Journal of Trends in Computer Science and Smart Technology (ISSN: 2582-4104)
www.irojournals.com/tcsst/

Comparative Study of Artificial Intelligence Models for Breast Cancer Detection

Tanvi Meet Dhruv[1], Vipul K. Dabhi[2]

Department of Information Technology, Dharmsinh Desai University, Nadiad–387002, Gujarat,India

Email: [1]tanvigoswami.it@ddu.ac.in, [2]vipuldabhi.it@ddu.ac.in

Abstract

The most prevalent type of cancer among women is breast cancer. According to the statistics given by the World Health Organization (WHO), breast cancer is the reason behind the death of about 2.3 billion women globally in 2020, accounting for 685.9 million deaths. Since they are thought to be useful approaches, machine learning and deep learning techniques have drawn attention from researchers in breast cancer detection. Also, it can significantly assist in the process of prior detection and prediction of breast cancer by extracting handcrafted features. However, in recent years, improvements in artificial intelligence (AI) have enabled the successful use of deep learning strategies like CNN and the transfer learning method for detection of breast cancer. A significantly large dataset is used for deep learning methods. It does not require human intervention for feature extraction, which, as a result, enhances the patient's chances of survival. This review paper is based on breast cancer detection using deep learning and machine learning-based cancer detection techniques to aid in the understanding of trends and challenges in cancer detection.

Keywords: Convolutional Neural Networks (CNN), Machine Learning (ML), Breast Cancer, Deep Learning (DL).

1. Introduction

Cancer is a condition that happens when there’s a problem with the way cells replicate and respond in the body, leading to abnormal cell growth. It can be categorized as malignant (cancerous) or benign (non-cancerous). Benign tumors do not get larger or spread to other areas

DOI: https://doi.org/10.36548/jtcsst.2024.1.002
Received: 13.01.2024, received in revised form: 03.02.2024, accepted: 18.02.2024, published: 01.03.2024

of the body. Malignant tumors, on the other hand, are cancerous growths that have the ability to invade and harm other tissues as well as target various body parts. Breast cancer affects women of all ages, but it is most common in women after puberty, with an increasing incidence in later life. Breast cancer can be categorized into four main types: normal, benign, non-invasive, and invasive carcinoma. Benign breast tissue does not pose a risk to health due to its ability to cause minimal alterations to the breast's structural components. Yet, non-invasive breast cancer does not spread to other tissues; instead, it stays contained within the ductal lobule system of the breast. If discovered earlier, it is possible to treat non-invasive types of cancer. By contrast, an invasive breast cancer is a cancerous tumor that can spread to other parts of the body [14]. There are major 4 basic breast cancer detection technologies: MRI, mammogram, ultrasound and microwave imaging. Where each technique has some pros and limitations [1]. Studies have shown that, depending on the density of the breasts, radiologists may be able to miss up to 30% of cases of breast cancer. Micro-calcification and masses are two potent markers that have been used to evaluate the accuracy of mammograms for breast cancer. Detection of masses is more difficult than that of micro calcification, not only due to the wide range of sizes and shapes that can appear on a mammogram, but also due to the fact that masses often display poor image contrast [12]. The mammographic detection of breast cancer presents a difficult image analysis task because healthy breast tissue can mask cancer cells. Also, breast cancer image represent heterogeneous features and each and every features play major role in disease detection, specifically if it is malignant and life-threatening. There is possibility in mammographic images where the normal tissue can appear as cancerous or vice-versa, which may lead to confusion and also affect patient's mental health. Figure 1 shows benign mammogram image of the breast versus malignant cancer infected breast images.

Due to major contribution of machine learning, it is now possible to identify serious health problems and give good outcomes. But, with machine learning requires handcrafted feature identification and extraction. As, in some cases of disease there may be variety of features representing the same disease and can be often confusing if model is not trained with enough data samples. However, deep learning is often used in the healthcare industry to identify and diagnose early illnesses. Image interpretation has advanced significantly with deep learning, facilitating the identification, classification, and measurement of patterns in breast cancer images. The survey of ML and DL based techniques for detecting breast cancer are summarized in this study. The study takes into account research articles published between 2016 and 2024. The study aims to understand and describe the factors influencing the breast cancer, the general methods for detection of the breast cancer and limitation associated with each method along with the possible measures. It also provides a brief overview of the publicly accessible dataset on breast cancer and discussion on strengths and the weakness of machine learning and the deep learning models in breast cancer detection.

