Predicting Breast Cancer Risk from Histopathology Images using Hybrid Deep Learning Classifier

G Sajiv , G. Ramkumar , S. Shanthi , Arunachalam Chinnathambi , Sulaiman Ali Alharbi

PII: S1350-4533(24)00050-X
DOI: https://doi.org/10.1016/j.medengphy.2024.104149
Reference: JJBE 104149

To appear in: *Medical Engineering and Physics*

Received date: 3 January 2024
Revised date: 17 February 2024
Accepted date: 9 March 2024

Please cite this article as: G Sajiv , G. Ramkumar , S. Shanthi , Arunachalam Chinnathambi , Sulaiman Ali Alharbi , Predicting Breast Cancer Risk from Histopathology Images using Hybrid Deep Learning Classifier, *Medical Engineering and Physics* (2024), doi: https://doi.org/10.1016/j.medengphy.2024.104149

Predicting Breast Cancer Risk from Histopathology Images using Hybrid Deep Learning Classifier

Sajiv.G[1], G. Ramkumar[2*], S. Shanthi[3], Arunachalam Chinnathambi[4], Sulaiman Ali Alharbi[5]

[1]Department of Electronics and Communication, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University, Chennai 602105, Tamil Nadu, India **Email: sajivg2003.sse@saveetha.com**

[2]Department of Electronics and Communication, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University, Chennai 602105, Tamil Nadu, India **Email: pgrvlsi@gmail.com**

[3]Department of Electronics and Communication Engineering, CARE College of Engineering, Tiruchirappalli, Tamil Nadu, India **Email: sshanthijj@gmail.com**

[4]Department of Botany and Microbiology, College of Science, King Saud University, PO Box -2455, Riyadh - 11451, Saudi Arabia. **Email: arun.biology@outlook.com**

[5]Department of Botany and Microbiology, College of Science, King Saud University, PO Box -2455, Riyadh - 11451, Saudi Arabia**. Email: sulaiman778@gmail.com**

Corresponding Author: G. Ramkumar (pgrvlsi@gmail.com)

Highlights

- Early breast cancer discovery can lead to successful treatment
- Automated examination of histological images in disease diagnosis
- Diagnosis and detection of breast cancer in clinical practice.
- Automated malignant breast cancer detection
- Multilayer Perceptron with LightGBM classifier based automated system for histopathology image processing

Abstract

Millions of people per year pass away from breast cancer (BC), which is a fatal illness. To deal with this issue more effectively, diagnoses can be made more scalable and less prone to mistakes by developing automated malignant BC detection systems applied on patient's imaging. Not less significant is the fact that this type of research can be expanded to include additional cancer types, having a greater influence on helping to save lives. Recent BC recognition results demonstrate that Convolution Neural Networks (CNN) can outperform manually crafted feature descriptors in terms of recognition rates. However, this higher recognition rate comes at the cost of a more complex system to develop, one that requires more time for training and specialized knowledge to fine-tune the CNN's architecture. The comprehensive understanding of the structural and cellular attributes of tissue samples that histopathological image analysis imparts is regarded as a crucial component in the rapid detection and diagnosis of breast cancer. By means of this analysis, pathologists are capable of discerning aberrant cellular proliferation, ascertaining the existence of malignant tumours, and evaluating the degree of cancer advancement. The utilization of sophisticated imaging technologies and computational algorithms in histopathological image analysis improves the precision and effectiveness of diagnosis, allowing

medical practitioners to expeditiously commence suitable therapeutic strategies. In conclusion, this instrument makes a substantial contribution towards enhancing the efficacy of breast cancer management in the field of clinical practice. However, because to its ineffectiveness, the identification of breast cancer is still an unresolved problem in medical image analysis. By combining deep learning and machine learning techniques, we were able to create a categorization system built on histology images that significantly improved the accuracy of early breast cancer diagnosis while relieving doctors of some of their effort. Multilayer Perceptron and LightGBM classifiers are used in this study's analysis of histology images of breast cancer (BC). Histopathology images are the gold standard when it comes to making a diagnosis of breast cancer. Here, a dataset of 3104 photos is used to train our model, and the image is subsequently successfully identified. This strategy also yields impressive analysis and findings. We test the suggested strategy using histology images from the Breast Cancer dataset, and we achieve a high classification accuracy of 98.28%. The results of the experiments demonstrate that our strategy can perform pretty favourably and exceed cutting-edge approaches.In the near future, using deep learning to predict breast cancer may show to be quite successful.

Keywords: Artificial intelligence; Deep Learning; Multilayer Perceptron; Histopathological Images; LightGBM; Breast Cancer.

Introduction

With over 30% of all female malignancies, breast cancer is the most frequent cancer in women globally and is regarded as a multifactorial disease. This condition has been more prevalent over the last 30 years, yet fewer people are dying from it. It is expected that mammography testing would lead in a 20% decrease in death and a 60% increase in cancer therapy [1]. As the name implies, breast cancer begins in the cells of the breast and primarily affects females. Breast cancer is the second leading cause of death among females after lung cancer. The microscopic appearance of the cancer cells is used to categorize breast cancer into its several subtypes. The two most common types of breast cancer are invasive Ductal carcinoma (IDC) and Ductal carcinoma in situ (DCIS). Ductal carcinoma in situ (DCIS) often develops slowly and has little effect on patients' daily life. The DCIS type accounts for a small portion of cases (between 20% and 53%); in contrast, the IDC form, which encircles the entire breast tissue, is more harmful. This category includes the majority of breast cancer patients, roughly 80% of them [2].

Early breast cancer discovery can lead to successful treatment. To identify the first sign of breast cancer, accurate screening techniques must be available. Several imaging modalities are used for screening purposes, with thermography, ultrasonography, and mammography being the most common. Some of the most useful methods for detecting breast cancer at an initial stage is mammography. Diagnostic sonography methods, such as ultrasound, are commonly employed as an alternative to mammography for breasts with solid tissue. These methods utilize high-frequency sound waves to produce images of the breast, allowing healthcare professionals to assess the presence of any abnormalities or potential tumors. Due to these factors, thermography might be preferable to ultrasonography for the diagnosis of smaller malignant masses [3], while radiography can safely ignore small masses.

Patients with modest and undetectable cancer indications can have diagnostic mammography performed to evaluate abnormal breast tissue. This method can't be utilized to evaluate places where cancer may be suspected due to the sheer volume of images. A study found that screenings of women with highly dense breast tissue missed about 50% of breast tumours. But within two years of screening, around 25% of breast cancer patients receive a negative diagnosis. That's why it's so important to detect breast cancer early [4]. The use of ultrasound to screen for breast cancer is uncommon. While not always necessary, it can be useful when looking for lumps or other changes in the breast. Ultrasound is especially beneficial for women due to its capability of detecting breast abnormalities that may elude detection by mammograms. Furthermore, it assumes a critical function in the investigation of suspicious regions detected by mammograms. In cases where a mammogram detects a potentially malignant region, an ultrasound may be employed to conduct a more comprehensive examination of that particular area. Ultrasound aids physicians in the assessment of the characteristics of the suspected region, enabling them to distinguish between benign and malignant growths and make well-informed judgments concerning additional diagnostic procedures or therapeutic interventions by means of real-time imaging. In many cases, ultrasound can tell the difference among a fluid-filled growth like a cyst and a solid tumor that needs further testing to rule out cancer. It has also been shown that ultrasound can be utilized to direct a biopsy needle into a breast tissue sample, facilitating the removal of cells for further analysis to detect cancer. This method can also be used to alleviate symptoms of swollen lymph nodes in the armpit. Ultrasound is easily accessible and does not expose the patient to any radiation. Additionally, it frequently costs less than alternative testing methods. The ability of thermography to diagnose or screen for early-stage breast

cancer on its own has not been demonstrated. The best primary screening approach for finding breast cancer in its earliest, most curable stages is still mammography.

