Research Article

Edge-Based Convolutional Neural Network for Improving Breast Cancer Prediction Performance

Madhu (ID) **and Raman Kumar** (ID)

Department of Computer Science and Engineering, I K Gujral Punjab Technical University, Kapurthala, Punjab 144603, India

Correspondence should be addressed to Madhu; madhu.dahiya2588@gmail.com

There are many research studies in the field of breast cancer prediction, but it has been observed that the time taken for prediction needs to be reduced. The problem in the existing research is space consumption by graphical content. The proposed research is supposed to minimize the prediction time and space consumption. In this paper, research has focused on the study of existing breast cancer research and techniques and eliminating their limitation. It has been observed that when the number of datasets increases, every comparison makes a huge gap in size and comparison time. This research proposes a methodology for breast cancer prediction using an edge-based CNN (convolutional neural network) algorithm. The elimination of useless content from the graphical image before applying CNN has reduced the time consumption along with space consumption. The edge detection mechanism would retail only edges from the image sample in order to detect the pattern to predict breast cancer. The proposed work is supposed to implement the proposed methodology. A comparison of the proposed methodology and algorithm with the existing algorithm is made during simulation. The proposed work is found to be more efficient compared to the existing techniques used in breast cancer prediction. The utilization of proposed in the work area of medical science is supposed to enhance the capability in case of CNN at the time of decision-making. The proposed work is supposed to be more accurate compared to the existing works. It has been observed that the proposed work is fourteen to fifteen percent more accurate. It is taking 9/4 times less space and 1.0849004/0.178971 times less time compared to the general CNN model. Accuracy might vary as per size of the image and alteration performed in dataset of the image.

1. Introduction

There are several research studies on breast cancer detection [1]. Such research studies are beneficial to capture the symptoms of breast cancer in patients. Such type of research plays a significant role in predicting the probability of breast cancer [2]. The applications of CNN can be found in medical imaging since the 1990s.

"Transferability" is set in pretrained CNN. It is an important aspect of CNN [3]. According to earlier research, in the field of medical imaging, transfer learning has two parts. Initially, a pretrained system is used for the removal of qualities, and in a second group, remaining system has been applied as in the first one except that the logistic layer is used in place of a fully connected layer. Thus, in the present research, proposal of applying an edge detection mechanism to enhance the efficiency in case of the previous convolutional neural network model has been illustrated. There have been limited works in the area of breast cancer [4] prediction model. This is because the nature of work is very complex. Research is considering CNN-dependent graphical processing to perform prediction of breast cancer [5].

Breast cancer has been considered a unique type of issue where cells of the breast get uncontrolled. Cancer stays nonaggressive until cells are not coming from the tubes [6]. Cancer may spread from the lymph glands. After that, it may spread to the other parts of the body. Mostly, it has been seen that breast cancer begins from ducts, which transfer the milk to the nipple. Such kind of cancer is referred to as ductal cancer. On the other hand, another kind of cancer [7] begins in the glands. Glands make the breast milk. It is lobular cancer. This type of disease also occurs in males. This disease

starts from the cells of the lobules. In these glands, milk formed. This is also known as the ducts. Milk flows from lobules to nipple. Medical examiner has been found eligible to provide data regarding cancer. According to several reports, examiner could check whether the cancer cells are inside or outside the milk ducts.

It has been observed that nonaggressive cancer stays inside lobes. These are growing in tissues that are not normal. Noninvasive cancers have been termed carcinoma in situ. It is also known as the precancerous stage.

Invasive cancers usually grow in the usual health tissues. Several cancers come under this category. The types of treatment that you have will depend on the type of cancer you have and how advanced it is. It depends on the response of a patient to treatments they got.

It is possible to observe both types of cancer through common samples. It means that a portion of cancer is growing in normal tissue. The portion of cancer is remaining inside milk ducts. This case is considered the case of invasive cancer.

In some cases, it is possible that breast cancer [8] may contain a combination of tumors. It means that it contains a combination of two or more cancerous cells. These cells may be lobular cells. These types of cancer have become famous in the form of aggressive cancer of the breast. It may be called undercover cancer of the breast. It is usually treated in the form of a ductal tumor.

When multiple tumors are present in the breast, then this disease becomes a cancer whose focus is unpredictable. In case of multifocal breast cancer, real tumor gives rise to other tumors. They are located in the same section of the breast. In case of multicentre cancer, the tumors are produced individually. They are produced on the various locationd of a breast.

In most cases, you can expect the breast cancer to be classified as one of the following:

(1) Cancer in the breast in actual status of duct: this has been defined as a nonaggressive condition which remains in the tube of milk.

(2) Cancer in the breast [9] in actual status of lobes: it is a type of cancer in which cells expand in the milk producing gland of the breast. It is similar to the above type because it also remains inside the milk tubes. Actually, it is not a cancer, but it only shows the warning sign of risk which increased continuously. When this risk increased, it will develop aggressive cancer. This cancer may occur in any breast in the future.

(3) Aggressive duct cancer is a type of breast cancer which is very common. It starts in the breast milk duct and after that, it attacked adjoining tissues of the breast.

(4) Varieties of aggressive duct cancer: they consist of tube-shaped bone marrow. They are also considered mucinous and papillary along with cribriform carcinomas of the breast [10]. Cells could look and behave in an abnormal way compared to invasive ductal carcinoma cells in such types of cancers.

(5) Aggressive lobe cancer: it usually begins inside the lobule and after that, it attacks adjoining tissues of the breast.

(6) Inflammatory breast cancer: it has been generally started with reddening/swelling of the breast. Usually, it is not started in a distinct lump.

(7) Male breast cancer occurred and then it becomes ductal carcinoma.

(8) Paget's disease of the nipple is a cancer that starts from the nipple duct. But as it grows, it will make a huge impact on the skin and areola of nipples. It is a form of breast cancer.

(9) Phyllodes cancer: it is the rarest type of cancer [11], which is related to mammary glands. It starts in the connective tissue of the breast. It keeps growing quickly. It grows in a leaf-like pattern. Many are cancerous; on the other side, several are not.

(10) Recurrent and metastatic breast cancer is returned after treatment. Sometimes, it spread from the breast to other parts of the body.

1.1. Stages of Cancer Which Is Related to Mammary Glands. On the basis of tumor size and its spread, this type of cancer is divided into two stages. Cancers in which the size of the tumor is large make a huge impact on adjoining tissues as well as organs, which are at upper status compared to other cancers. These might be small and stored inside the breast. Doctors should confirm the size of the tumor [12] at other stages of cancer. In zero stage, this disease remains in the tube of the breast. On the basis of tumor size and lymph node involvement, the stage (one, two, or three) has been decided. Up to the last stage, cancer related to mammary glands extends to other parts of the body.

2. Literature Review

A lot of researches have been done in order to predict breast cancer. Researchers proposed the latest mechanism to perform the detection of cancer in the breast. SVM and Bayesian networks have been used to perform prediction of cancer. Comparison of ML classifiers is performed to diagnose cancer in the breast on the selection of characteristics. They also wrote on Deep CNN Model to find reaction in case of breast cancer. Their work is explained as quantitative ultrasound as well as texture predictors in case of tumor in breast. Previous works proposed a mechanism to detect cancer in the breast with the support of classifiers.

In 2017, Tan et al. [1] proposed the research on breast cancer detection with the use of CNN in contrast to the breast imaging system. Here, the goal of breast cancer detection using CNN (BCNN) is to speed up the diagnosis process.

In 2016, Spanhol and Oliveira [2] have analyzed breast cancer histopathological image categorization. Convolution neural networks are used to classify breast cancer histopathological images from BreakHis, a public dataset available at http://web.inf.ufpr.br/vri/breast-cancer-database.

Cells related to folded layers are used to find traits of close pattern in case of graphics. In order to obtain the same traits for the input graphic channels, weights of each nerve have been distributed between the nodes in the convolutional layers. Such network is skilled. This could supply input in direction of the first level.

In 2018, Charan et al. [3] evaluated breast cancer detection in mammograms. In the research work, CNN is utilized. CNN had been instructed here with the complete set of unusual pictures that were one hundred and thirty-three. Different filter sizes in addition to preprocessing techniques have been applied. These techniques are applied on real data. It has been done to neglect the noise element. Proper segmentation has been considered mandatory. It has been done to efficiently extract the feature and classify them.

In 2018, Nawaz et al. [13] have proposed organization of cancer that is having various groups and it is related to breast. In this research, CNN is used. The basic objective of this approach is to make the classification of breast tumors.

