms, namely CAEs with dimensionality reduction, with CNNs, across diverse breast cancer image modalities.

Keywords – Convolutional Neural Networks, Convolutional Autoencoders, Dimensionality reduction.

I. INTRODUCTION

Universally, cancer is considered as the unsafe infections. There are a few screening tests and preparatory conclusion of cancer, the conclusive conclusion is made by examining a biopsy test of the suspicious tissue. The common signs and indications can be watched in patients with different cancers: misfortune of weight, Energy, sporadic dying, repetitive hack or alter in voice and fever. The guess of cancer ranges from amazing to destitute. It depends on the stages of the cancer.

(1) Convolution Neural Network

In picture categorization and acknowledgment, Convolutional Neural Systems (CNNs) are utilized. CNNs take pictures as input, handle them through a arrangement of convolution layers with channels, pooling layers, completely associated layers, and at last SoftMax enactment to allot probabilities to different categories. Each picture experiences control through convolution operations, where channels extricate highlights, taken after by pooling to diminish dimensionality and completely associated layers for classification.

a) Convolutional Layer: Convolutional layers use filters to extract features from input images and computing dot products. Each filter learns to detect specific patterns like edges, corners, textures etc. Multiple filters create feature maps representing different features.

b) Pooling: These layers are used to reduce the spatial dimensions of the feature maps which are obtained from the convolutional layers while preserving important information. Max pooling which selects the maximum value from each local region of the feature map. This helps in preserving important features while reducing computational complexity and avoiding overfitting.

c) Fully Connected Layer: Fully Connected layers joining each neuron in one layer to each neuron in the following layer. They are found at the conclusion of the organize and capable for high-level include learning and classification. Neurons in these layers get input from all neurons in the past layer, empowering them to learn complex designs and connections

(2) Autoencoder

Autoencoders are a distinct neural network variant where the input matches the output, focusing on data reconstruction. Trained without supervision, they aim to grasp fundamental, intricate features within the input. These compressed representations are then expanded to recreate the original data, enhancing understanding of its underlying structure. Autoencoders excel in data compression and denoising, proficiently extracting vital insights while generating precise reconstructions.

Types of Autoencoder are given below

1. Denoising Autoencoder: Denoising autoencoders are specialized neural networks designed to reconstruct clean data from noisy input. They operate by learning to filter out unwanted noise while preserving essential features during training. By corrupting the input data with noise and comparing it to the clean version, denoising autoencoders effectively learn to denoise data.
2. Sparse Autoencoder: Sparse autoencoders introduce sparsity constraints on the hidden layer activations. These constraints encourage the network to learn a compressed representation of the input data, where only a small subset of neurons is active at a time. It effectively learns robust features and reduce redundancy in the given data.
3. Deep Autoencoder: These networks consist of multiple encoding and decoding layers, enabling them to capture hierarchical features in the data. These can effectively compress and reconstruct high-dimensional input data, making them suitable for tasks such as feature learning, data compression, and dimensionality reduction in large datasets.
4. Convolutional Autoencoder: They are able to encode information into a set of signals, aim to reconstruct the input from those signals, and change the image's structure. These are the most advanced methods for convolutional filter unsupervised learning. These filters can be used to extract features from any input once they have been trained. Classification can be done using these features.

II. PROPOSED SYSTEM

(1) Autoencoder

In breast cancer detection, autoencoders are able to learn different features from medical images without requiring manual feature extraction. The autoencoder consists of an encoder network, which is able to compresses the input image into a low-dimensional latent space representation, and decoder network, which reconstructs the original image from this compressed representation. During training, the autoencoder learns to minimize the reconstruction error between the input and output images and capturing relevant features that are important for differentiating between cancerous and non-cancerous tissues.

These improve the classification accuracy. The image dataset is fed into the autoencoder and the output of the autoencoder is fed into the input of the Convolution Neural Network. The autoencoder takes the relevant features from the given image as an input and the classification accuracy is observed.

(2) Convolutional Autoencoder

These are able to learn how to decode the input into a series of signals and then reconstruct the input by changing the image's dimensions. Convolutional filters are learned which are used on any input only after the training to extract the features. These features are applied to any of the tasks which requires accurate representation of the input like classification.