The proper diagnosis of breast cancer must be made by a pathologist as part of a multidisciplinary approach to therapy. Mammography, magnetic resonance imaging, ultrasound, pathological imaging, and anomaly screening imaging were all employed to find a malignant breast tumor. The features of the breast tissue determine the type of radiographic imaging or histological pictures that are used. To reach the right conclusion, one must obtain a tissue test that is shown by a radiographic deviation from the norm. The pathologist assesses the tissue to determine its value and potential harm. Not all masses detected by imaging are cancerous growths. Some lumps in the breast, such fibro adenomas, are benign tumours. A biopsy entails the removal of the tissue's thin centre using a tiny needle, which is then processed in a pathology lab. The tissue is cut up into tiny pieces and set on glass slides. The glass slides that have been coloured in a dramatic way to stain the tissues pink and blue can be seen up close when seen via a magnifying glass. To determine whether the tissue test for a dangerous malignant growth is present, pathologists use a magnifying lens to inspect the slides [5].

Artificial intelligence (AI), machine learning (ML), and convolutional neural networks (CNN) are three of the medical industry's fastest-growing subsectors [6] [7]. The scientific field that works with and enhances technology systems to handle difficult problems by minimizing the need for human intelligence is where AI and ML are located. Artificial neural networks were a necessity for deep learning (DL), something that belongs to the family of automated learning systems. Deep learning architectures (DL architectures) including CNN, RNN, DBN, and DNN are typically utilized in these many areas. It is feasible to improve the speed and precision of cancer diagnosis by employing modern technologies, in particularly DL techniques [8].

In order to categorize a breast cancer Histopathological Image, we used Multilayer Perceptron (MLP) and LightGBM classifiers in our research. When compared to other significant types of neural network design, like Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Autoencoders (AE), and Generative Adversarial Networks, Multilayer Perceptron (MLP) is the most basic type. Perceptron (neurons) are layered one on top of the other in an MLP. Each layer's nodes are all linked to each other through the layer above. Nodes inside a layer are not connected to one another. Fully connected neural networks make up MLPs. We therefore add layers using Keras' Dense() class. Data flows through the layers of an MLP in a single direction, from input to output. In the fully connected last layer of the MLP structure is where the LGBM Classifier does its work. A well-known machine learning technique forms the foundation of LGBM, a rapid, distributed, and high-performance gradient lifting system. With the largest delta value selected for growth, this algorithm grows leaves at a time. The algorithms used by LightGBM are histogram-based. The latest benchmark methodologies are compared in a comparative analysis. The great performance of the suggested MLP-LGBM model has been shown by a significant comparison research.

Related Works

One of the most frequent forms of cancer among females is breast cancer, typically affecting those between the ages of 60 and 80. The conventional approach uses a mammography scan, followed by a number of other clinical tests, to manually determine the presence of cancer in the body. This method is error-prone and slow to diagnose cancer. It is frequently found using the biopsy technique, which involves removing tissue from the breast and examining it under a microscope. The histopathology department handles this entire procedure, and if he is not properly qualified, it could result in a misdiagnosis. Automatic analysis of histopathology images has aided pathologists in accurate diagnosis by allowing for adequate detection, which has improved diagnosis. Convolutional neural networks (CNN) are now the go-to deep learning technique for categorizing breast cancer. In paper [10], the researcher proposes a convolutional neural network (CNN) structure that takes Local Binary Pattern (LBP) pictures as input and compares its categorization outcomes to those of a traditional CNN that takes photos from the same origin. Here, a classification strategy is put out for the automatic classification of cancer into mild or moderate stages. This approach uses a 100 image dataset, with 80 percent of the photos being used for training and the remaining 20 percent being utilized for testing.

In terms of mortality rates, breast cancer is a major global health problem for women. Breast histopathology photos that have been H&E stained are frequently used for cancer diagnosis. Misdiagnosis risks could be increased by potential differences in slide preparation circumstances and stain colour in H&E pictures. A novel deep learning-based approach for breast H&E stained histopathology pictures was proposed by the author of [11]. The sample extraction strategy and colour normalization method used to pre-process H&E histology images are described in the designed approach's pipeline. It is also explained how to compute additional deep features using deep convolution neural networks (CNN) and how to classify data using different deep classifiers. On a difficult breast histological BreakHis dataset that is publicly available, the performance of

the designed technique is evaluated. The suggested system outperformed the state-of-the-art methods and showed considerable potential in its ability to categorize various cancers.

Women are prone to breast cancer more often than males. To make a diagnosis, pathologists study high-magnification images of breast histopathology slides. Error is more likely to occur since there are not sufficient pathologists to adequately serve the population in many countries. Today, deep learning is one of AI's most active study areas. Although deep neural networks had showed promise in image processing, their inability to analyse medical pictures is due to their insistence on a vast number of training data. In order to enhance the traditional models inmulticlass categorization, the researcher of [12] suggests a deep transfer learning-based model that heavily relies on pre-trained DCNN employing a sizable portion of the ImageNet dataset. We do this after comparing many cutting-edge deep learning techniques and techniques for analysing health records. The method is initially used with a deep classifier paired with data enrichment to identify among differentcancer and non-cancerous samples in binary and multi-class categorization. When importing a pre-trained DesneNet121 for Imagenet, the initial weights are what are copied over. This technique obtained a better accuracy in multi-class categorization. The results achieved for breast cancer CAD systems exceed the precision of past studies in several performance criteria. Also, the DCNN is flexible and extensible, thus it may be used to detect new types of sickness and combined with other CNNs to enhance generalization abilities.

To categorize breast cancer histology images, it is recommended to utilize an attention-guided convolutional neural network (CNN). It is often difficult to comprehend the reasoning behind the network's decisions. An effective CAD system will facilitate an open and transparent decision-making procedure. It is also critical that the network's verdict be grounded in histological features that are in agreement with those of a human expert. In order to better classify breast cancer data using CNN, researcher [13] proposes utilizing extra region-level supervision in the form of localized regions of interest (RoIs) to guide the attention of the categorization network. The proposed technique yields class activation maps that are highly consistent with what a seasoned pathologist would anticipate. In addition, on the BACH microscopy test database, the proposed method shows substantial improvement above the state-of-the-art.

Breast cancer, the most common form of cancer in women, poses a significant risk to female health. Promising research is being conducted on the application of deep learning algorithms for its early diagnosis. However, many popular Convolutional Neural Network (CNN) variants which are vulnerable to overfitting in breast cancer categorization because of the limited size of available breast pathology imaging datasets and the overconfidence inherent in the softmax-cross-entropy loss they employ. Overfitting is a common issue in categorization problems, and the AlexNet-BC model is a new framework proposed by the researcher [14] to address this issue and increase classification accuracy in breast pathology. Before fine-tuning, the model is pre-trained using the ImageNet dataset. We also built a better cross-entropy loss function to discourage overconfident low-entropy result probabilities and make the estimates appropriate for uniform distributions. The AlexNet method is subsequently verified by using various comparison tests conducted on various datasets. The practical findings show that the suggested technique outclassed the traditional techniques at a variety of magnifications. It is suitable for computer-aided clinical diagnosis systems in histopathology due to its great robustness and generalization capabilities.

Breast cancer is a devastating illness affecting women, but it may often be successfully treated if caught at an early stage. The accurate identification of BC is a major concern for clinicians and researchers. Many scientists have suggested using deep learning techniques to better diagnose breast cancer. However, these methods do not offer an accurate diagnosis of breast cancer. The researcher [15] proposes a three-layer convolutional neural network (CNN) architecture as a means of improving the accuracy of breast cancer diagnosis. The proposed model has been trained and validated using a dataset of breast histology images. The ideal model was selected, and the hyperparameters were fine-tuned, using the Hold out cross validation method. Additionally, various model evaluation measures have been employed to assess the performance of models. The findings of the experiment showed that the propped approach is better appropriate for treating breast cancer and that it can be successfully adopted into medical practice.

Problem Statement

Breast cancer is an example of a malignancy that can have its origins in the breast. The left or right breast might be the first site of development. Whenever cell growth becomes unchecked, cancer sets in. Although men can also develop breast cancer, breast cancer affects almost exclusively women. Based to data from [9], Table 1 details the top 10 malignancies diagnosed in the United States each year. Figure 1 shows the graphical representation of top 10 malignancies diagnosed in America.

