scientific reports

OPEN

Enhanced breast cancer diagnosis through integration of computer vision with fusion based joint transfer learning using multi modality medical images

S. Iniyan[1], M. Senthil Raja[1], R. Poonguzhali[2], A. Vikram[3], Janjhyam Venkata Naga Ramesh[4,5], Sachi Nandan Mohanty[6] & Khasim Vali Dudekula[6 ✉]

Breast cancer (BC) is a type of cancer which progresses and spreads from breast tissues and gradually exceeds the entire body; this kind of cancer originates in both sexes. Prompt recognition of this disorder is most significant in this phase, and it is measured by providing patients with the essential treatment so their efficient lifetime can be protected. Scientists and researchers in numerous studies have initiated techniques to identify tumours in early phases. Still, misperception in classifying skeptical lesions can be due to poor image excellence and dissimilar breast density. BC is a primary health concern, requiring constant initial detection and improvement in analysis. BC analysis has made major progress recently with combining multi-modal image modalities. These studies deliver an overview of the segmentation, classification, or grading of numerous cancer types, including BC, by employing conventional machine learning (ML) models over hand-engineered features. Therefore, this study uses multi-modality medical imaging to propose a Computer Vision with Fusion Joint Transfer Learning for Breast Cancer Diagnosis (CVFBJTL-BCD) technique. The presented CVFBJTL-BCD technique utilizes feature fusion and DL models to effectively detect and identify BC diagnoses. The CVFBJTL-BCD technique primarily employs the Gabor filtering (GF) technique for noise removal. Next, the CVFBJTL-BCD technique uses a fusion-based joint transfer learning (TL) process comprising three models, namely DenseNet201, InceptionV3, and MobileNetV2. The stacked autoencoders (SAE) model is implemented to classify BC diagnosis. Finally, the horse herd optimization algorithm (HHOA) model is utilized to select parameters involved in the SAE method optimally. To demonstrate the improved results of the CVFBJTL-BCD methodology, a comprehensive series of experimentations are performed on two benchmark datasets. The comparative analysis of the CVFBJTL-BCD technique portrayed a superior accuracy value of 98.18% and 99.15% over existing methods under Histopathological and Ultrasound datasets.

The WHO described that cancer is the crucial reason for non-accidental deaths globally; around 8.8 million people died from cancer in 2015[1]. BC is a fatal and common syndrome among women globally. Hence, initial and accurate diagnosis plays an essential part in increasing the prediction and enhancing the patient's survival rate with breast tumours from 30 to 50%. Generally, BC has two categories: malignant and benign. Malignant is cancerous (invasive), whereas Benign is a noncancerous (non-invasive) cancer type[2]. Both tumours have

[1]Department of Computing Technologies, School of Computing, SRM Institute of Science and Technology, Kaatankulathur, Chennai 603203, India. [2]Department of Computer Science and Engineering, Periyar Maniammai Institute of Science & Technology, Vallam, Thanjavur 613403, India. [3]Department of Computer Science and Engineering, Aditya University, Surampalem 533437, Andhra Pradesh, India. [4]Department of Computer Science and Engineering, Graphic Era Hill University, Dehradun, India. [5]Department of Computer Science and Engineering, Graphic Era University, Dehradun, Uttarakhand, India. [6]School of Computer Science and Engineering (SCOPE), VIT-AP University, Amravati, Andhra Pradesh, India. ✉email: khasim.vali@gmail.com

additional subcategories that want to be diagnosed individually because each might lead to various prediction and treatment strategies. Medical imaging conditions are generally more effective and adopted for BC recognition than other testing techniques. Standard medical imaging conditions for BC analysis are ultrasound imaging, magnetic resonance imaging (MRI), mammography, computed tomography (CT), and histopathology images[3]. Histopathology is the procedure of total estimation and microscopic review of biopsy tissue specimens performed by a diagnostician or specialist to learn about cancer growing in organs or tissues[4]. Regular histopathologic samples have other cells and structures surrounded and dispersed haphazardly by different tissue types. The manual examination of historical images and the vision inspection take time. The utilities of computer-based image investigation denote an efficient model to increase histopathology images' analytical and predictive skills[5]. Moreover, Ultrasound imaging can be helpful in the prognosis and identification of BC because of its real-time imaging, non-invasive, high image resolution, and non-radioactive. Nevertheless, the analysis of ultrasound images requires highly trained and proficient radiologists. Generally, the computer-aided diagnosis (CAD) method has been used in all imaging conditions and analysis types. CAD would be beneficial for radiologists in the classification and detection of BC[6].