III. IMPLEMENTATION

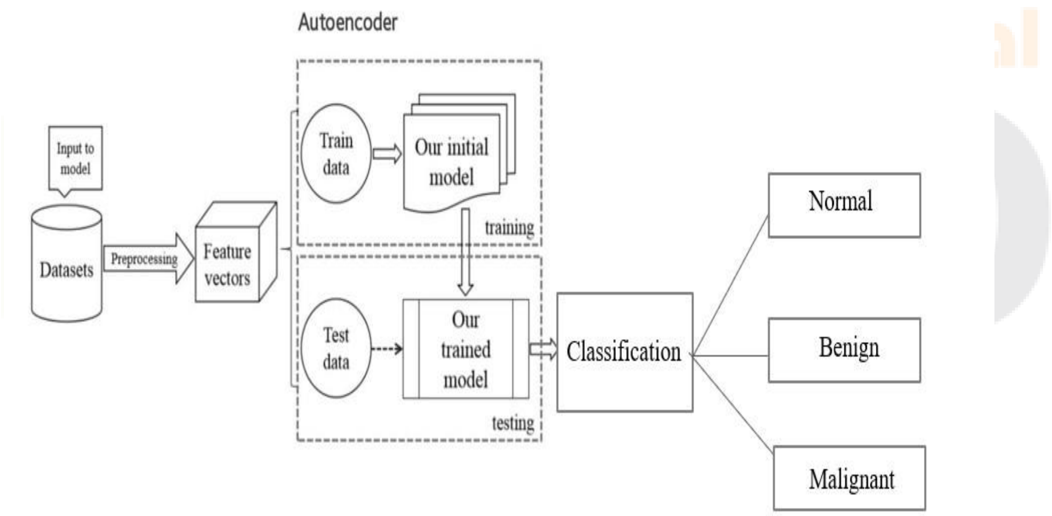

Fig.1) System Architecture for Breast Cancer Detection Using Autoencoder

Steps to implement the project:

Step 1: Upload Dataset: Users can upload a dataset containing Breast images. This dataset likely consists of normal, benign and malignant images

Step 2: Dataset Preprocessing: The uploaded dataset is preprocessed to extract features and labels for training deep learning model. This preprocessing step may involve loading images, extracting relevant features and preparing the data for training.
Step 3: Run DL Model (Autoencoder): Train a deep learning model using raw features extracted from the dataset. The DL model architecture involves convolutional layers and Autoencoder.
Step 4: Classification Model (CNN): Design and train a classification model using CAEs that incorporates the extracted features for accurate classification of breast cancer images.
Step 5: Detection and Prediction: Users can provide a test Breast images for prediction. The DL model predicts whether the given image contain cancer or not based on its learned features.
Step 6: Deploy System: Once trained and evaluated, the system can be deployed for real-world applications.

Algorithm Description

(1) Autoencoder Algorithm:

Encoder: The encoder part of the autoencoder is responsible for transforming the input image into a latent representation, also known as a bottleneck layer. This transformation typically involves reducing the dimensionality of the input image while extracting important features.
Here, we define the encoding layers of the autoencoder. We use convolutional layers with ReLu activation followed by max-pooling layers. These layers help in extracting features and reducing the spatial dimensions of the input image.
Activation functions:
1.ReLu:

- It is defined as f(x)=max (0, x), making it easy to compute and implement.
- It introduces non-linearity into the model, enabling it to learn complex patterns and representations.
- It outputs zero for all negative inputs, which can lead to sparse activations, reducing computational complexity.
- For positive inputs, the gradient is 1, which helps avoid the vanishing gradient problem, making training faster and more efficient.

2.Sigmoid:

- It takes any input value between 0 and 1.
- The sigmoid function is differentiable, which is crucial for gradient-based optimization methods.
- It gives the output values in the range (0, 1), which it suitable for binary classification problems.
- Mathematically, Sigmoid is defined as f(x)=1/1+e^(-x).

Decoder: It contains the decoding layers. A fully connected layer helps to design the encoded representation back to the original shape and then it reshapes it to the input shape using Reshape.