TABLE 1. Rates of New Cancer Cases United States

Cancer Type	Age-Adjusted Rate	Case Count
Leukaemia's	10.3	20952
Pancreas	11.6	26033
Kidney and Renal Pelvis	11.8	24673
Non-Hodgkin Lymphoma	15	31875
Melanomas of the Skin	18.2	36058
Thyroid	19.1	33009
Corpus and Uterus NOS	27.7	59450
Colon and Rectum	31.8	66881
Lung and Bronchus	48.1	109094
Female Breast	129.7	264121

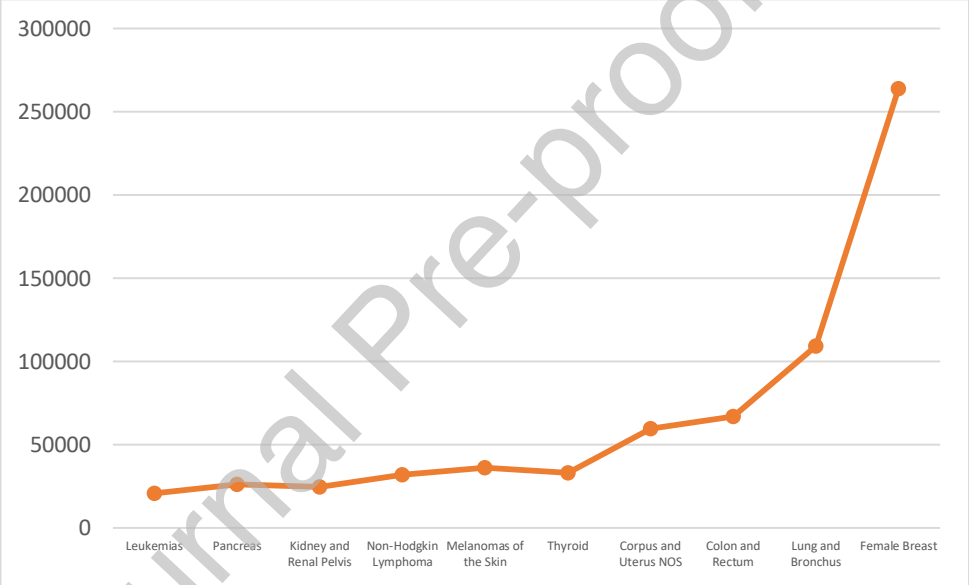


Figure 1. Number of new cancer cases in the USA

It is critical to remember that the vast majority of breast lumps are harmless and not malignant. Breast tumours that are not malignant are abnormal growths that do not spread to other parts of the body. Certain benign breast lumps can increase a woman's risk of acquiring breast cancer, despite the fact that they are not life-threatening. Cancer cells from breast cancer can travel to other parts of the body by penetrating either the circulatory or lymphatic systems. Once they enter these systems, they can be carried to distant organs and tissues, leading to the formation of secondary tumors. The immune system, which includes the lymphatic system, plays a crucial role in the spread of breast cancer. While the lymphatic system helps transport cancer cells to other parts of the body, the immune system's response is critical in recognizing and eliminating these abnormal cells. However, cancer cells can sometimes evade the immune system's surveillance, allowing them to establish secondary tumors. Clear lymph fluid is gathered and transported through the bodily tissues to the circulation via a network of organs called lymph nodes, which are tiny glands about the size of beans. There are immune system cells, debris from tissue, and waste items in the transparent lymph fluid that flows through the lymphatic system. It is more likely that cancer cells have metastasized to different parts of your body via the lymphatic system when they had migrated to your lymph nodes. However, not all lymph node-positive women go on to develop metastases, and some lymph node-negative females get metastases in the first place. Changes in risk factor exposure, adoption of screening tests, and treatment advancements all have an impact on the incidence of cancer diagnoses and cancer deaths. According to the lines below and the maps, where the colour is progressively becoming lighter over time, some cancer rates are decreasing. In some instances, the rate may be

decreasing, but there may be an increase in the number of new cases and fatalities. This occurs as a result of the aging and expanding population in our country.

Methodology

This work develops and evaluates a deep learning structure that combines machine learning and image categorization ideas for automated breast cancer detection. Different Deep Neural Network topologies, particularly those designed to process picture data, such Convolutional Neural Networks, have been discussed. In a manner analogous to digital staining, in which a classification system is used to highlight key parts of a picture informing diagnostic judgments, this function was implemented. The system labels the raw pixel input picture as either benign or malignant and then utilizes the highlighted visual features to distinguish between healthy and malignant tissue. By analyzing these visual features, the system can accurately identify and classify the tissue as either healthy or malignant. In order to boost breast cancer prediction accuracy, the system proposed a hybrid categorization model consisting of Multi-Layer Perceptron and LightGBM Classifier. The suggested method makes advantage of these deep learning approaches for improved results. The process of the recommended system is shown in Figure 2. The primary purpose of this research is to develop a model for improved breast cancer prediction.

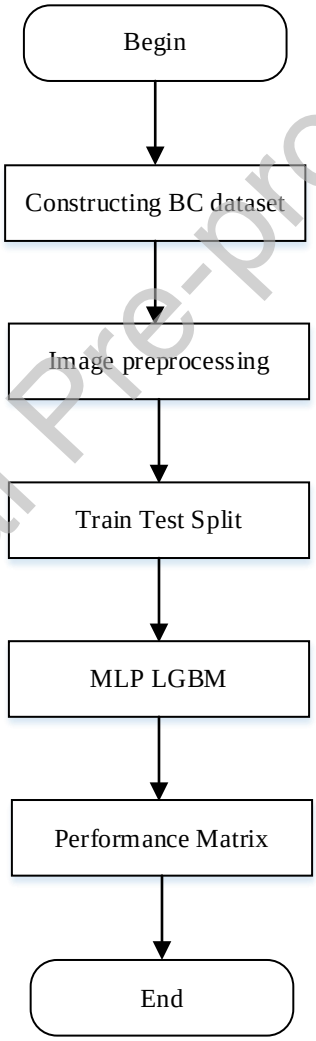


Figure 2.Workflow of Proposed Model

Data Collection

In 2021, there are estimated to be 2.7 million new cases of breast cancer, according to the American Cancer Society, and 721,000 people will die from the disease. About 77% of all breast cancer diagnoses can be attributed to invasive ductal carcinoma (IDC), the most frequent subtype of the disease. Making an accurate

diagnosis as soon as feasible is crucial for choosing the best course of treatment and boosting the overall survival rate of patients. Microscopic evaluation of histopathological stained tissue and subsequent digitalization is now a more viable option than it was in the past due to recent advancements in slide scanning technology and a decline in the cost of digital storage over the past few years. We have two sets of folders, train and test, in this Dataset [16] is collected from Kaggle. There are 1497 normal photographs and 542 illness images in the train folder. There are 928 normal images and 137 illness images in the test folder. A graphical depiction of the dataset sample photos for Normal and Diseases is shown in Figure 3. Image distribution the breast cancer dataset is shown in Table 2 and Figure 4.