In 2019, Jiang et al. [6] discussed breast cancer histopathological image categorization. In the research work, the convolutional neural network has been used. They make use of a small SE-ResNet module. The research work is proposing the design of a novel CNN. Convolutional layer has been considered here; along with this, the small SE-ResNet modules are also determined here. Here, the third and final are also determined which are a fully connected layer. This research has proposed a small SE-ResNet module. It has been known as the growth of a mixture of the residual module.

In 2019, Ragab et al. [7] have described cancer detection in the breast. CNN is used in research. They also consider the support vector devices. An innovative CAD system has been presented in the research work. They also used segmentation concepts. The first concept includes the determination of the region of interest (ROI) basically.

In 2019, Kontopodis et al. [11] investigated biomarkers in addition to their function. They also considered the model-free graphical biomarkers. The results have suggested the model-free DCE-MRI IBs. They are new and strong substitutes. These graphical biomarkers are very difficult.

In 2014, Thomas et al. [10] have written research to find breast cancer. They make use of weighted clinical classifiers in research work. Researchers are indicating an efficient mathematical layout and explained fusion with nonlinear categorization problems.

In 2018, Czamota et al. [9] explained quantitative ultrasound that is making use of texture predictors in case of breast tumor. They are analyzing quantitative ultrasound predictors. The issue is that the proposed system is working at a slow speed. Imaging biomarkers of 35 patients having breast cancer are taken into account. All of these 35 patients have undergone neoadjuvant chemotherapy (NAC). The data have been collected from the records of active contrast-enhanced MR imaging (DCE-MRI). With the help of two distant methods, these data have been collected by them. Out of these two methods, the first one is a model-free method that is derived from pattern recognition (PR), and the second one is model-based using pharmacokinetic compartmental modeling. It has been noticed by them from the evaluation of model-free as well as model-based biomarkers which can figure out the entire medical response as soon as the first cycle of NAC is over.

In 2018, Dai et al. [12] have done some work to find cancer that is related to the breast. In this work, a random forest algorithm was adopted. Present territories of medical management have gained a lot of advantages due to the availability of medical big data and machine learning. They can be utilized to gain knowledge. In the current work, by the adoption of a random forest algorithm, they have talked over another case of breast cancer identification and obtain high prediction accuracy. It is useful for secondary medical identification.

In 2018, Ahn et al. [14] submitted a calculation that indicates the impact of the genetic factors in ovarian cancer. A large amount of impactful information has been given by genetic factors. It is possible to forecast the future of several cancers such as breast and colon due to the effective implementation of application. It is implemented to the RNA sequence data which are obtained from the Cancer Genome Atlas (TCGA). From its consequences, it was recognized that whether the patient's future was at either higher or lower risk. The median standard is utilized by them to the arrangement of the available schemes. On the basis of these different schemes, the work was put forward by them for the next work.

In 2012, Anothaisintawee et al. [15] performed an organized survey over the achievement of a model, using which, the risk of breast cancer has been predicted. Provided models are having poor as well as analytical efficiency. The major reason for it has been incomplete data of factors that may cause risk. Another reason is that it is not possible for the community to identify the improvement in inequity itself. Due to this, it is necessary that some fresh methods like net reclassification index should be taken into account. This method can be used to study the improvement of the achievements of the newly manufactured model.

In 2016, Fan et al. [16] submitted work on structure-leveraged methods in order to find out how much helpful is the research in forecasting the risk of breast cancer. In the field of medical science, it was always required that the risk of breast cancer must be identified. If the risks are identified, then it is possible to give precision medicine to the patient. The main intention behind this research is to present fresh penalized methods. They were helpful because identification of breast cancer risk is improved. In order to achieve this, structure data which were available in electronic health records were exploited.

In 2017, Mulatu [17] conducted a study on the data mining techniques in order to find out how much helpful is the work in forecasting breast cancer repetition. At present, breast cancer becomes a very common disease in this universe. At first, it was assumed that it was female-affected disease, but in the situation, it is possible that males are also affected by this disease. The meaning of repetition is that it comes back. With the help of the work which was done in this research, it is easy to recognize the algorithm by which the reason for breast cancer repetition is efficiently

identified. In addition to this, it will also increase the efficiency of this algorithm.

In 2017, Hamouda et al. [18] reported that a brilliant arrangement for identification of breast cancer as well as for their treatment was submitted by them. A disease in which the growth of body cells does not handle is known as cancer. From the last few years, this type of cancer becomes a reason for women's death. It happens because of delayed identification of breast cancer. So, they have proposed a technique that contains two parts: in the primary part, it depends on the smart and effective use of the Rough Set Theory (RST) technique. With the help of this technique, they examine the dataset of breast cancer and figure out desired sets. They also increase the efficiency of diagnosis and figure out the essential aspects of data. In the next part, a program was manufactured by them in order to identify and cure the patient with breast cancer depending upon decision rules. It was produced on the basis of the Rough Set Theory dataset and the data gathered information from the patient record. Due to this system, it is a possibility that doctors can identify the symptoms of breast cancer in the early stages. In addition to this, it was also helpful in keeping track of the medical history of any patient.

In 2018, Lin et al. [19] explained the Fast Scan Net, speedy and opaque examination of multigiga pixel whole-slide images due to which growth of cancer cells was discovered. The growth of the lymph node becomes a significant sign. It can be easily noticed by pathologists with the help of the microscope. For the transformation of the model, a fresh method in which layers are held securely was submitted by them. It influences the effectiveness of a fully complex architecture in order to fulfill the essential requirement in clinical practice. In addition to this, they also made the network more skilled. The effectiveness of this method is confirmed on the standard dataset of the 2016 Camelyon Grand Challenge. If a comparison is made between their method and the state-of-the-art method, it is found that a serious boost is shown by their method on tumor localization accuracy at a very high rate. It can beat even human achievements on both challenging works.

In 2018, Chaurasia et al. [20] did forecasting of gentle and nasty breast cancer, which was done with the help of the data mining techniques. In the list of leading cancers, breast cancer comes in a second position by which females are affected in comparison to different types of cancers. From a historical record, it was found that there are near about 1.1 thousand cases in 2004. At present, this disease was noticed in its early stages due to the enhancement of industrialization as well as urbanization. In addition to this, it was also noticed in its early stages due to the availability of different types of equipment. In past, it was more general in nation where earning is high. But with the passage of time, it becomes common in the nation where earning is middle and low. From a record, it was observed that more than 50% of cases of breast cancer are fatal. Therefore, it becomes the primary reason of death from cancer in the female. One of the primary goals of this work is to submit details regarding breast cancer. In this work, they took the help of accessible scientific advancements. With the help of this, they create

prediction models for breast cancer survivability. For the formation of forecasting models, 3 trendy data mining algorithms were adopted by them by considering a large dataset of 683 breast cancer patients. In addition to this, to determine the balanced approximation of the three prediction models, 10-fold cross-validation methods were also adopted by them.

In 2018, Chauhan and Swami [21] conducted work on breast cancer forecast by adopting the assembly approach derived from the genetic algorithms. Identification of breast cancer is an unwrapped region of study. Breast cancer prediction is analytically difficult. It can be resolved by the adoption of machine learning models like a decision tree, random forest, and support vector machine. Each and every model possesses qualities and loopholes. For the forecasting of breast cancer, improvement in the exactness of models is required. A fresh method which was the weighted average ensemble method of categorization dataset was submitted in this work. They have worked on genetic algorithms. In this method, the loopholes which are present in the traditional weighted average method were eliminated.

In 2018, Fu et al. [22] predicted invasive disease-free survival in case of the initial stage of cancer in the breast. It can be achieved by the utilization of clinical data records. The females of China are dangerously exposed to breast cancer with high morbidity and deaths. It was almost impossible for doctors to organize a suitable treatment plan due to the lack of robust forecasting models that may make the life of a patient a little bit longer.

In 2018, Kaushik [23] stated the survival forecasting that is postsurgical in case of cancer in the breast. They have provided the novel concept related to the postsurgical survival forecasting of the breast cancer patient. Cancer is one of the most common deaths causing disease. Basically, the chance of this disease is more in women than men. The period after the breast cancer surgery of a patient has become very challenging. After the surgery, it is very difficult to make decrement in the death rate. The survival prediction has been connected to the life of a woman. Therefore, it is necessary to use efficient algorithms in order to predict breast cancer in women. Several prediction algorithms are introduced for breast cancer prediction. The existing algorithm such as the postsurgical survival (PSS) prediction has been used for the last three decades. Such concepts include statistical or machine learning techniques. These are used in order to estimate the life period of a patient that is suffering from breast cancer. The work has proposed the novel prediction concept related to prediction related to cancer in the breast. In research work, the SVM transmission is capable of sharing dual coordinate ascent. The authenticity of the proposed model has been presented with the use of SBT and Apache Flink. These are the tools used in the research work.