They have been improved to reduce expenses and enhance the radiologist's capacity for medical imaging analysis. The beginning of digital images in the medical field has given an edge to artificial intelligence (AI) for pattern recognition using a CAD process. CAD methods are intended to help doctors by automatically analyzing images. Therefore, such systems lessen human dependence, improve the diagnosis rate, and ease the complete treatment expenditures[7]. ML is abundantly required in object recognition, text classification, and image recognition. With the development of CAD technologies, ML has been efficiently applied to diagnose BC. Classification of histopathologic images connected to traditional ML methods and simulated extraction of features requires a physical features model; still, it doesn't require a tool through more efficacy, and it has advantages in computing time[8]. Still, the classification of histopathologic images associated with deep learning (DL), mainly convolutional neural networks (CNN), often requests more labelled training methods. It aids in classifying and detecting BC in an initial phase. The requirement for improved BC diagnosis stems from its high incidence and crucial impact on women's health worldwide. Early and precise detection of breast tumours is significant for enhancing treatment outcomes and survival rates[9]. Conventional diagnostic models often need to improve distinguishing between diverse tumours, accentuating the need for advanced methods. Leveraging computer vision (CV) and multi-modality medical images presents an innovative model to address these threats. Incorporating diverse imaging modalities can provide a more comprehensive understanding of tumour characteristics. This multifaceted strategy improves diagnostic accuracy and eases personalized treatment plans, ultimately aiming to enhance patient prognosis[10].

This study uses multi-modality medical imaging to propose a Computer Vision with Fusion Joint Transfer Learning for Breast Cancer Diagnosis (CVFBJTL-BCD) technique. The presented CVFBJTL-BCD technique utilizes feature fusion and DL models to effectively detect and identify BC diagnoses. The CVFBJTL-BCD technique primarily employs the Gabor filtering (GF) technique for noise removal. Next, the CVFBJTL-BCD technique uses a fusion-based joint transfer learning (TL) process comprising three models, namely DenseNet201, InceptionV3, and MobileNetV2. The stacked autoencoders (SAE) model is implemented to classify BC diagnosis. Finally, the horse herd optimization algorithm (HHOA) model is utilized to select parameters involved in the SAE method optimally. To demonstrate the improved results of the CVFBJTL-BCD methodology, a wide series of experimentations are performed on two benchmark datasets. The key contribution of the CVFBJTL-BCD methodology is listed below.

- The CVFBJTL-BCD model utilizes the GF approach for efficient noise reduction during the data preprocessing. This method also improves the quality of the input data, resulting in enhanced performance in subsequent analysis. Addressing noise effectively confirms more reliable results in BC classification.
- The CVFBJTL-BCD methodology utilizes a joint TL model that integrates three advanced models, namely DenseNet201, InceptionV3, and MobileNetV2. This incorporation also implements the merits of every technique to improve feature extraction and classification accuracy. As a result, it substantially enhances the overall efficiency of BC diagnosis.
- The CVFBJTL-BCD model employs the SAE technique to classify BC diagnoses precisely. This methodology effectually learns intrinsic patterns in the data, enabling greater accuracy in distinguishing between diverse cancer types. Optimizing the classification process contributes to more reliable diagnostic outcomes.
- The CVFBJTL-BCD technique optimizes the HHOA model to choose parameters for the SAE method. This optimization improves the approach's performance by fine-tuning crucial parameters and improving classification output. Effectively enhancing parameter selection confirms more precise and reliable BC diagnoses.
- The CVFBJTL-BCD method incorporates advanced noise removal, diverse TL models, and innovative optimization methods into a cohesive framework for enhanced BC classification. This overall methodology improves the model's data quality and performance and allows for greater adaptability across varying datasets. The novelty is its seamless integration of these elements, substantially enhancing diagnostic accuracy and reliability in clinical settings.