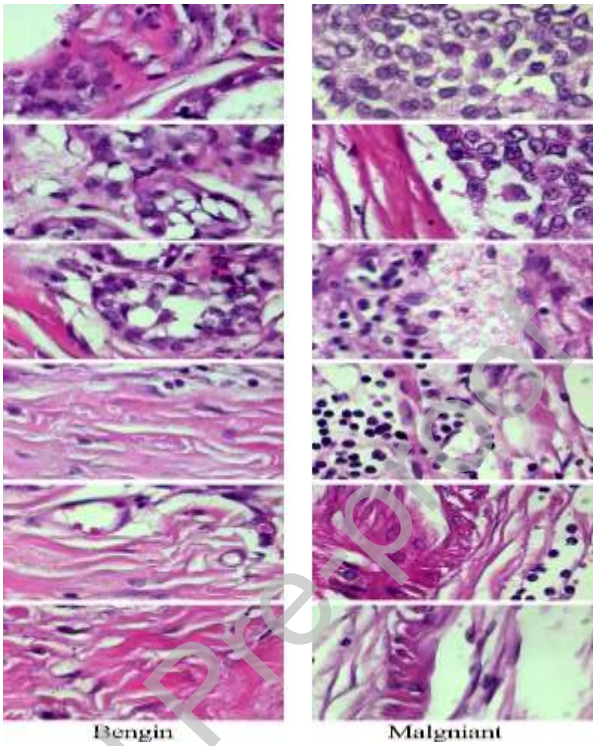


Figure 3. Sample Image from BC Dataset

TABLE 2. Histopathological Breast Cancer Datasets

SL No	Category	No of images
1	Cancer	2425
2	Non cancer	679

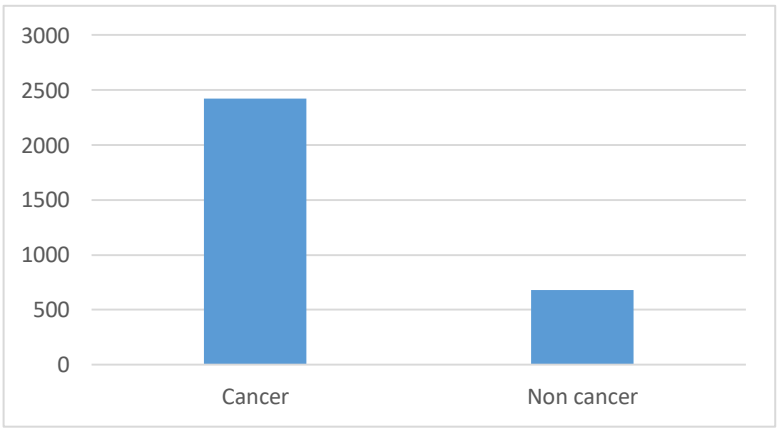


Figure 4.Data distribution in BC Dataset

Multi-Layer Perceptron

Deep learning includes a multilayer artificial neuron network as a key component. Artificial neural networks (ANN) are made up in part of the multi-layer perceptron (MLP) algorithm. An ANN's fundamental building block is a perceptron, which is a network with just one node. Inputs, weights, an activation function, and output make up a perceptron. Each input and its weight are given to the activation function, which multiplies them and outputs the outcome. However, the MLP requires minimum three nodes in order to function as an ANN. The nodes of an MLP are organized into three levels: the input layer, the hidden layers, and the output layer. The difficulty in designing an MLP lies in choosing the appropriate number of neurons for each layer and determining the optimal number of hidden layers. As a result, each time we modify a parameter, the MLP architecture is evaluated using evaluation metrics.

Multi-layer perception (or MLP) is a synonym. It consists of interconnected, thick layers capable of mapping any input measurement to the desired output. Multi-layer perception refers to a neural network that has several levels of processing. Some of the results of the neurons that make up a neural network are also used as inputs by other neurons in the network. Every output from a multi-layer perceptron is sent to a separate node in the outcome layer, and every input is processed by a single neuron (or node) in the input layer. The amount of hidden layers and their associated nodes is up to the designer. A Multi-Layer Perceptron (MLP) is shown schematically in Figure 5.

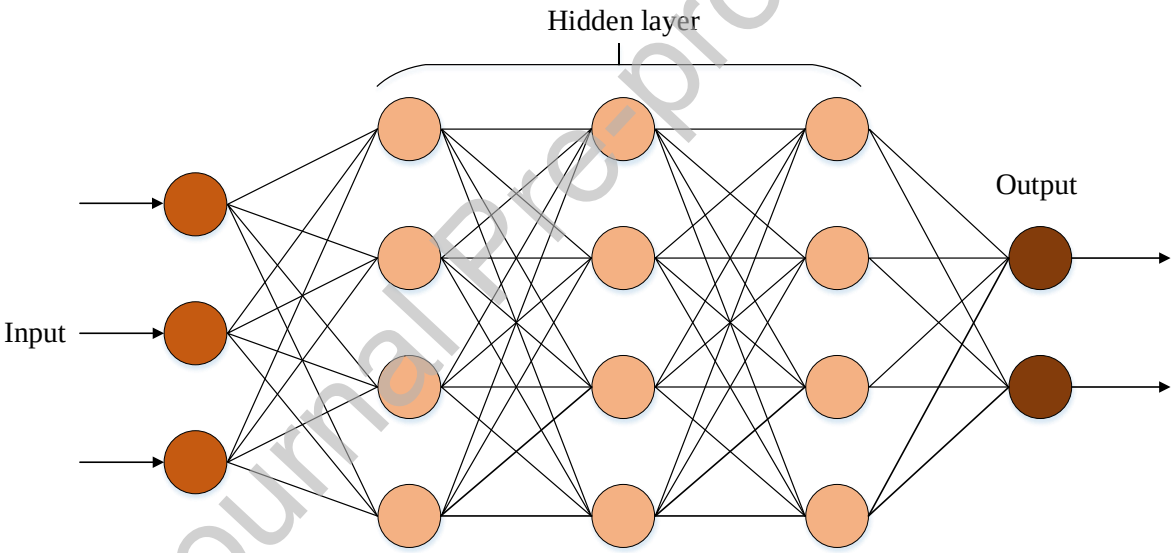


Figure 5. Architecture of MLP Model

The aforementioned multi-layer perceptron diagram has three inputs, three corresponding nodes, and three hidden layer nodes. Given that the output layer generates a pair of outputs, we can deduce that there are two nodes representing those findings. The input layer nodes in the accompanying layout transmit their information to every of the three hidden layer nodes, which then does some sort of processing on it before sending it on to the output layer. The information is passed on to the next layer of processing by the nodes in the input layer. Every multi-layer perceptual node uses a sigmoid activation function. The sigmoid activation function converts real information to a value among 0 and 1 using the sigmoid formula.

- The Sequential model restricts us to single-input, single-output stacks of layers and allows us to build models layer-by-layer as needed in a multi-layer perceptron.

- Without altering the batch size, flattening the input flattens it. Without a feature axis, input shapes that are (batch size,) include an additional channel dimension that is flattened to produce output shapes that are (batch size, 1).

- When utilizing the sigmoid activation function, activation is required.
- The first two dense layers—also known as the hidden layers—are what are used to create a fully connected model.
- The output layer, which determines which category the image belongs in, is the final dense layer.

Activation functions are of utmost importance in the learning phase of ANNs as they enable the networks to process inputs and produce meaningful outputs. By using sophisticated non-linear functions, activation functions allow ANNs to capture intricate patterns and relationships in the data, enhancing the network's ability to learn and make accurate predictions. Without these functions, the learning phase of ANNs would be severely limited, hindering their overall effectiveness. The activation functions most commonly used with MLP hidden layers include the identification function, the logistic sigmoid function, the hyperbolic tan function (*tanh*), and the corrected linear unit function. The hidden layer of NN uses this activation function, which is the most prevalent. (0,) is a deceptively straightforward formula. Contrary to its name and outward appearance, it is not linear and provides the same benefits as Sigmoid while also functioning better. The main advantage is that it effectively addresses the vanishing gradient issue while requiring less computational resources compared to *tanh* and sigmoid functions. Unfortunately, during the training process, a few gradients can become fatally unstable, necessitating their removal. This drawback highlights a significant limitation of the method. As a result, the neurons are killed out. Therefore, for activations in the region (x<0) of ReLu, the gradient will be 0, so the weights won't change throughout descent. The neurons in that state would no longer respond to random oscillations.