In 2018, Ma et al. [24] proposed the novel two-stage deep technique. This technology has been used for mitosis estimation in breast cancer histology graphics. It is very necessary to make detection accurately. Along with this, the counting of mitosis has been determined essential for computer-aided diagnosis. This has been physically performed with the help of a pathologist. It varies as per the

clinic experience of the patient. Therefore, it has been found that the procedure takes a lot of time. In addition, it provides the results in low agreement among the pathologists. If one makes the investigation, he will found that the use of several computer-aided detection techniques is available to use. In the research work, the researchers presented a novel two-stage detection technique. In the research work, the multiscale with similarity learning convents (MSSN) has been used. First, the huge sizes of possible candidates are created in the first stage. It has been done in order to decrease the FN. For this purpose, different square and nonsquare filters have been used. It is necessary to detect the spatial relation out of different scales. Second, a similar prediction model has been consequently executed on the required candidates. It has been done for the last detection. This step is made to reduce FP.

In 2018, Chang et al. [25] wrote on breast cancer prediction. Research is making use of Bayesian logistic regression. Work is considering various qualities. Forecasting that is related to cancer in the breast has been made.

In 2018, Rampun et al. [26] discussed breast mass classification in mammograms. In the research work, the ensemble convolutional neural networks have been used. The work proposed the quantitative results for the preliminary review. It has been undertaken as part of decision support with a data arrangement environment. These methods are related to breast cancer. DESIREE has been considered a project that is European-funded. This is utilized in order to make improvement in breast cancer. It is possible with graphic-based data architecture. In the research work, the researcher has provided an explanation of the use of ensemble deep learning. It has been done to classify the breast mass in mammograms. In the proposed work, Alex Net has been used in order to resolve the challenges. It has been done in order to deal with the issues related to the classification. Then, model selection has been executed in order to choose the best three results dependent on the highest validation accuracies. It has been done at the time of the validation phase. In conclusion, forecasting depends upon a standard chance in conjunction with this model. It has been clearly indicated from the practical outputs that precision commencing by a single system lies in the middle of seventy-five percent and seventy-seven percent.

In 2018, Sun et al. [27] provided a capable mixed mode. It is showing on behalf of genetics. This is expressing another way that is related to LR-RF. Research work has focused on breast cancer that is expressed genes. In the research work, the gene microarray datasets of breast cancer GSE15852 and GSE45255 are used. The downloading of such datasets is made from GEO. In the research work, the logistic regression with random forest algorithm is considered. In the research work, the novel technique that has been known by the name LR-RF has been proposed. Such work is proposed in order to select the severally expressed genes related to breast cancer on microarray data by the Bonferroni test of FWER error calculations. The result has indicated the facility of LR-FR. It is clear that it has a great role to select the genetics which is expressed in different ways. The accurateness of standard forecasting has

been obtained by the LR-RF. The average prediction accuracy of the proposed LR-RF from replicating random test 10 times surprisingly reaches 93.11 percent with variance as low as 0.00045. This research has provided a prediction accuracy rate of up to 95.57 percent. Additionally, in the research work, they analyzed the gene interaction networks. It has been observed that the most top 20 genes are included for developing breast cancer. The results of the research work have demonstrated the reliability for efficiency of LR-RF. It has been observed that the proposed work is very beneficial in the field of breast cancer prediction. They explained that LR-RF is able to offer new information with the techniques used by biologists. They also ensure that it is also made in use in the field of medical sciences. Along with these fields, it has been used in cognitive computing. The cause is that it is able to identify the disease-based genes of breast cancer.

In 2018, Xiao et al. [28] stated the breast cancer diagnosis. In the research work, they have used a feature extraction algorithm. The proposed algorithm is based on deep learning. Breast cancer has been known as a disease, which is mainly identified in females. It becomes very necessary to test and diagnose these types of cancer without any delay. Along with the use of detection technology, an increasing size of patients is handled by medical statistics and high level qualities. Such features are very helpful to treat the disease. Helpful contribution is provided by the high-level qualities before the commencement of analytical methodologies. Thus, here, the extraction of the essential data has been made. After data reduction, the feature dimension has been determined. Therefore, the traditional work is on breast cancer diagnosis. For this purpose, they have used the labelled data. It has been observed that it was very difficult to get these data. To resolve such issues, the research has provided a new scheme. The proposed work is the integration of deep learning with the stacked autoencoders. In deep learning, they have used the unsupervised feature extraction algorithm. In the research work, support vector machine model is also discussed. Such integration has been proposed in order to make the model for breast cancer diagnosis.

In 2018, M. Gupta and B. Gupta [29] offered an ensemble model. It has been proposed in order to provide assistance in breast cancer prediction. For this purpose, they have used the sequential least squares programming technique. Cancer has been known as the integration of several diseases. Cancer takes place in body when the cell expands in a way that becomes uncontrolled. Out of various cancers, cancer related to the breast has become a very common cancer. The reason is the abnormal growth of the breast. If there is cancer in one part of the body, it spreads to all parts of the body. Cancer related to the breast has different stages. The last is very critical. The research contains ensemble that has been proved to be more accurate than the traditional single classification system. Researchers have used the SLSQP technique. It has been used in order to allocate every form of organization. The forecasting done by all organizers has been mixed with the use of the soft voting methods.

3. Materials and Methods

3.1. Problem Statement. Existing researches made use of breast cancer categorization of graphical content with help of CNN [4], while some integrated deep convolution neural networks with support vector machines [7]. The authors have also used random forest algorithm for breast cancer diagnosis [12], genetic algorithm-based ensemble approach [22], and Bayesian logistic regression [26] for breast cancer prediction. Breast mass classification and its diagnosis have been made using mammograms using ensemble convolution neural networks [27]. Some researchers are performing the extraction of features in an unsupervised fashion with the help of deep learning [29].

According to a literature review made by different researchers, CNN is performing efficiently with graphical evaluations as well as categorization of contents that are graphical. So, there has been a requirement to perform more tasks on the breast cancer prediction model with the support of CNN. But the limitations of the existing CNN model are the space consumption and comparison time. It takes a lot of time during comparison of graphical content. Moreover, there is a need to reduce space and increase the accuracy during comparison. Therefore, the performance of the traditional CNN model needs to be improved.

3.2. Proposed Work. The proposed work is focusing on the study of existing breast cancer research and techniques. After considering the requirement of space and time minimization and accuracy maximization, research proposes a methodology for breast cancer prediction using an edge-based CNN (convolution neural network) algorithm. The proposed work is supposed to implement the proposed methodology. A comparison of the proposed methodology and algorithm with the CNN algorithm has been made. The proposed work is supposed to be more efficient compared to CNN used in breast cancer prediction. The use of edge detection would eliminate the useless content from graphical data. This would result in reduction of time during CNN processing. This will also reduce the requirement of storage space.

Table 1 represents how the proposed work is better than previous approaches.

3.3. Process Flow of Proposed Work. The following steps are involved in the work which is discussed here:

Phase 1: the image base of Benign, InSitu, and Invasive would be created.

During this phase, the image that consists of features to confirm breast cancer is considered. Phase 1 focuses on acquiring graphical contents for further processing.

Phase 2: apply a traditional CNN classifier in order to check the space and time consumption. This phase is considering the time and space consumed if the image is not edge-based.

Phase 3: apply the edge detection mechanism on the image set.

The use of an edge detector reduces the size of the image by eliminating the useless contents. Moreover, it takes less time to compare the graphical contents.

Phase 4: apply the proposed CNN classifier in order to check the space and time consumption.

The factors for comparison are found at this phase. The time and space consumption in case of the proposed dataset is considered.

Phase 5: compare the performance as well as space consumption in case of existing and proposed work. The result of the comparison confirms the time consumption, space taken, and accuracy for the proposed work.

Figure 1 represents the process flow of the proposed work.

3.4. Classification of Dataset. The image sample has been categorized into six sections, that is, Benign, InSitu, Invasive with edge detection, Benign, InSitu, and Invasive without edge detection. During the simulation process, the categories that are nonedge-based would take more time compared to the contents that are edge-based.

Several lumps have been benign. It is presenting that there is no cancer. Benign breast lumps usually are having smooth edges. It could be transferred when it is pushed against them. Such edges are usually available in case of every breast. Many causes are there that are including normal modification in breast tissue. There could be an infection or injury in the breast.