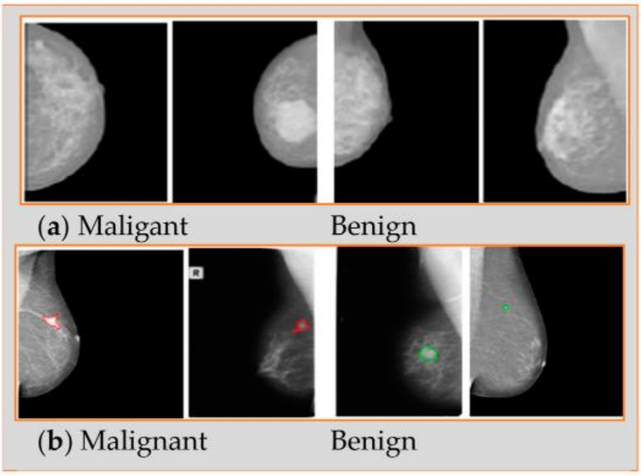

Figure 1. Benign versus Malignant Mammogram Samples [23].

2. Related Study

This section highlights the steps involved in breast cancer detection for ML and the DL based techniques.

2.1 Machine Learning Methods

Machine learning (ML) can be conceptualized as a branch of Artificial Intelligence which embeds the capacity to learn into a system based on a set of data used for the training process. The quality of classifiers depends heavily on feature extraction and selection. It is undoubtedly challenging to detect cancer using early-stage mammography. There are several complex features which may lead to misdiagnosis like: tumors appearing in low contrast areas, healthy tissues masking the cancer cells or a painless, hard, atypically shaped breast lump that differs from the surrounding breast tissue. The skin around the lump might get thicker, turn red, or change colour. There are a variety of approaches and techniques for creating systems that are capable of learning, which includes classification and clustering algorithms. The summary [38] of some frequently used machine learning techniques is given in Table 1.

There are 5 major steps of breast cancer detection using machine learning. These are: (1) Image acquisition (2) Image pre-processing (3) Segmentation (4) Feature Extraction and (5) Classification. A model is trained with labelled datasets in this supervised learning approach and based on its learning it is able to classify the breast cancer images as benign or malignant type of breast cancer.

Table 1. Summary of Machine Learning Algorithms

Algorithm	Advantages	Limitations/Assumptions
Decision Tree (DT)	Simple to comprehend, efficient, and unaffected by instance order	There must be mutual exclusion between classes. Missing values of the attribute DT depends on Order selection
Naive Bayes	Simple to comprehend, do not require more training data, handle both continuous and discrete data, highly scalable and fast.	Issue with zero-frequency problem Assume all attributes are independent which is not applicable to real-world problem
Neural Network	Able to handle noisy input, recognize intricate relationships between variables, and be helpful in regression and classification tasks.	Black-box nature Need more computational cost. Prone to over-fitting
Support Vector Machine	Can handle high-dimensional data, flexible and can give high quality results.	Do not perform well with large dataset Converge slowly with non-linear data. Sensitive to noise and outliers. cannot handle big data.

K-Nearest Neighbours	Learns slowly (lazy), No assumptions.	Outcome depends on choice of k and distance metric. High computational cost

2.2 Deep Learning Methods

Deep learning is an instance of machine learning that automatically deduces feature representations from unprocessed data by using learning representations. Deep learning does not necessitate the implementation of a human-engineered feature in order to achieve optimal performance. In recent years, a variety of deep learning techniques have been introduced such as Convolution Neural Network (CNN), Generative Adversarial Networks (GANs),Long Short Term Memory Networks (LSTMs), Recurrent Neural Networks (RNNs), Self-Organizing Maps (SOMs), Radial Basis Function Networks (RBFNs), Multilayer Perceptrons (MLPs), Deep Belief Networks (DBNs), auto encoders and many more[35]. There are major 7 types of layers in Deep Neural Networks named as: (1) Input Layer (2) Convolutional Layers (3) Activation (4) Pooling layer (5) Flattening (6) Fully Connected (FC) and (7) Output Layer