Neural networks require fine-tuning of their hyperparameters to optimize their performance and achieve better accuracy. By adjusting the number of layers and neurons, we can control the complexity and capacity of the network, allowing it to learn and generalize patterns effectively. Although backpropagation is a technique that is frequently utilized for the purpose of training neural networks to correct weights, it is not the only strategy that is accessible. The process of weight adjustment can also be accomplished by the utilization of other techniques, such as genetic algorithms and reinforcement learning, in certain circumstances. However, disappearing gradients may become a concern at deeper layers. In order to resolve this problem, special algorithms are needed. Calculate the hidden layer's activation unit$x_1^{(h)}$ in the first step.

$$a_1^{(h)} = x_0^{(in)} w_{0,1}^{(h)} + x_1^{(in)} w_{1.1}^{(h)} + \cdots + x_m^{(in)} w_{m,1}^{(h)} (1)$$

$$x_1^{(h)} = \emptyset\left(a_1^{(h)}\right) \qquad (2)$$

In the presence of an activation function the *a* value, the outcome is an activation unit∅. Gradient descent weight learning requires that it be differentiable. In many cases, the sigmoid (logistic) function serves as the activation function∅. It permits the nonlinearity required to solve difficult issues.

$$\emptyset(a) = \frac{1}{1 + e^{-z}} \qquad (3)$$

- Feed-forward pass :

$$a_0^{(l-1)}(r) \equiv 1 \qquad (4)$$

$$For\ r = 1\ to\ R, l = 1\ to\ L, i = 1\ to\ P(l),$$

$$a_i^{(l)}(r) = f\left(u_i^{(l)}(r)\right) = \frac{1}{1 + \exp\left[-u_i^{(l)}(k)\right]} (5)$$

$$u_i^{(l)}(r) = \sum_{j=0}^{N} v_{ij}^{(l)}(t) a_j^{(l-1)}(r) (6)$$

$$t: epoch\ index$$

$$r: sample\ index$$

- Weight update pass :

$$For\ r = 1\ to\ R, l = 1\ to\ L, i = 1\ to\ P(l),$$

$$-\frac{\partial E}{\partial v_{ij}^{(l)}(t)} = \sum_{k=1}^{K} \delta_i^{(l)}(r) a_j^{(l-1)}(r) \quad (7)$$

$$v_{ij}^{(l)}(t+1) = v_{ij}^{(l)}(t) - \mu \frac{\partial E}{\partial v_{ij}^{(l)}(t)} + \mu\left(v_{ij}^{(l)}(t) - v_{ij}^{(l)}(t-1)\right) \quad (8)$$

$A(l), l = 0,1, \dots, L: P(l) \times R$ matrix. i^{th}column represents neuronal responses to ther^{th}training sample in thel^{th} layer $if\ l > 0$, andthat which is used to train the feature vector if $l = 0$.

$V(l), l = 1,2, \dots, L : P(l) \times (P(l-1) + 1)$weight matrix. The$(i,j)^{th}$elements of $V(l)$is$v_{ij}^{(l)}$.

LightGBM Classifier

Light GBM, short for Light Gradient Boosting Machine, is a framework for gradient boosting that is based on decision tree algorithms. Additionally, it reduces memory usage by utilizing a histogram-based algorithm, which further enhances its efficiency in handling large-scale data. Light GBM is distinct from previous gradient boosting architectures primarily because it expands vertically and leaf-wise. Some techniques, on the other hand, expand outward in a horizontal, level pattern. Light GBM picks the leaf with the lowest error and most effectiveness. Error rates can be reduced by a much larger margin with this method. In other words, its leaves spread outward rather than up, unlike other plants. As a result of standard algorithms' subpar accuracy and increasing difficulty in producing quick results, Light GBM has grown in popularity in recent years. We require a new model that will be quicker and more effective than the Light GBM because data volume is growing constantly. Because of its high-speed training, we refer to it as Light. Despite handling a lot of data, Light GBM only uses a little amount of memory. Diagrammatic representation of leaf-wise tree growth is shown in Figure 6.

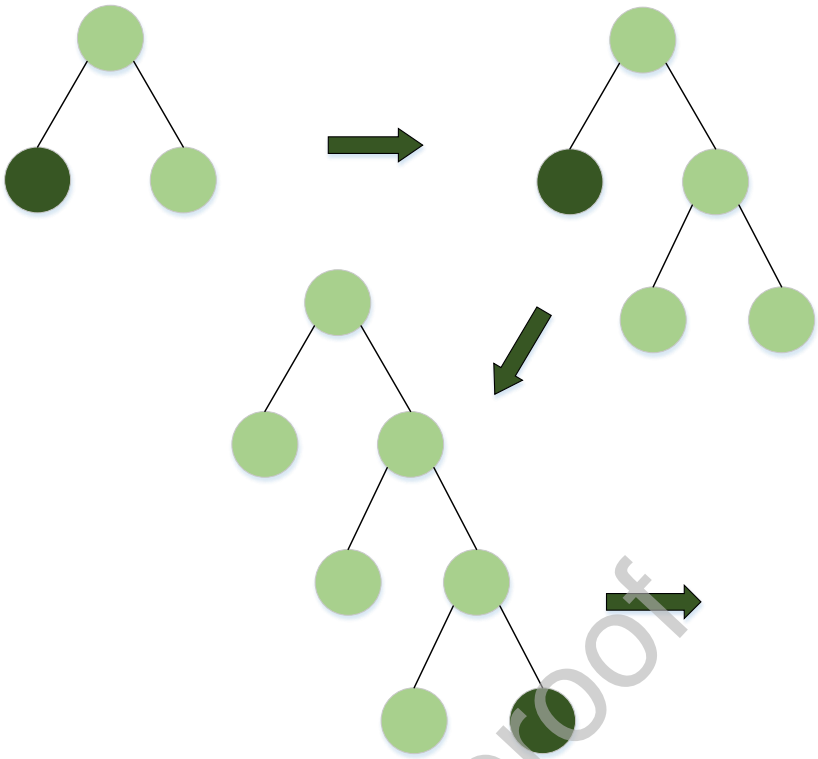


Figure 6 : Diagrammatic representation of Leaf-Wise Tree Growth

Because of Light GBM's overfitting issue, we are unable to employ it on tiny datasets. Overfitting is simple for LightGBM because it only needs a small sample size. Parameter tweaking is the most challenging and crucial aspect of putting the Light GBM into practice. When implemented, it involves about a hundred parameters. These are the benefits of doing so:

- Reduced Memory Use
- Lower communication costs for concurrent learning
- A decrease in the cost of determining the gain for each decision tree split.

MLP-LGBM Model

Using the Multi-Layer Perceptron and LightGBM algorithms, we build a model in this paper. Figure 3 displays the ensemble process. The MLP extracts and filters the features of the input data. Afterward, retrieve data and classify it once again using the LightGBM model using the flatten layer output. Enhance the model's predicting performance. Locating the ideal segmentation point prior to building a decision tree is crucial because it helps in effectively splitting the data and creating accurate decision boundaries. This ensures the decision tree's development aligns with the desired outcome. The general approach of enumerating all potential feature points involves sorting feature values, which consumes a significant amount of memory and time. This process becomes time-consuming and resource-intensive, making it a challenge for decision tree development. Utilizing a better histogram method is the LightGBM algorithm. The division points are chosen from a set of k values and the continuous eigenvalues are split into k intervals. Therefore, compared to the GBDT algorithm, it is faster at training and more space-efficient. While being a weak classifier, the decision tree. The histogram algorithm's use will result in a regularization effect and successfully prevent overfitting. The LightGBM method uses a leaf-wise generation approach to reduce the amount of training data. When growing the same leaf, the leaf-wise method can cut losses more than the traditional method, such as level (depth)-wise. Also used to limit decision tree depth and prevent overfitting is the additional parameter.Suggested MLP-LGBM Model Architecture is shown in Figure 7.