Table 2 represents types of images, that is, Benign, InSitu, and Invasive before edge detection and after edge detection.

4. Results and Discussion

The dataset of Benign, In Situ, and Invasive is stored in the relevant directory. Then sample taken from the patient is stored for breast cancer prediction. These samples are compared to stored samples in order to predict cancer considering features extracted from the present dataset using a convolution neural network mechanism [27]. The size of the dataset is taken along with the comparison time. Moreover, the percentage of matching in order to consider the accuracy is also taken. Then the edge detection mechanism is applied over the same dataset in order to eliminate the use of information from the graphical dataset. This results in a reduction in the size of dataset images. The image of fresh samples is also processed by an edge detection mechanism. Then the convolution neural network mechanism is applied over the dataset. Then dataset size, time of operation, and matching percentage are taken.

4.1. Image Dataset. The same dataset of Benign, In Situ, and Invasive has been considered in research. The dimension of the graphical sample is 204×154.

TABLE 1: Comparative analysis of previous research studies to the proposed work.

| Researches	Technology used	Key benefit	Performance factor	Storage space factor	Accuracy factor	Issues and challenges
[4]	CNN	Efficient image classification	No	No	No	Time consumption and space consumption need to be reduced
[7]	CNN and SVM	Fast performance	Yes	No	No	The issue of space consumption and accuracy not resolved
[12]	Random forest algorithm	Handling missing values, no feature scaling required, and less impacted by noise	No	No	No	Complexity and long training period
[22]	Genetic algorithm	It provides good quality solutions in a less time	Yes	No	No	Does not provide the optimal solution
[26]	Bayesian logistic regression	It provides better result that is unbiased, with lower variances	No	No	Yes	Logistic regression is capable of predicting a categorical outcome
[27]	Ensemble convolution neural networks	It offers increased flexibility	Yes	No	No	Does not consider the space and accuracy
[29]	Unsupervised feature extraction algorithm	Ideal to explore raw and unknown data	Yes	No	No	There is a lack of accuracy due to unavailability of labels
Proposed work	CNN and edge detection		Yes	Yes	Yes	The integration of multiple technologies is quite challenging
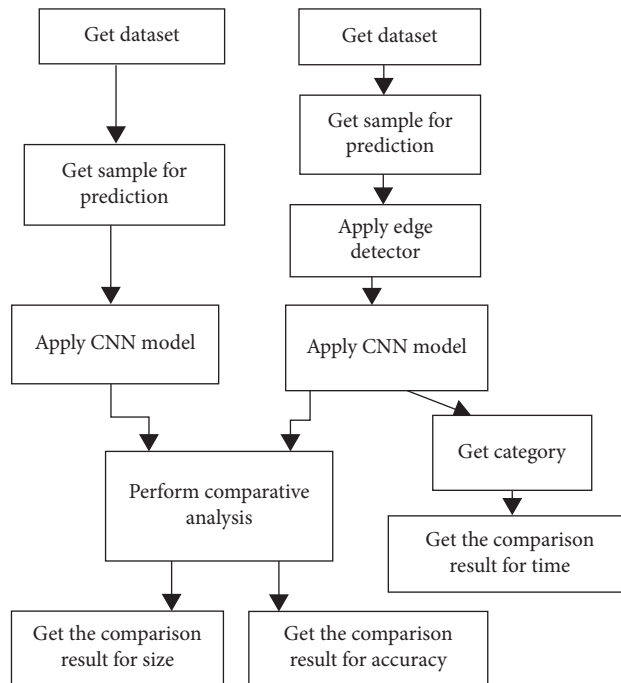
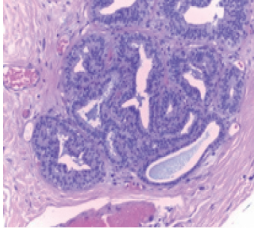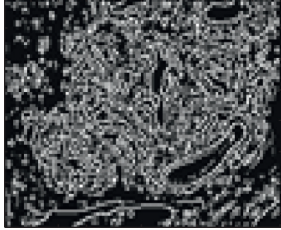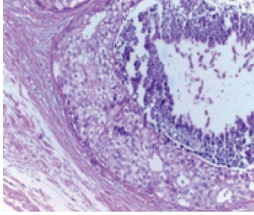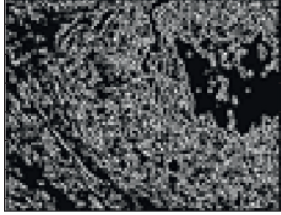FIGURE 1: Process flow of the proposed work.

Table 2: Classification chart.
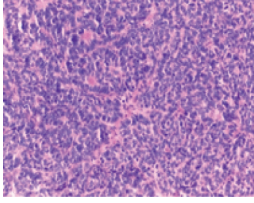
Types of images	Before edge detection	After edge detection
Benign (benign)		
InSitu (malignant)		
Invasive (malignant)		

Figure 2 shows dataset for Benign, In Situ, and Invasive cancer images.

4.2. Implementation of the Proposed Canny-Based CNN Classification

4.2.1. Edge Detection Model.
Researchers have utilized a Gaussian smoothed step edge. Thus, a one-dimensional image f that has exactly one edge positioned at $x=0$ is modeled as

$$f(x) = \frac{I_r - I_l}{2}\left(\text{erf}\left(\frac{x}{\sqrt{2}\,\sigma}\right) + 1\right) + I_l. \tag{1}$$

On the left side of the edge, intensity has been considered as $I_l = \lim_{x \to \infty} f(x)$.

Right of the edge has been considered as

$$I_l = \lim_{x \to \infty} f(x). \tag{2}$$

Scale parameter σ is considered as blur scale of edge. Such scale parameter is needed to be set according to image quality in order to restrict damage of true edges.

4.2.2. First-Order Mechanism.
Various gradient operators are used to check image gradients input graphics. The mechanism is to make use of central differences:

$$Q_x(x, y) = -\frac{1}{2}Q(x - 1, y) + 0 \cdot Q(x, y) + \frac{1}{2} \cdot Q(x + 1, y), \tag{3}$$

$$Q_y(x, y) = -\frac{1}{2}Q(x, y - 1) + 0 \cdot Q(x, y) + \frac{1}{2} \cdot Q(x, y + 1). \tag{4}$$

which are corresponding to the application of the following filter masks of image data:

$$Q_x = \begin{bmatrix} +\frac{1}{2} & 0 & -\frac{1}{2} \end{bmatrix},$$

$$Q_y = \begin{bmatrix} +\frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{bmatrix} Q. \tag{5}$$

Well-known and previous Sobel operator is depending on filters specified as follows:

$$Q_x = \begin{bmatrix} +1 & 0 & -1 \\ +2 & 0 & -2 \\ +1 & 0 & -1 \end{bmatrix},$$

$$Q_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}. \tag{6}$$
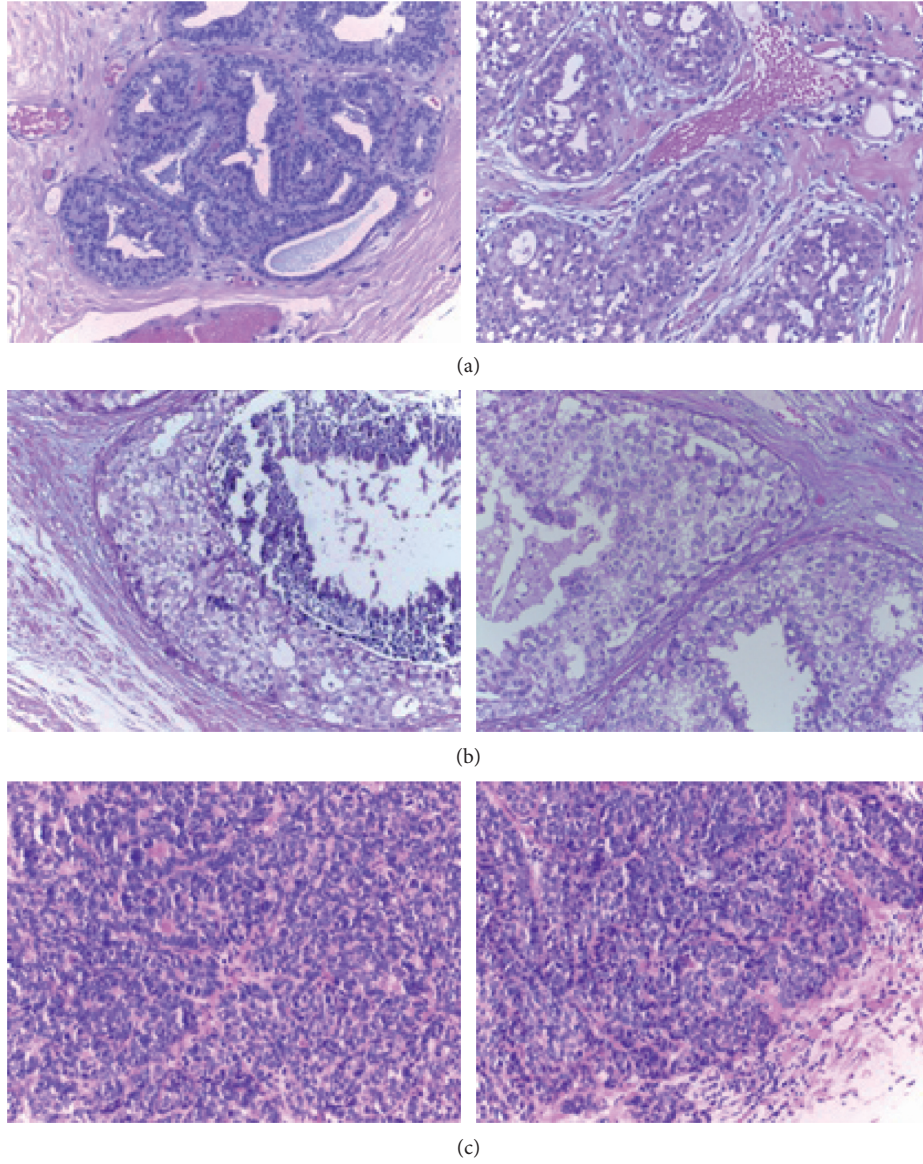
(a)