As CNNs are good at extracting features from images and learning to identify patterns, it is widely used in image processing for tasks like object detection, image segmentation, and classification. However, tuning of the hyperparameters in CNN plays major role in classification accuracy. These are the external configuration for example number of nodes, number of layers, learning rate, batch size, number of epochs and many more. The internal parameters of CNN can be learned during the learning process and they are set automatically. While, the hyper parameters that can be set externally which contribute majorly in the model performance and accuracy. The transfer learning method is highly used nowadays where pre-trained models are fine-tuned if needed and are used with new problem with less dataset. The training time of CNN can be reduced by using various pre-trained architectures like VGG16, VGG19, ResNet50, and InceptionV3. Some of these CNN architectures [36][37] are summarized in Table 2.

Table 2. CNN Architectures [37]

CNN Architec-ture	Input	Output	No. of Layers	No. of Parameters	Activation Function	Top 5 error Rate	Key Features	Developed By	Year
LeNet	32x32x1	10	5	60K	tanh	_	CNNs were successfully applied for the first time with five layers (convolutional and pooling layers alternate). identifying letters that have been machine-printed and handwritten.	Yann LeCun et. al.	1998
AlexNet	227x227x3	1000	8	60M	ReLU	15.30%	Wider and deeper than LeNet; dropout layers were added; GPUs were used for training. Large-scale image identification assignments	Alex Krizhevsky et. al.	2012
VGG16	224x224x3	1000	16	138M	ReLU	7.30%	Deep Network, small filter size (3×3) and The depth of each convolutional layer is the same. Later configuration: VGG19. Used for Large-scale image identification assignments	Simonyan and Zisserman	2014
Inception v1	224x224x3	1000	22	7M	ReLU	6.67%	The introduction of the Inception module, which enables deeper networks and more effective computation. First GoogleNet version, later versions are:Inception v2, v3 and v4 .Used for extensive image recognition assignments. Secured 1st place in ILSVRC-	Google	2014
ResNet50	224x224x3	1000	50	26M	ReLU	3.60%	Shortcuts or skipping of connections were introduced to allow for the training of deeper networks. ResNet-101 and ResNet-152 later versions, which took first place in the ILSVRC-2015	Microsoft Research	2015
MobileNet	224x224x3	1000	28	4.2M	ReLU	_	A depthwise convolution and a pointwise convolution make up each depthwise separable convolution layer. Later Versions are MobileNet v2(2018) and MobileNet v3(2019).	Google	2017

3. Literature Survey

This section reviews some of the prior work on the diagnosis of breast cancer conducted by researchers utilizing various ML and DL techniques and it is tabulated in Table 3 and Table 4.

3.1 Machine Learning based Cancer Detection in Breast

Sharma et al. [20] has done comparative analysis on KNN, Naive Bayes, and Random forest algorithms and determined time complexity of each algorithms along with the accuracy. It was observed that KNN provides high accuracy as compared to other two on WDBC dataset. Hadidi et al [6] transformed mammography images from time domain to frequency domain using DWT. This transformation results in the wavelet decomposition, which is made up of the diagonal detail coefficient matrix, the vertical detail coefficient matrix, the horizontal detail matrix, and the approximation coefficient matrix. At a later stage, Logistic regression models were used and compared with Back-propagation neural Network models. However, there was no clear discussion on the outcome. Nallamala et al [16] predicted using an ensemble voting ML technique with Logistic Regression and SVM classifier, the proposed strategy has achieved 98.50% accuracy. Muhammet et al [4] suggests a comparative study for the diagnosis and detection of breast cancer utilizing machine learning and data visualization approaches. Logistic regression (LR), k-nearest neighbour (kNN), support vector machine (SVM), naive bayes, decision tree (DT), Random forest and rotation forest are analyzed. Simulation results show that LR is the most accurate classifier as it has the highest classification accuracy of 98.1% on WBCD dataset. Das et. al.[34] proposed ML based Intelligent System – MLISBCP for breast cancer detection. The class imbalance problem is handled by ‘K-Means and ‘Boruta’ technique is used for selecting most prominent features from the dataset. The proposed system has achieved the accuracy of 97.53%.