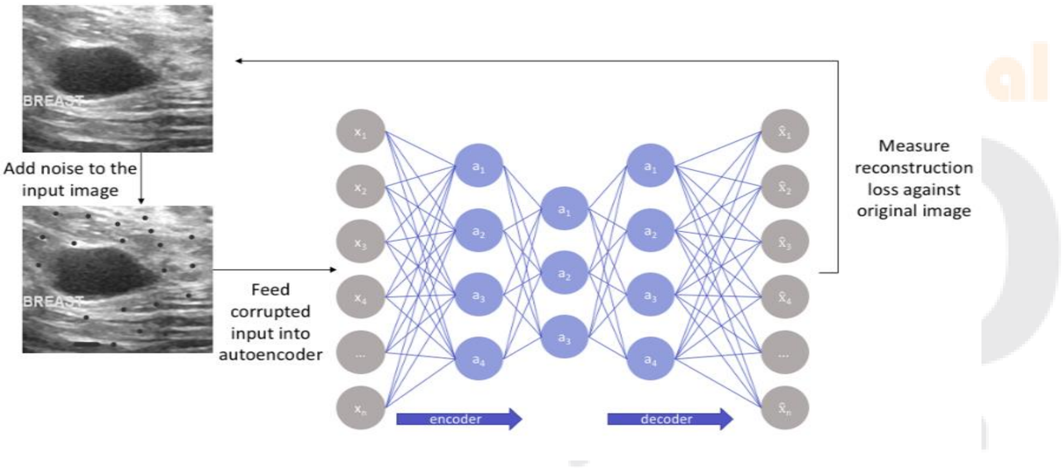

Fig.2) Autoencoder Architecture

(2) CNN Algorithm:

Convolutional Layers:

- Convolution2D: 2D convolution is performed on the input data by these layers. The first convolutional layer has 32 filters of size 3x3. These are used on input images.
- MaxPooling2D: These layers perform max-pooling, discard the unnecessary input representation. It is used to collect necessary features from the previous convolutional layer.
- Flatten Layer: This layer flattens the input, transforming the input into a 1D array.

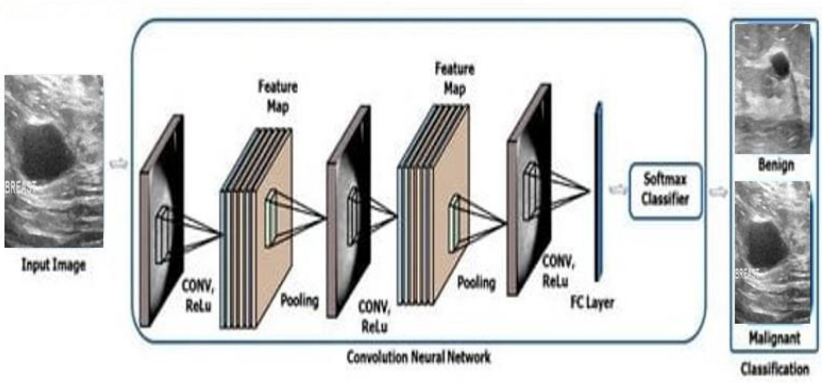

Fig.3) CNN Architecture

IV. RESULTS

The input is given as Breast image with (256,256,3) dimensions and reduced to image size with (32,32,2) dimensions using Autoencoder and CNN and displayed output as Benign cancer.

Input:

Fig. 4) Input image for medical dataset

Output:

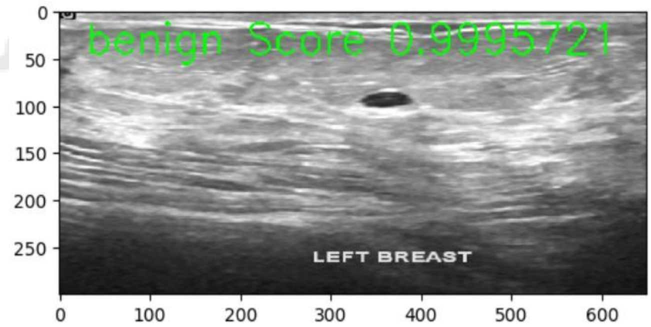

Fig. 5) Output image as Benign Breast Cancer

we got the result with 99% accuracy and the below shown figure shows predicted labels and prediction count.

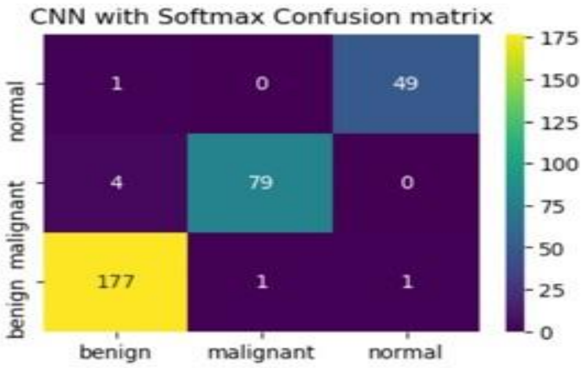

Fig. 6) Confusion Matrix

The below figure shows whether the predictions or correct or not based on the blue and yellow lines arrangement.