(b)

(c)

FIGURE 2: Dataset of Invasive, InSitu, and Benign. (a) Dataset for Benign. (b) Dataset for InSitu. (c) Dataset for Invasive.

Considering first-order image derivatives, gradient magnitude has been calculated as follows:

$$|\nabla Q| = \sqrt{Q_x^2 + Q_y^2}. \tag{7}$$

On the other hand, gradient orientation could be calculated as follows:

$$\theta = a \tan 2 \left(Q_y + Q_x\right). \tag{8}$$

4.2.3. Second-Order Approaches.

Edge detection operators might consider second-order derivatives of intensity.

(1) Differential. A more refined second-order edge detection mechanism finds edges automatically using subpixel accuracy. Such mechanism is utilizing the following differential mechanism of finding zero crossings of second-order directional derivative in case of gradient direction.

Following the differential geometric way of expressing the requirement of nonmaximum suppression proposed by Lindeberg, let us introduce at every image point a local coordinate system $(u; v)$, with a direction parallel to gradient direction. Assuming that image has been presmoothed by Gaussian smoothing and a scale space that represent $Q(x, y, t)$ equation at scale has been computed, we can require that gradient magnitude of scale space representation, which is equal to first-order directional derivative in vv-direction Q_v, should have its first-order directional derivative in v-direction equal to zero

$$\partial_v \left(Q_v\right) = 0, \tag{9}$$

while second-order directional derivative in vv-direction $Q_v Q_v$ should be negative; that is,

$$\partial_{vv}(Q_v) \le 0. \tag{10}$$

Written out as an explicit expression in terms of local partial derivatives Q_x, Q_y, ..., Q_{yyy}, this edge definition can be expressed as zero-crossing curves of differential invariant:

$$Q_v^3 Q_{vvv} = Q_x^3 Q_{xxx} + 3Q_x^2 Q_y Q_{xxy} + 3Q_x Q_y^2 Q_{xyy} + Q_y^3 Q_{yyy} \le 0, \quad \text{where } Q_x, Q_y, \ldots, Q_{yyy}. \tag{12}$$

Denote partial derivatives computed from a scale space representation QQ obtained by smoothing the original image with a Gaussian kernel.

$$Q_{xx}(x, y) = Q(x - 1, y) - 2Q(x, y) + Q(x + 1, y),$$

$$Q_{xy}(x, y) = \frac{1}{4}(Q(x - 1, y - 1) - Q(x - 1, y + 1) - Q(x + 1, y - 1) + Q(x + 1, y + 1)), \tag{13}$$

$$Q_{yy}(x, y) = Q(x, y - 1) - 2Q(x, y) + Q(x, y + 1),$$

which are corresponding to the following filter masks:

$$Q_{xx} = [+1 \ -2 \ \ 1]Q,$$

$$Q_{xy} = \begin{bmatrix} -\dfrac{1}{4} & 0 & \dfrac{1}{4} \\ 0 & 0 & 0 \\ \dfrac{1}{4} & 0 & -\dfrac{1}{4} \end{bmatrix}Q, \tag{14}$$

$$Q_{yy} = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}Q.$$

4.3. Initial Setup. In Figure 3, the initial setup is shown. There are three folders named benign, malignant, normal which have been created. The graphical contents are stored in this folder. Then the CNN classifier module has been created that would perform CNN-based breast cancer prediction.

4.4. Results of Simulation. During simulation, the image sets of Benign, InSitu, and Invasive are considered. In Figure 4, the status of samples before applying edge detection is shown.

As the canny edge detector is applied to the above dataset. Then, in Figure 5, the edge-based image dataset is produced.

$$Q_v^2 Q_{vv} = Q_x^2 Q_{xx} + 2Q_x Q_y Q_{xy} Q_x^2 Q_{yy} = 0. \tag{11}$$

That satisfies a sign condition on the following differential invariant:

In practice, the first-order derivative approximations can be computed by central differences as described above, while second-order derivatives can be computed from scale space representation QQ according to

4.4.1. Canny Edge Detector. To smooth the image, a Gaussian filter is applied as follows:

$$s = G_\sigma * I,$$

$$G_\sigma = \frac{1}{\sqrt{2\pi\sigma}} e^{-(x^2 + y^2)/2\sigma^2}. \tag{15}$$

Compute X and y derivatives

$$\nabla S = \left[\frac{\partial}{\partial x}S \ \ \frac{\partial}{\partial y}S \right]^T = \left[S_x \ \ S_y \right]^T. \tag{16}$$

Compute gradient magnitude and orientation

$$|\nabla S| = \sqrt{S_x^2 + S_y^2},$$

$$\theta = \tan^{-1}\frac{S_y}{S_x}. \tag{17}$$

As the canny edge detector is applied to the above dataset, the following edge-based image dataset is produced.

Figure 6 shows the convolutional layer model that converts images into Benign, InSitu, and Invasive tumors. CNN is used to extract features from images and then extracts features supplied to a fully connected neural network.

4.5. Simulation of Time Complexity. Time complexity during the CNN operation is lesser in case the proposed work for edge detection has been performed. The algorithm to simulate the time is discussed as follows:

(1) Get the normal sample for breast cancer and apply the CNN module to detect cancer
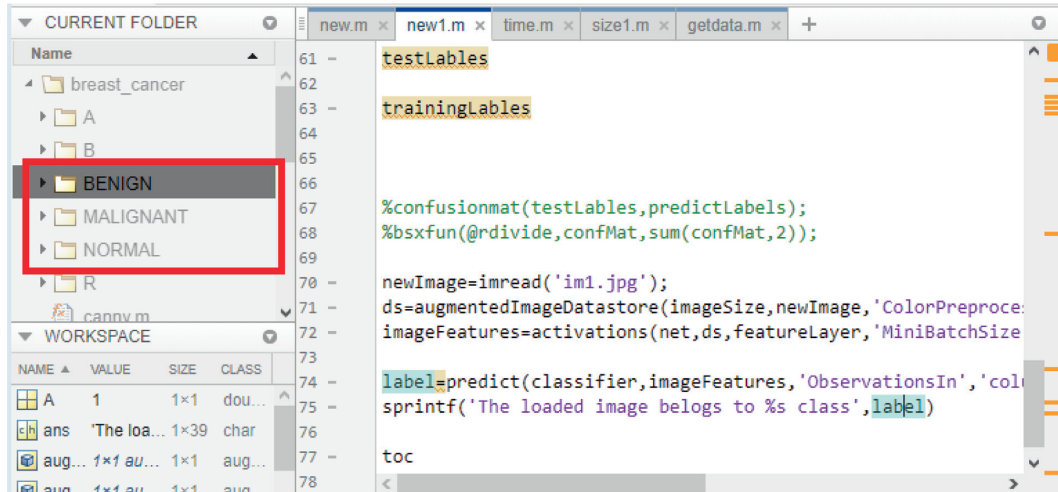
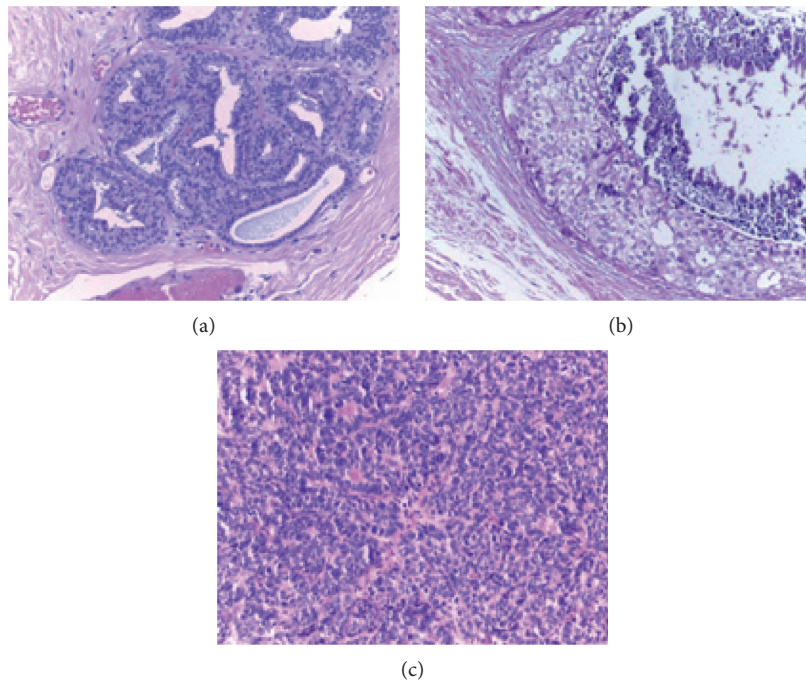(2) Get the time $t1$ during detection