3.1.1 Discussion on Machine learning Methods

The major strength of ML algorithms is that it can be trained with less computational cost with less number of training samples and therefore it needs less training time. However, it cannot scale well with large dataset and it is challenging to choose appropriate feature extraction technique for given features like color, texture and morphological features of the disease. Also, there is a need of hand-crafted feature extraction and it may lead to inaccurate result if appropriate features are not taken into consideration. Appropriate feature selection is the most important aspect to enhance the outcome of ML model. By pre-processing of images/data the unnecessary noise from the dataset can be eliminated and also reduces the number of input features through feature selection. The outcome can be improved by choosing appropriate set of parameters and its tuning.

Table 3. Review of Literature: Machine Learning based Approach

Ref	Dataset	No of Samples Considered	ML Algorithm used	Performance Measures	Limitations
[6]	-	209 mammogram m images	Backpropogation Neural Network (BPNN), Logistic Regression(LR)	Regression value >93.7%, Mean Square Error(MSE) <0.07	Number of features considered with LR were less due to limitations in MATLAB memory Very few samples and may not cover all the possibilities of features. No clear description of the dataset used.
[10]	WDBC	300 (csv)	Relevance vector machine (RVM)	Accuracy:97% Sensitivity: 98% Specificity: 98%	For the given dataset the numbers of features are reduced below 5. With no clear description of features considered. No feature selection mechanism.
[12]	DDMS	899	SVM, Decision Tree & Bayesian classifier	Decision Tree AUC: 100% Texture & EFD features: Bayes, Morphology & entropy: SVM- RB	No Single classifier for all the features like texture, morphology etc. For different features different classifier and kernels were used which produces good result.
[20]	WDBC	569 instances and 32 attribute & 2 classes (csv)	kNN (k-NearestNeighbor), Random Forest(RF), Naıve Bayes(NB)	Accuracy: kNN: 95.90% RF : 94.74%, NB: 94.47%	NB assumes all the features are independent due to which more malignant cases were misclassified as compare to benign cases.

3.2 Deep Learning based Cancer Detection in Breast

Nasser et.al.[17] has thoroughly reviewed the use of deep learning methods in breast cancer detection . Mahmud et.al [14] evaluated the capabilities of pre-trained, transfer learning models, including ResNet50 and ResNet101, as well as VGG16 and VGG19, to detect breast cancer using a dataset of 2453 histopathological images. The images in the dataset were divided into two groups: those with and those without invasive ductal cancer (IDC). It was observed and concluded that ResNet50 has a loss of about 3.5%. However, the VGG19 model was found to be the least robust of the four predefined transfer learning models. Gayathri et. al.[11] concluded that InceptionV3 has a reputation for capturing different levels of abstraction in medical images. While, ResNet50 was developed to address the vanishing gradient issue that can occur in deep neural networks. As a lightweight model optimized for mobile and embedded devices, MobileNet is an ideal choice for applications with limited resources. On the other hand, NASNet is a model that utilizes reinforcement learning to automatically identify neural network architectures that are superior to those designed by humans. Sahu et. al.[33] has integrated AlexNet, MobileNet2 and ResNet pre-trained models. The system's speed was enhanced by using residual learning, depthwise separable convolution, and inverted residual structure. Some of the connections were dropped, which made optimization easier. They obtained overall accuracy of 99.17% with 97.75% of malignancy detection on mini-DDSM dataset. 96.92% and 97.50% of accuracy is achieved on ultrasound dataset (BUSI) and BUS2 respectively. Luyang et. al. [32] has provided the in-depth view on various breast cancers imaging like mammogram, digital pathology, ultrasound and MRI along with the major profound learning strategies and applications on imaging-based screening, determination, treatment reaction expectation.

Table 4. Review of Literature: Deep Learning based Approach

Ref	Dataset Ref	No. of Samples considered for research work.	DL Algorithm used	performance Measures
[3]	Database for Mastology Research	266 thermal images.	CNN + Bayesian Network	Accuracy of CNN: 75.40%, ResNet50: 90.74%, CNN +BN:93%