The article is structured as follows: Section "2" presents the literature review, Sect. "3" outlines the proposed method, Sect. "4" details the results evaluation, and Sect. "5" concludes the study.

Related works

Sushanki et al[11]. present an overview of the recent techniques and developments in multi-modal imaging for BC detection. While radiomics, a quantifiable research of imaging data, has been combined using ML and DL methods, breast lesions are displayed as much as possible. These methods could assist in differentiating between

malignant and benign tumours, offering doctors critical information. During the several stages of BC recognition, novel techniques are advanced for improvement, classification, segmentation, and feature extraction, utilizing numerous image modalities. Abdullakutty et al[12]. examined the developing area of multi-modal methods, especially combining HI with non-image data. This article uses multi-modal data and highlights explainability to improve the precision of diagnosis, patient engagement, and clinician confidence, eventually raising more customized treatment tactics for BC. Oyelade et al[13]. propose a new DL method incorporating TwinCNN methods for addressing the challenges of BC image identification from multi-modalities. Initially, modality-based feature learning was attained by extracting either high- or low-level features utilizing the networks embedded by TwinCNN. In addition, to tackle the infamous issue of highly dimensional features related to the extracted features, the binary optimizer technique has been improved to remove nondiscriminant features efficiently in the search space. Hu et al[14]. present new methods for the survival prediction of BC, called GMBS. GMBS initially describes a sequence of multi-modal fusing components to incorporate various modalities of patient data, compromising robust early embeddings. Afterwards, GMBS proposes a patient-patient graph creation component, intending to describe the relationship of interpatients efficiently. Finally, GMBS combines a GCN method to utilize the complex architectural data encrypted in the created graph.

Yang et al[15]. present a new multiple-attention interaction network for BC identification based on the images of DWI and ADC. A three-intermodality interaction mechanism has been demonstrated to ultimately unite the multi-modality data. Three model connections are performed over the advanced intermodality relative component, multilevel attention fusion component, and channel interaction module to study discriminative, correlation, and complementary information. Guo et al[16]. present a multi-modal BC diagnosis technique based on Knowledge Augmented DL called KAMnet. This technique combines the 3 forms of previous knowledge into DNN over various incorporation tactics. Initially, a temporal segment-selecting approach led by Gaussian sampling over data level incorporation was developed to direct the method to concentrate on keyframes. Next, the author creates a feature fusion network for structure-level incorporation and attains collective interpretation by decision-level incorporation, enabling the exchange of multi-modal data. Zhang et al[17]. present the MoSID method that extracts inter or intra-modality attention maps as previous knowledge to conduct cancer segmentation from multi-parametric MRI. This can be attained by separating modality-specific data, which offers corresponding signs to the segmentation task and creating modality-specific attention mapping in a creation method. Yan et al[18]. present the TDF-Net for identifying BC utilizing partial multi-modal ultrasound data. Firstly, this technique proposes a double branch feature extraction component to seizure modality-specific information. Karthik et al[19]. introduce a tumour lesion segmentation methodology using advanced CNNs on combined MRI and PET scans. It also employs four pre-trained models, namely VGG 19, ResNet 50, SqueezeNet, and DenseNet 121—fine-tuned on ISL datasets, and utilizes a weighted voting ensemble for improved prediction accuracy.

Atrey, Singh, and Bodhey[20] propose a hybrid framework that integrates a residual neural network (ResNet) with ML for classifying BC images from mammography and ultrasound. Features are extracted, fused, and classified as benign or malignant using a support vector machine (SVM). Subaar et al[21]. utilize TL from a Residual Network 18 (ResNet18) and Residual Network 34 (ResNet34) architectures to detect BC. Chi et al[22]. introduce a dual-branch network with a cross-modality attention mechanism for thyroid nodule diagnosis, incorporating CEUS videos and ultrasound images. It employs two transformers—UAC-T for segmenting CEUS features and CAU-T for guiding classification—optimized through multi-task learning. Li et al[23]. present nn-TransEC, a 3D TL framework based on nnU-Net that integrates segmentation and classification for endometrial cancer T-stage classification. It features nn-MTNet for multi-task learning and a knowledge-embedded ROI tokenization method (KRT) replicating doctors' workflows by cropping key regions from CT volumes based on prior segmentation. Rahman, Khan, and Babu[24] introduce the "BreastMultiNet" framework by utilizing the TL model to detect diverse BC types using two public datasets. Hasan et al[25]. present a computer-assisted methodology for classifying ultrasound lung images using a fuzzy pooling-based CNN (FP-CNN) model. The Shapley Additive Explanation (SHAP) method is employed to clarify the FP-CNN predictions of the model and illustrate the decision-making process. Hossain et al[26]. introduce a fast, automatic COVID-19 detection technique using chest CT scans. After preprocessing, the model also fuses features from VGG-19 and ResNet-50, identifies key features with Principal Component Analysis (PCA), and classifies them using Max Voting Ensemble Classification (MVEC) methods.