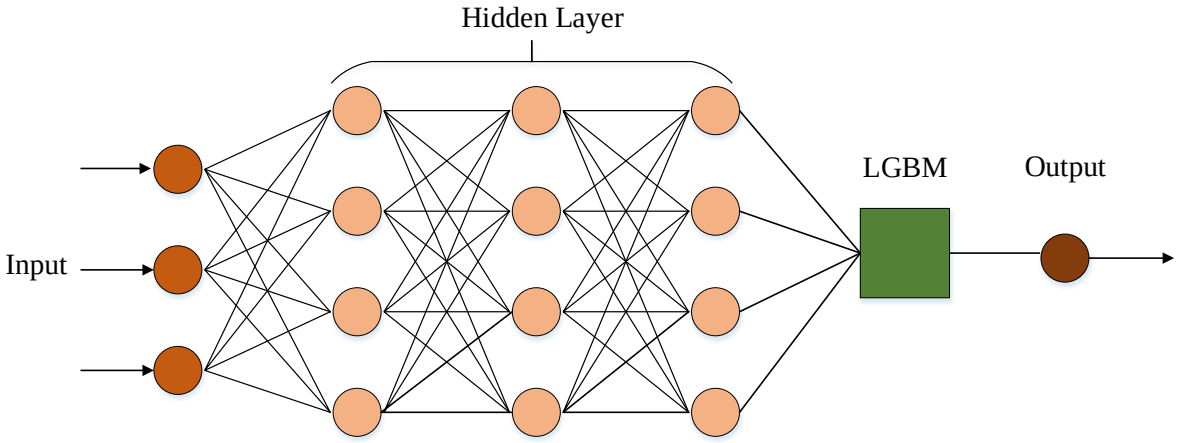


Figure 7. Suggested MLP-LGBM Model Architecture.

The model comprises of a LightGBM classifier and a straightforward MLP architecture. A 32x32 matrix of normalized images from the Breast Cancer dataset is used as the MLP input layer's starting point. The input image's distinctive features are extracted by the feature map layer as values. After running through multiple epochs and until the training process converges, the MLP is trained. Here, the LightGBM classifier takes the place of the final layer of the MLP. The dense_3 layer's characteristics from the input image are used as an input for the LightGBM classifier. These new automatically created training picture characteristics are used to train the LightGBM classifier at first. Finally, the digits utilized for testing are recognized by the trained LightGBM classifier.

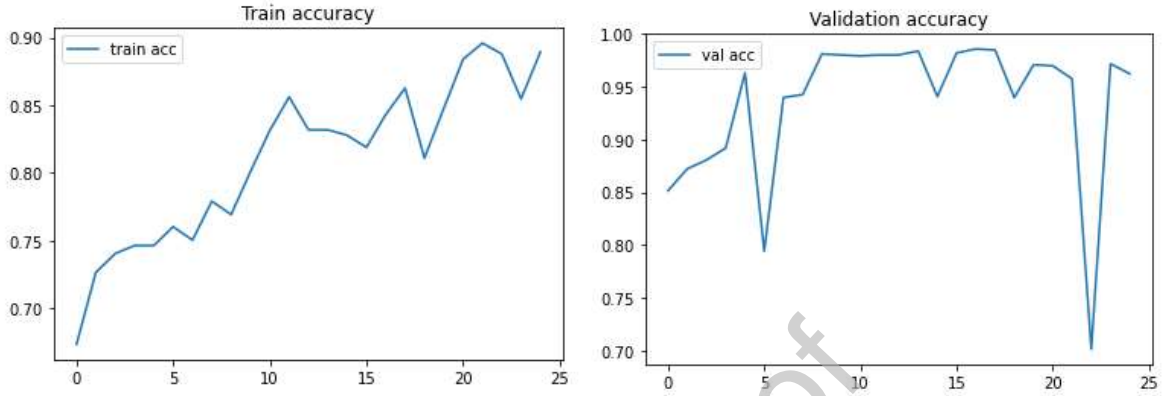
Results and Discussions

Breast Cancer Histology For this effort, breast cancer images gathered from Kaggle are classified. Jupyter Notebook, a highly effective program for dynamically creating and reporting information science projects, is used for all of the experiments. In a Jupyter Notebook, the script and its results may be seen side by side alongside visualizations, narrative text, and scientific calculations in a single document.

The loaded breast cancer image goes through early pre-processing. To ensure that learning and inference happen on the same picture attributes, preprocessing should be applied to our training, validation, and testing set. Images require pre-processing prior to being utilized in model training and inference. Alterations to the scale, direction, and hue are all in scope here. Image scaling is an essential pre-processing step in computer vision. Our machine learning techniques learn faster and better when presented with smaller images. More time is needed for our system to learn from an input picture that is four times larger and contains twice as numerous pixels. Furthermore, most deep learning network structures require that our photos be the same size, even though the raw acquired pictures may vary in size. The main benefit of feature extraction is that it allows for resource reduction without sacrificing important data. This means that machine learning models can become more effective and accurate by extracting only the relevant features from the input data and Feature extraction helps in dealing with the laborious task of working with a large amount of data by reducing the amount of time and storage required. Since a significant portion of the input data is often redundant, feature extraction helps in eliminating this redundancy and extracting only the essential features, making the process more efficient and manageable. Feature extraction is a method for distilling important features from a huge input data collection. Dimensionality reduction is used in this procedure to break up enormous input data sets into more manageable processing units. The train-test-split methodology allows you to replicate your model's performance on new information as part of the validation process. Separate the dataset in halves to use as a training set and a test set. To do this, you will need to add an 80% non-replaceable random selection of the information to your training set. The last 20% of the test group is completed by you. For a specific train test split, the graphics in "Features" and "Labels" show where their data will go (X_train," "X_test," "y_train," "y_test").

Multilayer Perceptron will be used in this study to train the model. The photos from the histology of breast cancer provide the input for the various deep learning models. Each neural network's anticipated input was scaled appropriately for these photos. It was done using an Adam optimizer with a 105 learning rate. For each model, a total of 25 iterations were performed using the categorical_crossentropy as the loss function during the training process. The Multilayer Perceptron model should be made more accurate and effective.

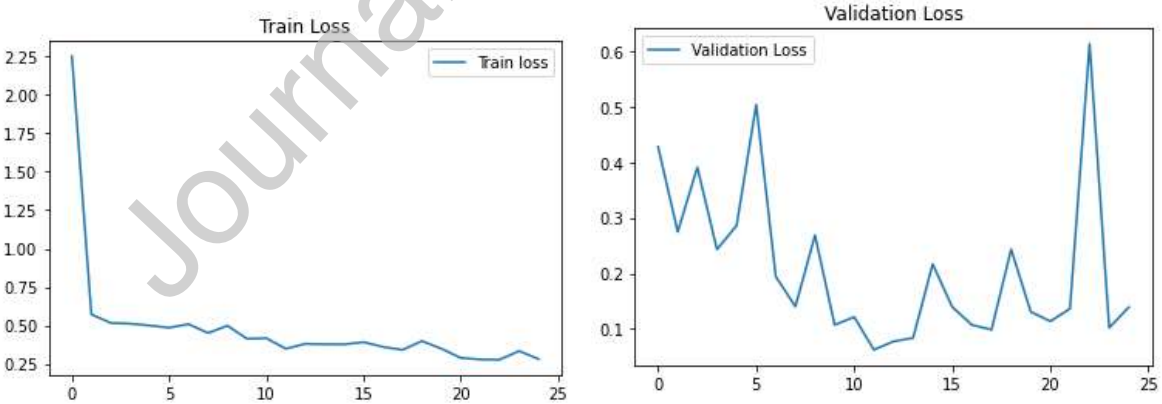
Using the LightGBM Classifier, we extract the characteristics of the final layer of the Multilayer Perceptron model. Faster training times and greater efficacy characterize the LightGBM classifier. Less memory is used, and the results are more accurate. In this case, a multilayer perceptron model will be used for feature extraction, and a classification model called the LightGBM Classifier will be used. The accuracy of the suggested model, which is 98.28%, is significantly higher than that of previous deep learning and machine learning models.



(a) **(b)**

Figure 8. Model Accuracy and Validation Accuracy of suggested model

Both the training and validation accuracies are shown in Figure 8. Figure 9 displays the model's training loss as well as its validation loss. It's possible that the keras model's accuracy and loss on validation information will change based on the specifics of the training process. Loss should typically decrease and accuracy should increase as each period increases. Our model is cramming values rather than learning as validation loss starts to rise and validation accuracy to fall. In situations when softmax is being utilized in the output layer, the rise in validation accuracy and loss could be caused by overfitting or a variety of probability values. It's okay if validation loss starts to decline and validation accuracy starts to rise because it shows that the model is adapting and operating properly.