[19]	MIAS	322 digitized films	Sparse Autoencoders(SAE), and Stacked Sparse Autoencoder (SSAE unsupervised), CNN (supervised).	Accuracy: SAE: 98.5%, SSAE: 98.9 %, CNN : 97%
[22]	MIAS	322 digitized films	MobileNet and Inception V3	Accuracy: MobileNet : 58%, InceptionV3: 83%
[13]	WDBC	569 instances and 11 at- tribute	Hand-crafted CNN with Number of input=12, No. of Neurons=12, epochs=20	Accuracy: 99.67%
[18]	DDSM and CBIS-DDSM	CBIS-DDSM: No. Of microcalcification case:753, No. Of mass cases: 891. DDSM: 2620 images.	Fine-tuned AlexNet with SVM Classifier	Accuracy=79%, AUC=0.88, Specificity=0.82, sensitivity=0.763, Precision=0.85
[21]	CBIS-DDSM	2478 mammography images	Resnet50 and VGG16	AUC=0.91, sensitivity: 86.1%, specificity:80.1%
[5]	Kaggle 162 H&E	162 scanned images of Breast Cancer. No. of patch extracted:, 277,524 patch size: 50 x 50	five-layer CNN in Model	Accuracy=87%
[2]	BreakHis	Microscopic images: 7909 of breast cancer tissue. No. of benign samples: 2480 and malignant samples:5429	Xception	F1- Score=97.58%, Precision=97.60%, Recall=97.6%.

[7]	Dataset from kaggle	1000	ML algorithms=SVM & RF, DL algorithm=CNN	Accuracy: CNN: 99.67%, SVM: 89.84% , RF:90.55%
[14]	BreakHis	2478 mammography images	ResNet50, ResNet101, VGG16, and VGG19	Highest performance achieved by ResNet50 Accuracy : 90.20%, AUC: 90%, recall: 94.7% and loss: 3.5 %
[11]	UM-BMID	200 scans of MRI derived breast phantoms	DenseNet201, ResNet50, InceptionV3, MobileNetV2, InceptionResNetV3, NASNetMobile and NASNetLarge	Highest performance achieved by NASNet Accuracy:88.41% , loss:27.82%, AUC:0.786 %
[9]	BreakHis	Total Samples: 1305 benign samples:407 malignant samples:898	"15 layers CNN architecture: convolution layers:3, max pooling layers:3, dense (hidden) layers:3, dropout layers:3, and flatten layer:1"	Accuracy=83.9%

3.2.1 Discussion on Deep Learning Methods

The DL can eliminate the problem of feature extraction of cancer imaging as it automatically extract the features by selecting appropriate filters. Also, it can be trained in a comprehensive manner and can be highly precise and potentially transferable across multiple types of mammography. These algorithms scale well with large amount of datasets. The transfer learning approach eliminates the requirement of excessive training time if it is appropriate for given dataset. Despite of many advantages of DL is high in computational cost, training time, and requires large dataset for training. Model overfitting is the major problem with deep learning. The computational resources are provided by various platforms where a free GPU is allocated for limited time period for example: Google Colab, Kaggle and Gradient. The details about these resources are given by [31]. Also, the number of training samples can be increased by data augmentation and hyperparameter tuning also plays major role for better performance and they are crucial for obtaining the optimal results. Also, transfer learning can be used to train a neural network and this approach has the benefit of reducing the training time for the learning model and can lead to a decrease in generalization errors. Designing of lightweight CNN architecture may take fewer resources which may leads to improvement in the performance.

4. Dataset Availability

There are very few and openly accessible datasets for breast cancers available and are summarized in Table 5. The majority of researchers who use the DDSM, for instance, do not utilize all of its images. The decompression code is necessary for non-standard DDSM image compressed which is no longer updated or maintained for use with modern computers. Lastly, while a general position of lesions was indicated by the ROI (region of interest) annotations for the irregularity in the DDSM, a precise segmentation for those lesions was not provided. For accurate feature extraction, many researchers must thus use segmentation algorithms. The creation and planned release of the Curated Breast Imaging Subset of DDSM abbreviated as CBIS-DDSM, an upgraded DDSM offers better ROI segmentation and readily accessible data [27]. One benchmark dataset for breast cancer that is more frequently used is the Wisconsin Breast Cancer Diagnostic (WDBC).This dataset was acquired by fine-needle aspiration of the breast mass, yielding 569 images and their corresponding cell nuclei characteristics and for every cell nucleus, ten real-valued features were calculated. Every image has been diagnosed and classified as either "Benign" or "Malignant". However, this dataset is limited by number of samples [28].