Hasan et al[27]. developed a model integrating hand-crafted and automated feature extraction. The study also utilized multi-criteria frame selection. Hossain et al[28]. present a variational quantum circuit (VQC)-based technique using an optimized feature set. Features are optimized with minimum redundancy, maximum relevance (mRMR), and PCA methods. Furthermore, a rule-based expert system is presented. Mostafiz et al[29]. integrate curvelet transform and CNN features extracted from the preprocessed MRI sequence of the brain. Relevant features are chosen from the feature vector using mutual data based on the minimum redundancy maximum relevance (mRMR) methodology. The detection is done by employing the ensemble classifier of the bagging method. Ai et al[30]. propose the Self-and-Mutual Attention (SAMA) model. This technique features a module with three stages: improving modality-specific features, bidirectional cross-guidance for refinement, and optimized feature aggregation for precise predictions. Nakach, Zerouaoui, and Idri[31] employs multi-scale ensembles. Three pre-trained models, DenseNet 201, MobileNet v2, and Inception v3, are utilized across four magnification factors. Predictions are fused with weighted voting and seven meta-classifiers, with top models selected by using the Scott–Knott test and ranked by Borda count voting. Mukhlif, Al-Khateeb, and Mohammed[32] utilize advanced TL models. The modified Xception model is used to classify histological images. Mukhlif, Al-Khateeb, and Mohammed[33] present Dual TL (DTL), utilizing pattern convergence between source and target domains. It utilizes this method for VGG16, Xception, ResNet50, and MobileNetV2, fine-tuning

unclassified and limited classified images and using data augmentation for enhanced sample size and class balance. Ali and Mohammed[34] aim to review and analyze the efficiency of AI methods in omics data analysis.

The existing studies exhibit limitations in methodologies for BC detection. A reliance on imaging data may restrict applicability when datasets are insufficient, while multi-modal methods can complicate integration and reduce effectiveness. Methodologies like TwinCNN may need help with high-dimensional data, resulting in overfitting. Some models need massive datasets for effectual training, and intrinsic interaction mechanisms can increase computational demands. Concerns also arise regarding the robustness of models that rely on prior knowledge integration, as segmentation methods may falter with diverse cancer types. Moreover, performance can vary based on the choice of pre-trained models, and reliance on predefined ROI tokenization might hinder feature capture. Variability in ultrasound imaging conditions additionally complicates classification. These threats underscore the requirement for more adaptable and efficient diagnostic frameworks in clinical settings. Despite enhancements in BC detection models, crucial research gaps remain, specifically in addressing the incorporation of multi-modal data and optimizing model robustness across diverse datasets. Many existing models need help with high-dimensional data and overfitting, while the reliance on predefined features often limits their adaptability. Furthermore, the variability in imaging conditions and the requirement for massive datasets hinder the development of universally applicable outcomes. Thus, there is a significant requirement for more efficient, flexible, and interpretable frameworks in clinical practice.

Proposed methodology

This research proposes a novel CVFBJTL-BCD technique using multi-modality medical imaging. The presented CVFBJTL-BCD technique uses feature fusion and DL models to effectively detect and identify BC diagnosis. It comprises four stages: image preprocessing, feature fusion, classification, and parameter tuning. Figure 1 illustrates the entire flow of the CVFBJTL-BCD approach.