(a) **(b)**

Figure 9. Model Loss and Validation Loss of suggested model

To evaluate the performance of a categorization model, scientists create a confusion matrix of size Sx S, where S is the amount of groups being evaluated. The confusion matrix plays a crucial role in determining how well a categorization model performs. By analyzing the number of correct and wrong predictions made by the classifier, it provides valuable insights into the model's accuracy, precision, recall, and overall performance. This technique helps in understanding the strengths and weaknesses of the machine learning model. The accuracy, precision, recall, and F1-score are only few of the efficiency measures that may be calculated to evaluate a classification model's quality. Confusion matrices are commonly employed in place of classification

accuracy since they offer a more comprehensive overview of a model's effectiveness. Confusion matrix of our proposed model is shown in Figure 10.

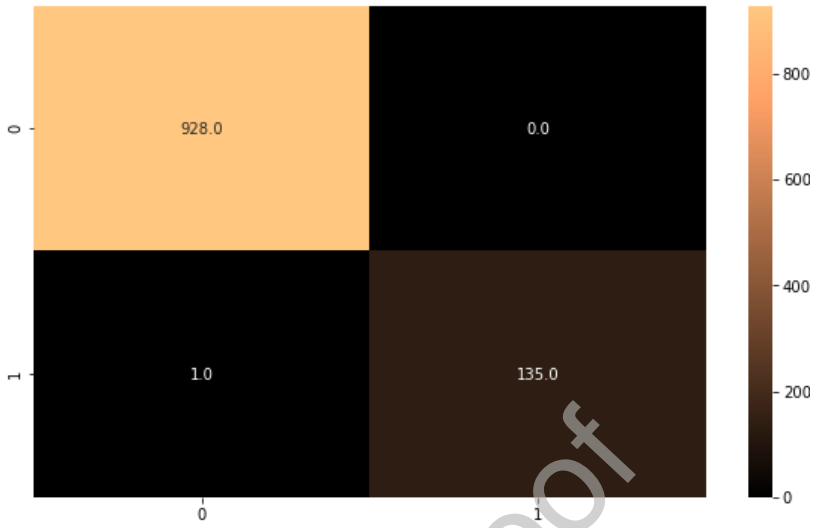


Figure 10. Confusion Matrix of Suggested model

An enlarged form of the confusion matrix is what a classification report is. Other metrics exist that can aid in a deeper comprehension and analysis of our model's operation. These metrics are not limited to the confusion matrix. Precision can be defined as the ratio of correctly classified positive classes to the entire amount of predicted positive classes. The recall measure is the proportion of positive classes that were properly identified, expressed as a percentage of the total amount of positive classes. The proportion of correctly predicted positive classes can also be calculated. Precision and recall for your classifier are sort of kept in balance by the F1 score. Our F1 score will also be low if our precision and recall are both weak. Classification report of the MLP-LGBM model is shown in Figure 11.

	precision	recall	f1-score	support
0	0.98	0.98	1.00	928
1	0.99	0.99	0.98	136
accuracy			1.00	1064
macro avg	0.98	0.99	0.98	1064
weighted avg	0.98	0.98	0.99	1064

Figure 11. Classification Report of Suggested Model

TABLE 3.PerformanceEvaluation of the MLP-LGBM Model

Model	Accuracy	Precision	Recall	F1
DCNN [17]	85.45	-	-	-
KNN [18]	95.90	0.98	0.90	0.94
LR [19]	93	-	-	-
SLSQP [20]	97.88	0.95	0.90	0.97

SVM [21]	0.96	0.98	0.94	0.92
Proposed Model	98. 28	0.97	0.98	.99

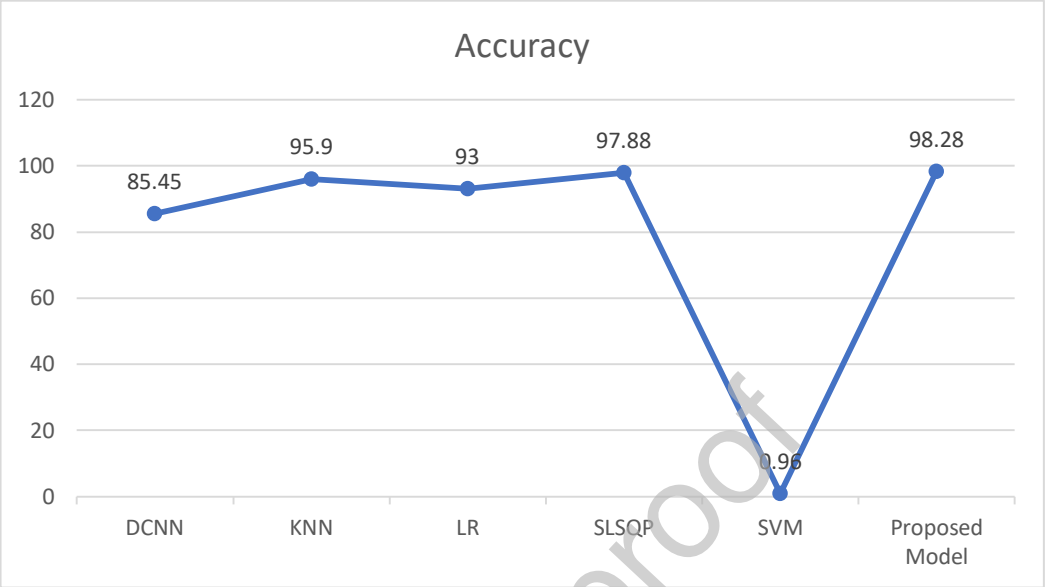


Figure 12. Accuracy Comparison between MLP-LGBM and Existing Model

Existing methods including Deep Convolutional Neural Network, KNN, Logistic Regression, Sequential Least Squares Programming Method (SLSQP), and Support Vector Machine, are contrasted with our proposed model, MLP- LGBM. The accompanying table makes it very evident that our proposed model has outperformed every other model in use. A 98.28% accuracy rate was attained by our model. Accuracy comparison between MLP-LGBM and existing models are shown in Table 3 and Figure 12.

Conclusion

When it comes to malignancies affecting women, breast cancer is among the most deadly and aggressive. It has a high mortality rate, which is unfortunate. Automated examination of histological images can aid in disease diagnosis. Therefore, sophisticated picture analysis supported by computer assisted diagnostics can aid in the diagnosis of this condition. Here, we provide an automated method for histopathology image processing that is based on a Multilayer Perceptron and a LightGBM classifier. In order to attain higher recognition rates, it is necessary to design an ensemble of networks rather than simply one to accomplish this. All of the separate networks that make up the ensemble are used to compute a consensus decision in order to do this. In our study, the LightGBM classifier served as the classification model, and the Multilayer Perceptron was used to extract features. LightGBM takes the place of the Multilayer Perceptron's last layer. The outcomes of the suggested strategy have demonstrated that it is feasible to use the information from other datasets to enhance the accuracy of a specific classification task. In terms of accuracy, our proposed model is contrasted with five other models: Deep Convolutional Neural Network, KNN, Logistic Regression, Sequential Least Squares Programming Method (SLSQP), and Support Vector Machine. However, our model outperformed the current models with an accuracy of 98.28%. A more robust architecture, better hyper-parameter management, and the usage of larger datasets can all be used to enhance this model. These findings are quite good at identifying and categorizing tumours in histopathological pictures, which helps doctors diagnose breast cancer.

Author Contribution

Sajiv G: Investigation, Methodology, Writing - review & editing. G. Ramkumar: Conceptualization, Formal analysis, Writing - review & editing. S. Shanthi: Conceptualization, Formal analysis, Writing - original draft

Writing - review & editing. Arunachalam Chinnathambi: Conceptualization, Sulaiman Ali Alharbi: Formal analysis, Writing - review & editing.

Data Availability

The data used to support the findings of this study are included within the article.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Ethical approval

Not required

Funding

This project was supported by Researchers Supporting Project number (RSP2024R383), King Saud University, Riyadh, Saudi Arabia.