FIGURE 3: Initial setup.

FIGURE 4: Before edge detection. (a) Benign. (b) Invasive. (c) InSitu.

(3) Apply the edge detection on the same breast cancer sample

(4) Get the edge detected sample for breast cancer and apply the CNN module to detect cancer

(5) Get time $t2$ during detection

(6) Get the ratio of $t1$ and $t2$

(7) Create a simulation module to represent the difference in time

It has been observed that the time consumption in case of the CNN approach is relatively more compared to the case when the edge detection mechanism is applied to the samples. The various samples 10 to 100 are taken to perform CNN-based detection. The ratios of both cases are shown as follows:

$$\frac{\text{time taken by CNN}}{\text{time taken by edge} - \text{based CNN}} = \frac{1.0849004}{0.178971}. \qquad (18)$$

Considering (1), the simulation to compare the time consumption in both cases has been performed. Table 3 represents the time consumption in case of CNN and edge-based CNN.

Considering Table 3, Figure 7 compares delay in CNN and CNN with edge detection. The steep time is increasing in previous CNN due to an increase in count of images. Dataset of 70 images has increased the time consumption
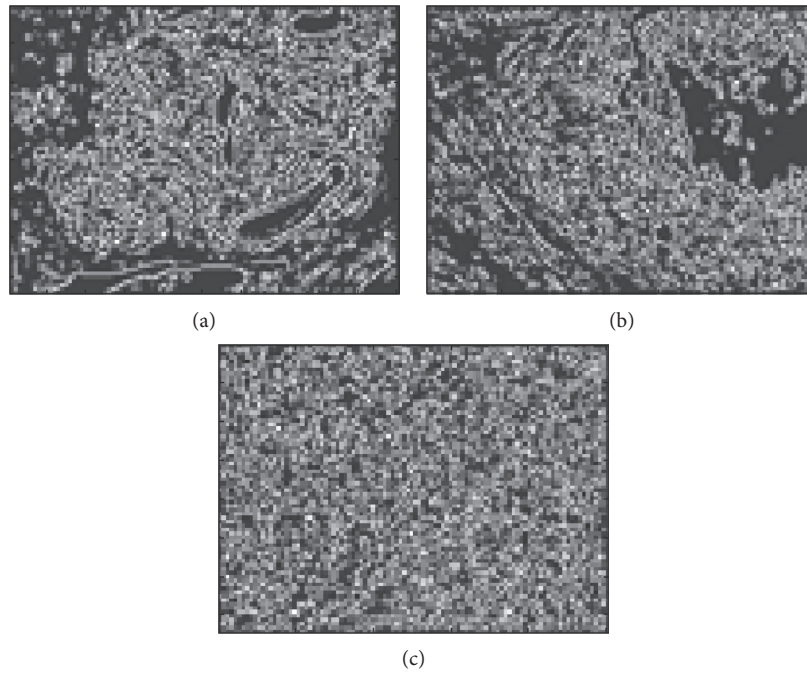
(a)

(b)

(c)

FIGURE 5: After edge detection. (a) Benign. (b) Invasive. (c) InSitu.
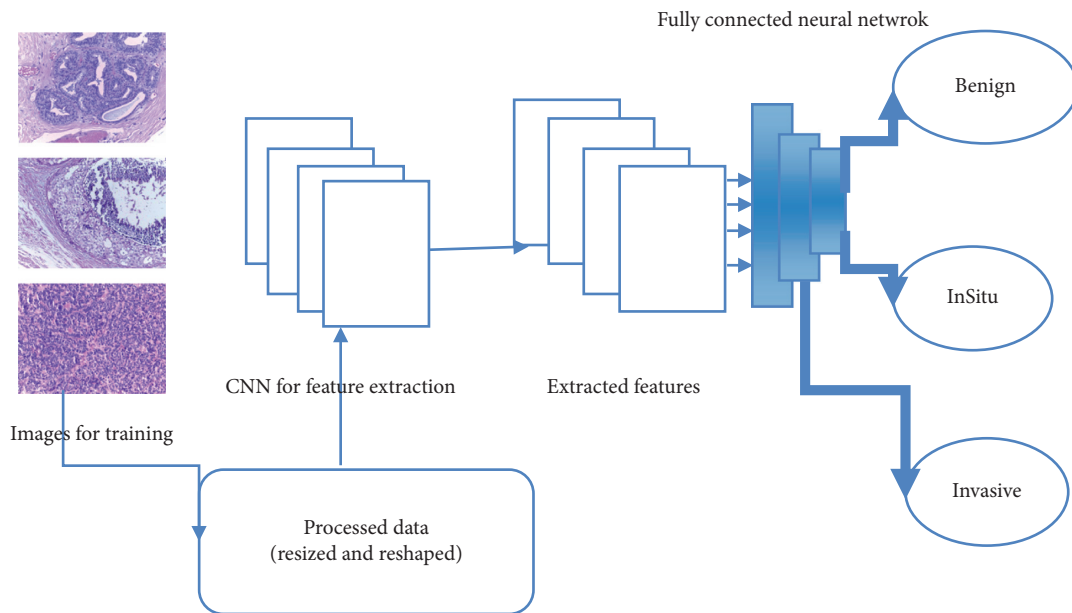

FIGURE 6: Convolutional layer model.

TABLE 3: Time consumption in case of CNN and edge-based CNN.

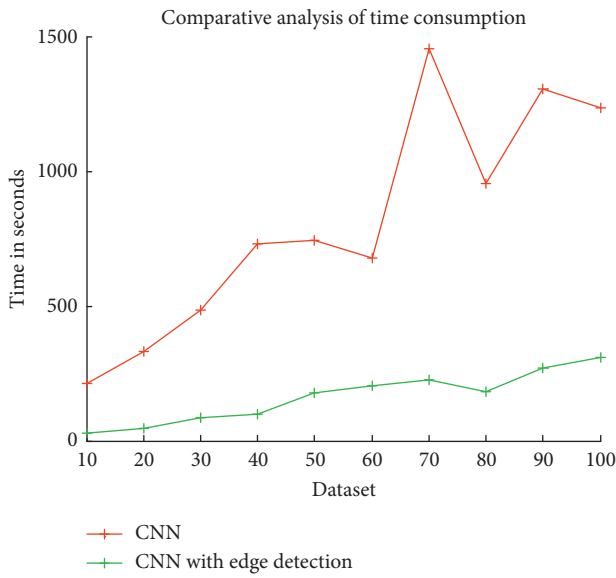| Number of samples	CNN approach	CNN with edge detection
10	$10.849004 * (1 + Rnd)$	$1.78971 * (1 + Rnd)$
20	$21.698008 * (1 + Rnd)$	$3.57942 * (1 + Rnd)$
30	$32.547012 * (1 + Rnd)$	$5.36913 * (1 + Rnd)$
40	$43.396016 * (1 + Rnd)$	$7.15884 * (1 + Rnd)$
50	$54.24502 * (1 + Rnd)$	$8.94855 * (1 + Rnd)$
60	$65.094024 * (1 + Rnd)$	$10.73826 * (1 + Rnd)$
70	$75.943028 * (1 + Rnd)$	$12.52797 * (1 + Rnd)$
80	$86.792032 * (1 + Rnd)$	$14.31768 * (1 + Rnd)$
90	$97.641036 * (1 + Rnd)$	$16.10739 * (1 + Rnd)$
100	$108.49004 * (1 + Rnd)$	$17.8971 * (1 + Rnd)$

FIGURE 7: Simulation of time consumption in case of CNN and CNN with edge detection.

dramatically due to load on CPU due to external process that gets reduced in a dataset of 80 images. But in case of the proposed work, this change has been smartly handled.