Noise reduction

The presented CVFBJTL-BCD technique primarily applies GF for noise removal[35]. This model is chosen for its effectiveness in conserving significant spatial frequency data while reducing unwanted noise. GFs are appropriate for image processing tasks, specifically medical imaging because they can efficiently detect textures and patterns at diverse scales and orientations. This adaptability allows for an enhanced feature extraction, which is significant for precise diagnosis. Compared to other noise removal approaches, such as Gaussian filtering, the GF model provides superior edge preservation, confirming that critical features are not blurred. Moreover, its mathematical foundation enables precise control over the filter parameters, making it versatile for diverse imaging modalities. Overall, utilizing GF improves the quality of the input data, resulting in enhanced outcomes in subsequent analysis and classification tasks. Figure 2 illustrates the structure of the GF model.
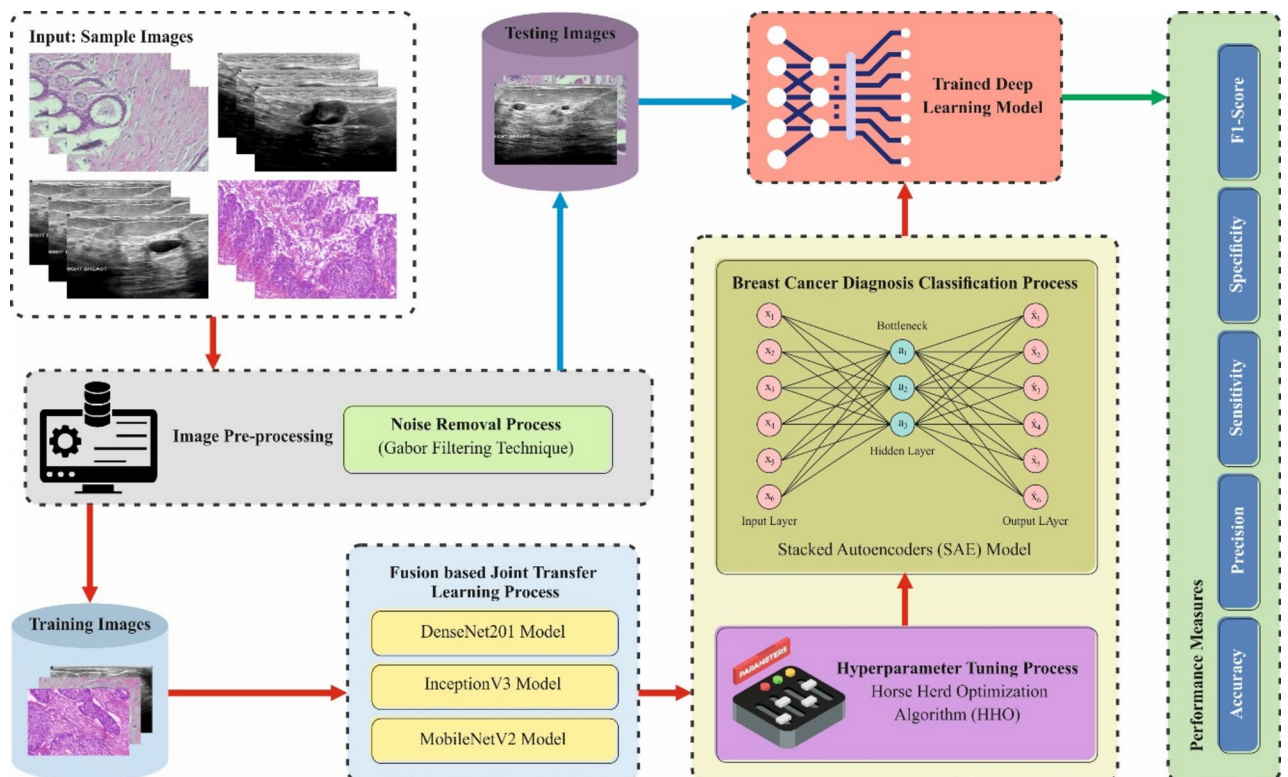