Table 5. Publicly Available Datasets on Breast Cancer

Dataset used	Number of Samples	Dataset Format	Classes/ samples count in each class	Dataset Description	Dataset URL
Wisconsin Diagnosis Breast Cancer dataset (WDBC)	569 records of patient data, 9 features	csv	2 / benign: 357 malignant :212	"The features are estimated from a digital image of the breast mass fine needle aspiration (FNA). They represent the cellular nuclei found in the image".	https://archive.ics.uci.edu/dataset/17/breast+cancer+wisconsin+diagnostic
Database for Screening Mammography (DDSM)	2,620	LJPG	4/ Malignant Cases:914 Benign : 870 Benign Without Call-back: 141 Normal: 695 Cases	"2620 scanned film collection breast cancer studies. The database includes normal, benign and malignant breast cancer cases with validated pathology data.[39]"	"http://www.eng.usf.edu/cvprg/mammography/database.html"
"Curated Breast Imaging Subset of Digital Database for Screening Mammography (CBIS-DDSM)"	1644	Diacom	2/calcification cases:753, and mass cases: 891	Updated and standardized version of DDSM. Enough number of samples for mammography analysis and decision[27]	https://www.kaggle.com/datasets/awsaf49/cbis-ddsm-breast-cancer-image-dataset
Mammographic Image Analysis	322	PGM	-	The Black and white scans are normalized to 1024x1024 pixels. Due to the size of the dataset, it was not suitable for	"https://www.repository.cam.ac.uk/handl

Society (MIAS)				training. However, it was used for experimental data analysis and as an additional test data set.[23]	e/1810/250394”
Breast Cancer Histopathological Database (BreakHis)	7909 microscopic images	PNG	2/ Benign: 2,480 and Malignant : 5,429	RGB image with resolution 700X460 pixels, 8-bit depth per channel, PNG format. Has been generated in association with Pathological Anatomy and Cytopathology Laboratory [24].	https://data.mendeley.com/datasets/jxwvdwhpc2/1
“University of Manitoba Breast Microwave Imaging Dataset (UM-BMID)”	249 (Gen 1) + 1008 (Gen 2) MRI derived breast phantoms scans	.txt, .mat and .pickle	2/ Healthy samples:50% and Tumor samples: 50%	Available in three file formats: raw data : raw.txt, for MATLAB/Octave: .mat file, for python: .pickle file[25]	https://bit.ly/UM-bmid.
“The Database for Mastology Research with Infrared Image (DMR-IR)”	3,534 thermal images	JPEG	2/ healthy and sick.	It is the most popular breast thermogram database to date because it is the only one that is made available publicly.[26]	https://figshare.com/articles/dataset/unet_database_rar/21225386

5. Discussion

The novel insight gained from the study is that the deep learning approach is transforming the healthcare to the next level. Therefore, with promising amount of dataset a model can be trained well which may leads to better results in disease detection. In future multimodal learning can be useful for improving accuracy of the breast cancer detection which may include various views of the breast cancer images, and by considering different modalities like patients clinical report, age, genetic information, therapy, and stages of therapy.

6. Conclusion

Artificial intelligence has the potential to revolutionize the comprehensive diagnosis of cancer. However, healthcare professionals may find it difficult to understand complex AI functions, and this "black box" aspect may make AI programs less acceptable from an ethical and legal standpoint. By examining various ML and DL based algorithm's potential in breast cancer detection, researchers can gain a better understanding of how the technology might improve patient's breast cancer disease outcomes and reduce the number of cancer deaths. The advances in deep learning have discovered major texture as well as morphological patterns in images, providing new insights into the analysis of medical images. Deep learning is considered to be the state-of-the-art technique which has given prominent results in a variety of medical applications, applying them to further refinement could be a significant advancement in the field of medical computing.

References

[1] Ahmad Ashraf Abdul Halim, Allan Melvin Andrew, Mohd Najib Mohd Yasin, Mohd Amiruddin Abd Rahman, Muzammil Jusoh, Vijayasarveswari Veeraperumal, Hasliza A Rahim, Usman Illahi, Muhammad Khalis Abdul Karim, and Edgar Scavino. Existing and emerging breast cancer detection technologies and its challenges: a review. Applied Sciences, 11(22):10753, 2021.