Reference

[1] P. K. Diderot, N. Vasudevan and V. R. Prakash, "Analysis of Mammography Techniques for Breast Cancer Detection," 2019 International Conference on Wireless Communications Signal Processing and Networking (WiSPNET), Chennai, India, 2019, pp. 127-130, doi: 10.1109/WiSPNET45539.2019.9032722.

[2] M. P. Beham, N. Kayalvizhi, R. Tamilselvi and A. Nagaraj, "MAMMSIT: A Database For The diagnosis and detection of Breast Cancer in Mammography images," 2020 Sixth International Conference on Bio Signals, Images, and Instrumentation (ICBSII), Chennai, India, 2020, pp. 1-5, doi: 10.1109/ICBSII49132.2020.9167662.

[3] N. S. Ismail and C. Sovuthy, "Breast Cancer Detection Based on Deep Learning Technique," 2019 International UNIMAS STEM 12th Engineering Conference (EnCon), Kuching, Malaysia, 2019, pp. 89-92, doi: 10.1109/EnCon.2019.8861256.

[4] N. Wu et al., "Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening," in IEEE Transactions on Medical Imaging, vol. 39, no. 4, pp. 1184-1194, April 2020, doi: 10.1109/TMI.2019.2945514.

[5] M. Ahmed and M. R. Islam, "A Multiple-Input Based Convolutional Neural Network in Breast Cancer Classification from Histopathological Images," 2021 24th International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 2021, pp. 1-5, doi: 10.1109/ICCIT54785.2021.9689856.

[6] M. Ahmed and M. R. Islam, "Breast Cancer Classification from Histopathological Images using Convolutional Neural Network," 2021 International Conference on Computer, Communication, Chemical, Materials and Electronic Engineering (IC4ME2), Rajshahi, Bangladesh, 2021, pp. 1-4, doi: 10.1109/IC4ME253898.2021.9768615.

[7] A. K. Arslan, Ş. Yaşar and C. Çolak, "Breast cancer classification using a constructed convolutional neural network on the basis of the histopathological images by an interactive web-based interface," 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 2019, pp. 1-5, doi: 10.1109/ISMSIT.2019.8932942.

[8] W. Zou, H. Lu, K. Yan and M. Ye, "Breast Cancer Histopathological Image Classification using Deep Learning," 2019 10th International Conference on Information Technology in Medicine and Education (ITME), Qingdao, China, 2019, pp. 53-57, doi: 10.1109/ITME.2019.00023.

[9] https://gis.cdc.gov/Cancer/USCS/#/AtAGlance/. (Accessed on December 20th 2022)

[10] A. Madduri, S. S. Adusumalli, H. Sri Katragadda, M. K. Reddy Dontireddy and P. Sarah Suhasini, "Classification of Breast Cancer Histopathological Images using Convolutional Neural Networks," 2021 8th International Conference on Signal Processing and Integrated Networks (SPIN), Noida, India, 2021, pp. 755-759, doi: 10.1109/SPIN52536.2021.9566015.

[11] T. Kausar, M. A. Ashraf, A. Kausar and I. Riaz, "Convolution Neural Network based Approach for Breast Cancer Type Classification," 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST), Islamabad, Pakistan, 2021, pp. 407-413, doi: 10.1109/IBCAST51254.2021.9393249.

[12] Y. Yari and H. Nguyen, "A State-of-the-art Deep Transfer Learning-Based Model for Accurate Breast Cancer Recognition in Histology Images," 2020 IEEE 20th International Conference on Bioinformatics and Bioengineering (BIBE), Cincinnati, OH, USA, 2020, pp. 900-905, doi: 10.1109/BIBE50027.2020.00153.

[13] H. Yang, J. -Y. Kim, H. Kim and S. P. Adhikari, "Guided Soft Attention Network for Classification of Breast Cancer Histopathology Images," in IEEE Transactions on Medical Imaging, vol. 39, no. 5, pp. 1306-1315, May 2020, doi: 10.1109/TMI.2019.2948026.

[14] M. Liu et al., "A Deep Learning Method for Breast Cancer Classification in the Pathology Images," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 10, pp. 5025-5032, Oct. 2022, doi: 10.1109/JBHI.2022.3187765.

[15] A. U. Haq et al., "3DCNN: Three-Layers Deep Convolutional Neural Network Architecture for Breast Cancer Detection using Clinical Image Data," 2020 17th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP), Chengdu, China, 2020, pp. 83-88, doi: 10.1109/ICCWAMTIP51612.2020.9317312.

[16] https://www.kaggle.com/datasets/akhilbs/breastcancer (Accessed on 5th December 2022)

[17] X. Zhou et al., "A New Deep Convolutional Neural Network Model for Automated Breast Cancer Detection," 2020 7th International Conference on Behavioural and Social Computing (BESC), Bournemouth, United Kingdom, 2020, pp. 1-4, doi: 10.1109/BESC51023.2020.9348322.

[18] S. Sharma, A. Aggarwal and T. Choudhury, "Breast Cancer Detection Using Machine Learning Algorithms," 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), Belgaum, India, 2018, pp. 114-118, doi: 10.1109/CTEMS.2018.8769187.

[19] M. R. Al-Hadidi, A. Alarabeyyat and M. Alhanahnah, "Breast Cancer Detection Using K-Nearest Neighbor Machine Learning Algorithm," 2016 9th International Conference on Developments in eSystems Engineering (DeSE), Liverpool, UK, 2016, pp. 35-39, doi: 10.1109/DeSE.2016.8.

[20] M. Gupta and B. Gupta, "An Ensemble Model for Breast Cancer Prediction Using Sequential Least Squares Programming Method (SLSQP)," 2018 Eleventh International Conference on Contemporary Computing (IC3), Noida, India, 2018, pp. 1-3, doi: 10.1109/IC3.2018.8530572.

[21] V. A. Telsang and K. Hegde, "Breast Cancer Prediction Analysis using Machine Learning Algorithms," 2020 International Conference on Communication, Computing and Industry 4.0 (C2I4), Bangalore, India, 2020, pp. 1-5, doi: 10.1109/C2I451079.2020.9368911.

Declarations

The following additional information is required for submission. Please note that this form runs over two pages and failure to respond to these questions/statements will mean your submission will be returned to you. **If you have nothing to declare in any of these categories then this should be stated.**

Conflict of interest

All authors must disclose any financial and personal relationships with other people or organisations that could inappropriately influence (bias) their work. Examples of potential conflicts of interest include employment, consultancies, stock ownership, honoraria, paid expert testimony, patent applications/registrations, and grants or other funding.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Please state any sources of funding for your research

This project was supported by Researchers Supporting Project number (RSP2024R383), King Saud University, Riyadh, Saudi Arabia.

Ethical Approval

Work on human beings that is submitted to *Medical Engineering & Physics* should comply with the principles laid down in the Declaration of Helsinki; Recommendations guiding physicians in biomedical research involving human subjects. Adopted by the 18th World Medical Assembly, Helsinki, Finland, June 1964, amended by the 29th World Medical Assembly, Tokyo, Japan, October 1975, the 35th World Medical Assembly, Venice, Italy, October 1983, and the 41st World Medical Assembly, Hong Kong, September 1989. You should include information as to whether the work has been approved by the appropriate ethical committees related to the institution(s) in which it was performed and that subjects gave informed consent to the work.

DOES YOUR STUDY INVOLVE HUMAN SUBJECTS? Please cross out whichever is not applicable.

~~**Yes**~~

No

If your study involves human subjects you MUST have obtained ethical approval.

Please state whether Ethical Approval was given, by whom and the relevant Judgement's reference number

DOES YOUR STUDY INVOLVE ANIMAL SUBJECTS? Please cross out whichever is not applicable.

~~**Yes**~~

No

If your study involves animals you must declare that the work was carried out in accordance with your institution guidelines and, as appropriate, in accordance with the EU Directive 2010/63/EU**.** http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32010L0063

This information must also be inserted into your manuscript under the acknowledgements section prior to the References.

If you have no declaration to make please insert the following statements into your manuscript:

Competing interests: declared

Funding: declared

Ethical approval: Not required