Fig. 1. Overall flow of CVFBJTL-BCD approach.
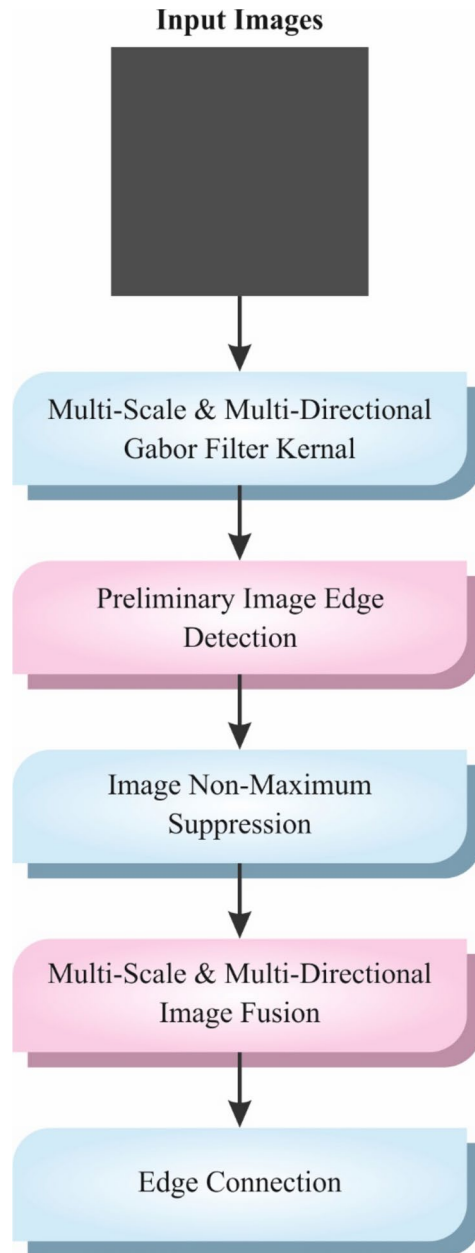
Input Images

Fig. 2. Workflow of GF technique.

The texture is crucial for differentiating the ROIs of dissimilar image classes. Texture study is essential for the analysis of computerized distribution. In ROIs, cancers contain micro patterns in numerous directions and frequencies. These forms are significant in classifying critical areas in a CAD method. GFs are effectively employed to recognize these patterns. GFs are linear filters used in numerous works in CV issues like face identification, texture analysis, and cancer diagnosis. An inspiring feature of GFs is that they have the finest united localization in spatial and frequency fields. To get the mammogram value, GFs of distinguished routes are employed to transfer mammograms by defining the finest extent of GF parameters, and the standardized mammogram is an output. These filters are denoted in mathematical formulations as given below:

$$g = \exp\left[-\frac{(x\cos\theta + y\sin\theta)^2 + \gamma^2(y\cos\theta - x\sin\theta)^2}{2\sigma^2}\right].$$

$$\exp\left[i\left[\frac{2_{Tr}(x\cos\theta + y\sin\theta)}{\lambda} + \varnothing\right]\right] \qquad (1)$$

whereas θ denotes filter orientation, γ means the ratio of spatial feature, λ represents sinusoidal wavelength, σ refers to standard deviation, and ø refers to the off-set of phase.

Feature fusion process

Next, the CVFBJTL-BCD method utilizes a fusion-based joint TL process comprising three models, namely DenseNet201, InceptionV3, and MobileNetV2. These models are chosen due to their merits in complementary feature extraction and classification. The DenseNet201 model outperforms at capturing complex patterns through its dense connectivity, which fosters better gradient flow and feature reuse. Also, the InceptionV3 method presets a multi-scale approach, effectually detecting features at diverse resolutions, making it robust against discrepancies in input data. MobileNetV2 is designed for effectiveness, giving high accuracy with mitigated computational costs, making it ideal for resource-constrained environments. By incorporating these models, the methodology employs diverse architectures to improve performance, increase generalization across diverse datasets, and improve overall predictive accuracy in complex classification tasks. This ensemble method also reduces the weaknesses of individual models, resulting in a more robust and efficient outcome.