[2] Basem S Abunasser, Mohammed Rasheed J ALHiealy, Ihab S Zaqout, and Samy S Abu-Naser. Breast cancer detection and classification using deep learning xception algorithm. International Journal of Advanced Computer Science and Applications, 13(7), 2022

[3] Nurduman Aidossov, Vasilios Zarikas, Yong Zhao, Aigerim Mashekova, Eddie Yin Kwee Ng, Olzhas Mukhmetov, Yerken Mirasbekov, and Aldiyar Omirbayev. An integrated intelligent system for breast cancer detection at early stages using ir images and machine learning methods with explainability. SN Computer Science, 4(2):184, 2023.

[4] Muhammet Fatih Ak. A comparative analysis of breast cancer detection and diagnosis using data visualization and machine learning applications. In Healthcare, volume 8, page 111. MDPI, 2020.

[5] Saad Awadh Alanazi, MM Kamruzzaman, Md Nazirul Islam Sarker, Madallah Alruwaili, Yousef Alhwaiti, Nasser Alshammari, Muhammad Hameed Siddiqi, et al. Boosting breast cancer detection using convolutional neural network. Journal of Healthcare Engineering, 2021, (2021).

[6] Abdulsalam Alarabeyyat, Mohannad Alhanahnah, et al. Breast cancer detection using k-nearest neighbor machine learning algorithm. In 2016 9th International Conference on Developments in eSystems Engineering (DeSE), pages 35–39. IEEE, 2016.

[7] Viswanatha Reddy Allugunti. Breast cancer detection based on thermographic images using machine learning and deep learning algorithms. International Journal of Engineering in Computer Science, 4(1):49–56, 2022.

[8] Hakim El Massari, Noreddine Gherabi, Sajida Mhammedi, Zineb Sabouri, Hamza Ghandi, and Fatima Qanouni. Effectiveness of applying machine learning techniques and ontologies in breast cancer detection. Procedia Computer Science, 218:2392– 2400, 2023.

[9] Bhavin Gami, Khushi Chauhan, and Brijeshkumar Y Panchal. Breast cancer detection using deep learning. In Mobile Radio Communications and 5G Networks: Proceedings of Third MRCN 2022, pages 85–95. Springer, 2023.

[10] BM Gayathri and CP Sumathi. Comparative study of relevance vector machine with various machine learning techniques used for detecting breast cancer. In 2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), pages 1–5. IEEE, 2016.

[11] Gayathri Girish, Ponnathota Spandana, and Badrish Vasu. Breast cancer detection using deep learning. arXiv preprint arXiv:2304.10386, 2023.

[12] Lal Hussain, Wajid Aziz, Sharjil Saeed, Saima Rathore, and Muhammad Rafique. Automated breast cancer detection using machine learning techniques by extracting

different feature extracting strategies. In 2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE), pages 327–331. IEEE, 2018.

[13] Naresh Khuriwal and Nidhi Mishra. Breast cancer diagnosis using deep learning algorithm. In 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), pages 98–103. IEEE, 2018.

[14] Md Ishtyaq Mahmud, Muntasir Mamun, and Ahmed Abdelgawad. A deep analysis of transfer learning based breast cancer detection using histopathology images. In 2023 10th International Conference on Signal Processing and Integrated Networks (SPIN), pages 198–204. IEEE, 2023.

[15] Mohammed Amine Naji, Sanaa El Filali, Kawtar Aarika, EL Habib Benlahmar, Rachida Ait Abdelouhahid, and Olivier Debauche. Machine learning algorithms for breast cancer prediction and diagnosis. Procedia Computer Science, 191:487–492, 2021.

[16] Sri Hari Nallamala, Pragnyaban Mishra, and Suvarna Vani Koneru. Breast cancer detection using machine learning way. Int J Recent Technol Eng, 8(2-3):1402–1405, 2019.

[17] Maged Nasser and Umi Kalsom Yusof. Deep learning based methods for breast cancer diagnosis: A systematic review and future direction. Diagnostics, 13(1):161, 2023.

[18] Dina A Ragab, Maha Sharkas, Stephen Marshall, and Jinchang Ren. Breast cancer detection using deep convolutional neural networks and support vector machines. PeerJ, 7:e6201, 2019.

[19] D Selvathi and A Aarthy Poornila. Deep learning techniques for breast cancer detection using medical image analysis. Biologically rationalized computing techniques for image processing applications, pages 159–186, 2018.