4.6. Simulation of Size. Size consumption during the CNN operation is lesser in case proposed work for edge detection has been performed. The algorithm to simulate the size is discussed as follows:

(1) Get the normal sample for breast cancer to apply the CNN module to detect cancer

(2) Get the size $s1$ of the sample

(3) Apply the edge detection on the same breast cancer sample

(4) Get the edge detected sample of the breast cancer to apply the CNN module to detect cancer

(5) Get size $s2$ of sample

(6) Get the ratio of $s1$ and $s2$

(7) Create a simulation module to represent the difference of size in the dataset

In order to simulate the space consumption, the sample size is considered. The image sample used for breast cancer has been taken and their size is checked. The edge detected samples are also considered. From observation, it has been found that the size taken by samples with edge detection is more compared to edge detected images. The ratio of normal sample to edge detected sample is 9 : 4. That is to say,

$$\frac{\text{size of normal sample}}{\text{size of edge} - \text{based sample}} = \frac{9}{4}. \tag{19}$$

Considering (2), Table 4 has been produced to represent the comparison of size between CNN-based approach and edge detection-based CNN approach. The (1 + Rnd) has been

TABLE 4: Space consumption in case of CNN and proposed CNN with edge detection mechanism.

Number of samples	CNN approach	CNN with edge detection
10	$900 * (1 + \text{Rnd})$	$400 * (1 + \text{Rnd})$
20	$1800 * (1 + \text{Rnd})$	$800 * (1 + \text{Rnd})$
30	$2700 * (1 + \text{Rnd})$	$1200 * (1 + \text{Rnd})$
40	$3600 * (1 + \text{Rnd})$	$1600 * (1 + \text{Rnd})$
50	$4500 * (1 + \text{Rnd})$	$2000 * (1 + \text{Rnd})$
60	$5400 * (1 + \text{Rnd})$	$2400 * (1 + \text{Rnd})$
70	$6300 * (1 + \text{Rnd})$	$2800 * (1 + \text{Rnd})$
80	$7200 * (1 + \text{Rnd})$	$3200 * (1 + \text{Rnd})$
90	$8100 * (1 + \text{Rnd})$	$3600 * (1 + \text{Rnd})$
100	$9000 * (1 + \text{Rnd})$	$400 * (1 + \text{Rnd})$

added in order to tackle the variation in results. The value of Rnd is lying between zero and one and 1 + Rnd means that 1+ (value between 0 and 1). The table representing the space consumption in case of CNN and proposed CNN with edge detection is shown in the following and the simulation has been made in Figure 8.

Considering Table 4, Figure 8 represents a comparison of space consumed for CNN and CNN with edge detection.

4.7. Simulation of Accuracy. Here, the simulation of accuracy before edge detection and after edge detection has been found. The existing sample of InSitu has been taken to check the accuracy during the comparison process. The algorithm to simulate the accuracy is discussed as follows:

(1) In this step, the sample dataset with slight modification are read.

(2) The comparison of both matrixes is made using the image comparison module that would find the modification in the sample dataset.

(3) The result of mismatch is shown in the following if edge detection is not applied:

Ans = 95.4436.

(4) Now apply edge detection on the sample dataset.

(5) Perform the comparison of both datasets.

(6) The difference/mismatch has been shown as follows:

Ans = 83.4231.

(7) Get the difference between both datasets:

Ans = 12.

(8) Calculate the accuracy with the support of given equation ((old_matching% new_matching %) × 100)/new_matching % ((95 – 83) × 100)/83:

$$\frac{1200}{83} = 14.4578. \tag{20}$$

The sample of 10, 20, . . ., 90 has been considered for simulation. The CNN has been applied and its accuracy is found. The accuracy is multiplied by (1 + Rnd) where Rnd is a random value lying between 0 and 1. Then the accuracy that is 14.45% from equation (3) is added in case of CNN
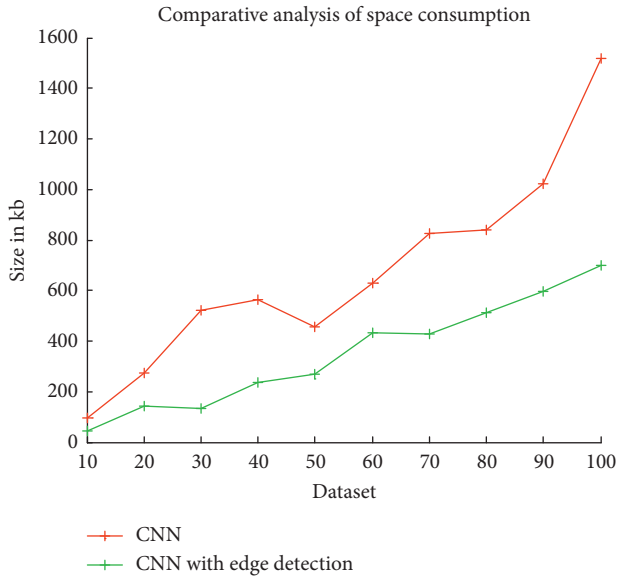
FIGURE 8: Simulation of space consumption in case of CNN and CNN with edge detection.

TABLE 5: Accuracy in case of CNN and proposed CNN with edge detection mechanism.

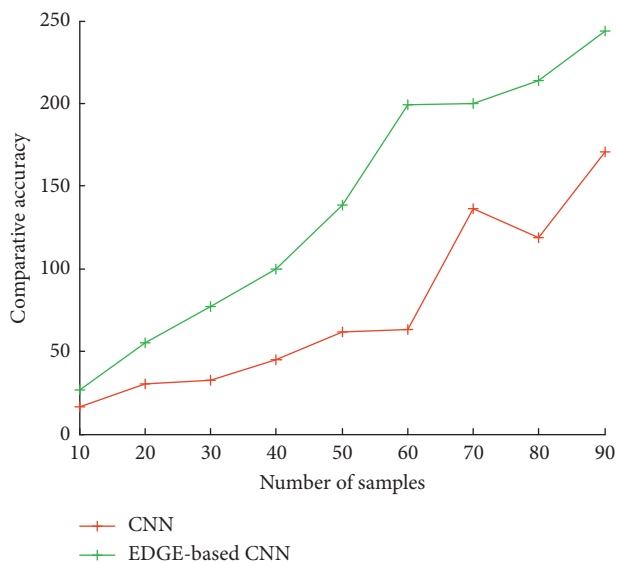| Number of samples	CNN approach	CNN with edge detection
10	$10 * (1 + Rnd)$	$11.4458 * (1 + Rnd)$
20	$20 * (1 + Rnd)$	$22.8916 * (1 + Rnd)$
30	$30 * (1 + Rnd)$	$34.3373 * (1 + Rnd)$
40	$40 * (1 + Rnd)$	$45.7831 * (1 + Rnd)$
50	$50 * (1 + Rnd)$	$57.2289 * (1 + Rnd)$
60	$60 * (1 + Rnd)$	$68.6747 * (1 + Rnd)$
70	$70 * (1 + Rnd)$	$80.1205 * (1 + Rnd)$
80	$80 * (1 + Rnd)$	$91.5662 * (1 + Rnd)$
90	$90 * (1 + Rnd)$	$103.012 * (1 + Rnd)$

FIGURE 9: Simulation of accuracy in case of CNN and CNN with edge detection.

with edge detection and multiplied by (1 + Rnd). The simulation results are shown in Table 5 and the plotting is made in Figure 9.

Table 5 represents the accuracy in case of CNN and proposed CNN with an edge-based mechanism.

The chart in Figure 9 is representing a comparison of accuracy for CNN and CNN with edge detection. The accuracy is influenced as the number of datasets increases.