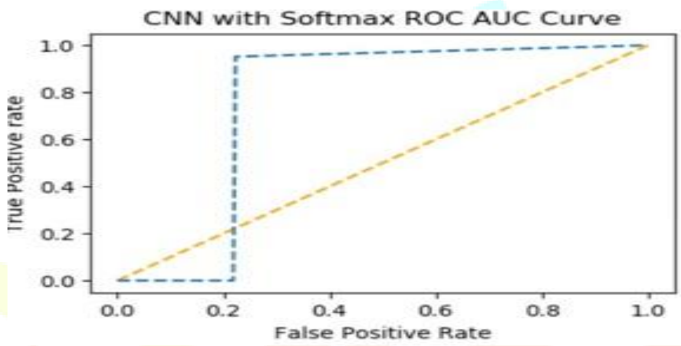

Fig. 7) ROC AUC Curve

The training and validation accuracy of the input data of images is shown below.

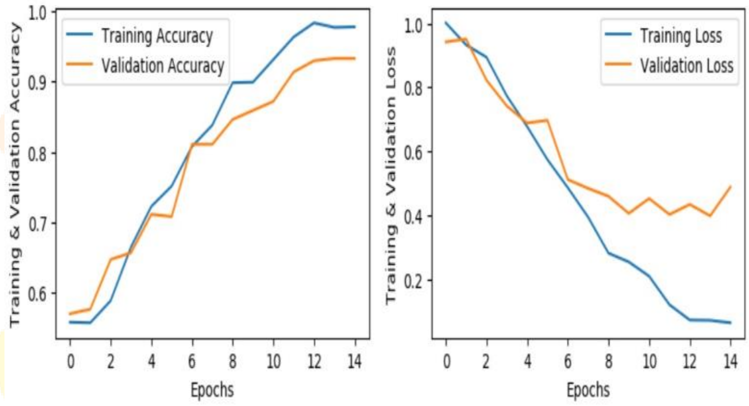

Fig. 8) Training and validation Graph

V. CONCLUSION

In conclusion, the integration of Convolutional Autoencoders (CAEs) with Convolutional Neural Networks (CNNs) presents a promising approach for enhancing the accuracy and efficiency of breast cancer image classification. By leveraging CAEs for feature extraction with CNNs, this project has demonstrated the potential to capture intricate patterns within diverse breast cancer image modalities. The combined framework not only improves classification accuracy but also facilitates more efficient processing of large-scale datasets, thereby contributing to earlier detection.

VI. REFERENCES

[1] Poonam Kathale, Snehal Thorat (2020): “Breast Cancer Detection and Classification”.in 2020 International Conference on Emerging Trends in Information Technology and Engineering IEEE.

[2] S¸tefan Nit¸ica, Gabriela Czibula and Vlad-Ioan Tomescu: “A comparative study on using unsupervised learning-based data analysis techniques for breast cancer detection”. 2020 IEEE.

[3] Naderan M., Zaychenko Yu “Using convolutional neural networks for breast cancer diagnosing”. System research and information technologies, 2019.

[4] Yajush Tewari, Eshant Ujjwal, Lalit Kumar: “Breast Cancer Classification Using Machine Learning”.in 2022 Advance Computing and Innovative Technologies in Engineering. IEEE, 2022.

[5] Maheshwar, Gautam Kumar: “Breast Cancer Detection Using Decision Tree, Naïve Bayes, KNN and SVM Classifiers: A Comparative Study”. (ICSSIT 2019) IEEE Xplore.

[6] Hajer Kamel, Dhahir Abdulah, Jamal M.Al-Tuwaijari: “Cancer Classification Using Gaussian Naive Bayes Algorithm”. IEEE, 2019 - (IEC2019).

[7] V. Kumar, B. K. Mishra, M. Mazara, D. N. H. Thanh, and A. Verma (2019): “Prediction of malignant & benign breast cancer: A data mining approach in healthcare applications”.

[8] M. Togac¸ ar, B. Ergen, (2020):“Application of breast cancer ¨ diagnosis based on a combination of convolutional neural networks, ridge regression and linear discriminant analysis using invasive breast cancer images processed with autoencoders”.

[9] Toa, C.K., Elsayed, M. and Sim, K.S., 2022. Deep Residual Learning with Attention Mechanism for Breast Cancer Classification.

[10] Sahu, A., Das, P.K. and Meher, S., 2023. High accuracy hybrid CNN classifiers for breast cancer detection using mammogram and ultrasound datasets. Biomedical Signal Processing and Control, 80, p.104292.