[20] Shubham Sharma, Archit Aggarwal, and Tanupriya Choudhury. Breast cancer detection using machine learning algorithms. In 2018 International conference on

computational techniques, electronics and mechanical systems (CTEMS), pages 114–118. IEEE, 2018.

[21] Li Shen, Laurie R Margolies, Joseph H Rothstein, Eugene Fluder, Russell McBride, and Weiva Sieh. Deep learning to improve breast cancer detection on screening mammography. Scientific reports, 9(1):12495, 2019.

[22] K Shwetha, SS Sindhu, M Spoorthi, and D Chaithra. Breast cancer detection using deep learning. International Journal of Engineering Research and Technology (IJERT), 2018.

[23] Zebari DA, Ibrahim DA, Zeebaree DQ, Mohammed MA, Haron H, Zebari NA, Damaševičius R, Maskeliūnas R. Breast cancer detection using mammogram images with improved multi-fractal dimension approach and feature fusion. Applied Sciences. 2021 Dec 20;11(24):12122.

[24] https://data.mendeley.com/datasets/jxwvdwhpc2/1

[25] Reimer T, Krenkevich J, Pistorius S. An open-access experimental dataset for breast microwave imaging. In2020 14th European Conference on Antennas and Propagation (EuCAP) 2020 Mar 15 (pp. 1-5). IEEE.

[26] Pérez-Martín, Jorge, and Raquel Sánchez-Cauce. "Quality analysis of a breast thermal images database." *Health Informatics Journal* 29, no. 1 (2023): 14604582231153779.

[27] . Lee, Rebecca Sawyer, Francisco Gimenez, Assaf Hoogi, Kanae Kawai Miyake, Mia Gorovoy, and Daniel L. Rubin. "A curated mammography data set for use in computer-aided detection and diagnosis research." *Scientific data* 4, no. 1 (2017): 1-9.

[28] Strelcenia, Emilija, and Simant Prakoonwit. "Effective feature engineering and classification of breast cancer diagnosis: a comparative study." *BioMedInformatics* 3, no. 3 (2023): 616-631.

[29] Alshuhri M, Al-Musawi SG, Al-Alwany AA, Uinarni, Rasulova I, Rodrigues P, Alkhafaji AT, Alshanberi AM, Alawadi AH, Abbas AH. Artificial intelligence in cancer diagnosis: Opportunities and challenges. Pathology-Research and Practice. 2023 Nov 29:154996

[30] Verde F, Romeo V, Stanzione A, Maurea S. Current trends of artificial intelligence in cancer imaging. Artificial Intelligence in Medical Imaging. 2020 Sep 28;1(3):87-93.

[31] https://towardsdatascience.com/free-gpus-for-training-your-deep-learning-models-c1ce47863350

[32] Luo L, Wang X, Lin Y, Ma X, Tan A, Chan R, Vardhanabhuti V, Chu WC, Cheng KT, Chen H. Deep learning in breast cancer imaging: A decade of progress and future directions. IEEE Reviews in Biomedical Engineering. 2024 Jan 24.

[33] Sahu A, Das PK, Meher S. An efficient deep learning scheme to detect breast cancer using mammogram and ultrasound breast images. Biomedical Signal Processing and Control. 2024 Jan 1;87:105377.

[34] Das AK, Biswas SK, Mandal A, Bhattacharya A, Sanyal S. Machine Learning based Intelligent System for Breast Cancer Prediction (MLISBCP). Expert Systems with Applications. 2024 May 15;242:122673.

[35] https://developer.ibm.com/articles/cc-machine-learning-deep-learning-architectures/

[36] https://medium.com/analytics-vidhya/cnns-architectures-lenet-alexnet-vgg-googlenet-resnet-and-more-666091488df5

[37] https://vitalflux.com/different-types-of-cnn-architectures-explained-examples/

[38] Suryachandra P, Reddy PV. Comparison of machine learning algorithms for breast cancer. In2016 International Conference on Inventive Computation Technologies (ICICT) 2016 Aug 26 (Vol. 3, pp. 1-6). Ieee.

[39] Logan J, Kennedy PJ, Catchpoole D. A review of the machine learning datasets in mammography, their adherence to the FaIR principles and the outlook for the future. Scientific Data. 2023 Sep 8;10(1):595.