5. Conclusions

During the research work, dataset of benign, in situ, and invasive has been taken from the patient. The data are then stored for breast cancer prediction. Using a convolution neural network mechanism, breast cancer prediction has been done. CNN was also used in previous research for cancer prediction. As the objective of the research is to reduce the time and space and to improve the accuracy of the proposed work compared to previous work, the size of the dataset is taken along with the comparison time and percentage of matching. Then dataset size, time of operation, and matching percentage are taken after performing experiments considering edge detection-based conventional neural network.

The research has concluded that utilization of canny edge detection is minimized. Moreover, the time consumption of the existing CNN model is also reduced. The storage consumption is also reduced with the support of a graphical sample. The utilization of proposed work in the field of medical science is capable of enhancing the capability of CNN during decisions making. It has been observed that the proposed work is fourteen to fifteen percent more accurate. It is taking 9/4 times less space and 1.0849004/0.178971 times less time compared to the general CNN model. Accuracy might vary as per size of the image and alteration performed in dataset of the image.

The research would provide the study of the existing research in the field of breast cancer prediction. It would study the present research objectives and their benefits. The research work would investigate the limitation of existing researches. CNN-based edge detection has been used in order to predict breast cancer. It would perform a simulation in order to present the output of the proposed work. The research would perform a comparative analysis of CNN with proposed work in order to represent how the proposed model is better than the previous. In future, the optimization mechanism could be used.

Data Availability

All datasets are included within the paper.

Conflicts of Interest

The authors declare no conflicts of interest.

References

[1] Y. J. Tan, K. S. Sim, and F. F. Ting, "Breast cancer detection using convolutional neural networks for mammogram

imaging system," in *Proceedings of the 2017 International Conference on Robotics, Automation and Sciences (ICORAS)*, Melaka, Malaysia, November 2017.

[2] F. A. Spanhol and L. S. Oliveira, "Breast cancer histopathological image classification using convolutional neural networks," in *Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN)*, Vancouver, Canada, July 2016.

[3] S. Charan, M. J. Khan, and K. Khurshid, "Breast cancer detection in mammograms using convolutional neural network,," pp. 0–4, 2017.

[4] A. Chandra and S. Rao, "Breast cancer classification of image using convolutional neural network," in *Proceedings of the 2018 4th International Conference on Recent Advances in Information Technology (RAIT)*, pp. 1–6, Dhanbad, India, March 2018.

[5] K. Liu, G. Kang, N. Zhang, and B. Hou, "Breast cancer classification based on fully-connected layer first convolutional neural networks," *in IEEE Access*, vol. 6, pp. 23722–23732, 2018.

[6] Y. Jiang, L. C. Id, H. Zhang, and X. Xiao, "Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module," *PLoS One*, vol. 14, no. 3, Article ID e0214587, 2019.

[7] D. A. Ragab, M. Sharkas, S. Marshall, and J. Ren, "Breast cancer detection using deep convolutional neural networks and support vector machines," *PeerJ*, vol. 7, p. e6201, 2019.

[8] S. Dencks, M. Piepenbrock, T. Opacic, B. Krauspe, E. Stickeler, and F. Kiessling, "Relative blood volume estimation from clinical super-resolution US imaging in breast cancer," in *Proceedings of the 2018 IEEE International Ultrasonics Symposium (IUS)*, no. 1, pp. 1–4, Kobe, Japan, October 2018.

[9] G. Czamota, H. Tadayyon, M. Gangeh et al., "Quantitative ultrasound and texture predictors of breast tumour response to chemotherapy," in *Proceedings of the 2018 IEEE International Ultrasonics Symposium (IUS)*, no. 1, pp. 1–4, Kobe, Japan, October 2018.

[10] M. Thomas, K. De Brabanter, J. A. K. Suykens, and B. De Moor, "Predicting breast cancer using an expression values weighted clinical classifier," *BMC Bioinformatics*, vol. 15, no. 1, p. 411, 2014.

[11] E. Kontopodis, M. Venianaki, G. Manikis et al., "Investigating the role of model-based and model- free imaging biomarkers as early predictors of neoadjuvant breast cancer therapy outcome," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 5, pp. 1834–1843, 2019.

[12] B. Dai, R. Chen, S. Zhu, and W. Zhang, "Using random forest algorithm for breast cancer diagnosis," in *Proceedings of the 2018 International Symposium on Computer, Consumer and Control (IS3C)*, Taichung, Taiwan, December 2018.

[13] M. Nawaz, A. A. Sewissy, and T. H. A. Soliman, "Multi-class breast cancer classification using deep learning convolutional neural network," *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 9, no. 6, pp. 316–322, 2018.

[14] T. Ahn, N. Kang, Y. Kim, and T. Park, "Gene expression based prediction of prognostic outcome in ovarian cancer," in *Proceedings of the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Madrid, Spain, December 2018.

[15] T. Anothaisintawee, Y. Teerawattananon, C. Wiratkapun, V. Kasamesup, and A. Thakkinstian, "Risk prediction models of breast cancer: a systematic review of model performances: a systematic review of model performances," *Breast Cancer Research and Treatment*, vol. 133, no. 1, pp. 1–10, 2012.

[16] J. Fan, D. Page, and E. Burnside, "Structure-leveraged methods in breast cancer risk prediction," *Journal of Machine Learning Research*, vol. 17, pp. 1–15, 2016.

[17] D. Mulatu, "Survey of data mining techniques for prediction of breast cancer recurrence," *International Journal of Computer Science and Information Technologies*, vol. 8, no. 6, pp. 599–601, 2017.

[18] S. K. M. Hamouda, H. R. Abo El-Ezz, and M. E. Wahed, "Intelligent system for predicting, diagnosis and treatment of breast cancer," *International Journal of Biomedical Data Mining*, vol. 6, no. 128, 2017.

[19] H. Lin, S. Member, H. Chen, S. Graham, and S. Member, "Fast ScanNet: Fast and dense analysis of multi-gigapixel whole-slide images for cancer metastasis detection," *IEEE Transactions on Medical Imaging*, vol. 38, no. 8, pp. 1948–1958, 2018.

[20] V. Chaurasia, S. Pal, and B. B. Tiwari, "Prediction of benign and malignant breast cancer using data mining techniques," *Journal of Algorithms & Computational Technology*, vol. 12, no. 2, pp. 119–126, 2018.

[21] P. Chauhan and A. Swami, "Breast cancer prediction using genetic algorithm based ensemble approach," in *Proceedinds of the 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, pp. 1–8, Bengaluru, India, July 2018.

[22] B. Fu, P. Liu, J. Lin, L. Deng, K. Hu, and H. Zheng, "Predicting invasive disease-free survival for early-stage breast cancer patients using follow-up clinical data," *IEEE Transactions on Biomedical Engineering*, vol. 9294, 2018.

[23] D. Kaushik, "Post-surgical survival forecasting of breast cancer patient: a novel approach," in *Proceedings of the 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 37–41, Bangalore, India, September 2018.

[24] M. Ma, Y. Shi, W. Li, Y. Gao, and J. Xu, "A novel two-stage deep method for mitosis detection in breast cancer histology images," in *Proceedings of the 2018 24th International Conference on Pattern Recognition (ICPR 2018)*, pp. 3892–3897, Beijing, China, August 2018.

[25] M. Chang, R. J. Dalpatadu, D. Phanord, A. K. Singh, W. F. Harrah, and H. Administration, "Breast cancer prediction using bayesian logistic regression," *Annals of Community Medicine and Practice*, vol. 2, pp. 2–6, 2018.

[26] A. Rampun, B. W. Scotney, P. J. Morrow, and H. Wang, "Breast mass classification in mammograms using ensemble convolutional neural networks," in *Proceedings of the 2018 IEEE 20th International Conference on e-Health Networking, Applications and Services (Healthcom)*, pp. 1–6, Ostrava, Czech Republic, September 2018.

[27] M. Sun, T. Ding, X.-Q. Tang, and Y. Keming, "An efficient mixed-model for screening differentially expressed genes of breast cancer based on LR-RF," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 16, no. 1, pp. 124–130, 2019.

[28] Y. Xiao, J. Wu, Z. Lin, and X. Zhao, "Breast cancer diagnosis using an unsupervised feature extraction algorithm based on deep learning," in *Proceedings of the 2018 37th Chinese Control Conference (CCC)*, pp. 9428–9433, Wuhan, China, July 2018.

[29] M. Gupta and B. Gupta, "An ensemble model for breast cancer prediction using sequential least squares programming method (SLSQP)," in *Proceedings of the 2018 Eleventh International Conference on Contemporary Computing (IC3)*, Noida, India, August 2018.