While global features cover comprehensive data, the present redundancy frequently declines attention to the primary information[36]. To evade this concern, some approaches trust local features to remove terminated data intrusion by segmentation but might disregard the link. Removing local and global features analogously and openly combining them is a usual feature extraction tactic. On the other hand, this approach generally trades the significance of local and global feature extraction. To improve the correlation amongst dissimilar features, an effective technique uses a weighted sharing fully connected (FC) layer in a multi-task joint learning system. In the meantime, a multi-loss device is designed to assist network training in stabilizing the significance of local and global features. For local features, the segmentation can remove the terminated data in the feature for improvement. So, the joint learning network employs a patch layer for segmenting an input image (signified as $I \in R^{n \times n \times 3}$) into numerous patch-level imageries through the form of $\frac{n}{7} \times \frac{n}{7} \times 3$ and loads them as per the size of the channel. The attained patch-level tensor $\widehat{I}$ is denoted as:

$$\widehat{I} = [i_0, i_1, \ldots, i_k], \tag{2}$$

Meanwhile, i_k signifies the kth patch-level image.

Next, a technique is used to pick the convolution (Conv) layer with the kernel dimension $3x3$ in the local branch to remove local features. The $PRelu$ task is used to familiarize non-linearity. The $PRelu$ function presents few trainable parameters and has improved the drive of network weights and increased detection accuracy.

To decrease the intricacy, the global average pooling layer is utilized to compact the local branch output (signified as $1 \in R^{\frac{n}{112} \times \frac{n}{112} \times 512}$) and acquire the compacted feature $L \in R^{1 \times 512}$:

$$L = GAP(\bullet), \tag{3}$$

Here, $GAP(\bullet)$ indicates the global average pooling layer. Next, the local feature $\left(F_1 \in R^{1 \times 1024} \right)$ is mined as:

$$F_1 = FC(L), \tag{4}$$

Here, FC specifies the FC layer.

Likewise, by following the FC and global average pooling layers, the global feature $F_g \in R^{1 \times 1024}$ is attained as:

$$F_g = FC(G), \tag{5}$$

where

$$G = GAP(g), G \in R^{1 \times 512}, \tag{6}$$

$FC(\bullet)$ denotes the FC layer, and $GAP(\bullet)$ signifies the global average pooling layer.

A straight mixture of local and global features can cover both. The method restrains the training of multi-task joint learning networks depending upon the backward loss of straight fusion feature to ensure that the critical local and global feature data were well-kept after the synthesis. L and G are attained by compression through the global average pooling layer to achieve local and global features. When equated with F_g and F_1, G and L are not managed by FC layer training weights, so they could more efficiently mine local and global data. Now, G and L are combined directly with a similar ratio as $\left(F_c \in R^{1 \times 1024} \right)$:

$$F_c = Concat(L, G), \tag{7}$$

Meanwhile, $Concat(\bullet)$ denotes the concatenate operation.

DenseNet201

The design of the DenseNet201 structure uses densely connected blocks, permitting every layer to obtain straight input from each prior layer[37]. This helps feature reprocess and recover the flow of gradient. The structure includes numerous dense blocks, each containing densely associated convolutional layers. Transition layers were employed to allow downsampling over pooling and dimensional reduction. Non-linearity and constant training are endorsed by using $ReLU$ and BN after every convolutional layer. The network is determined with
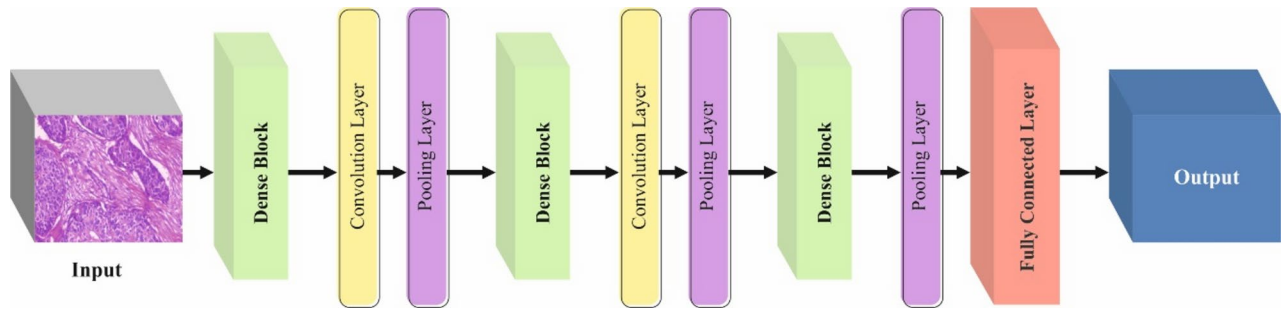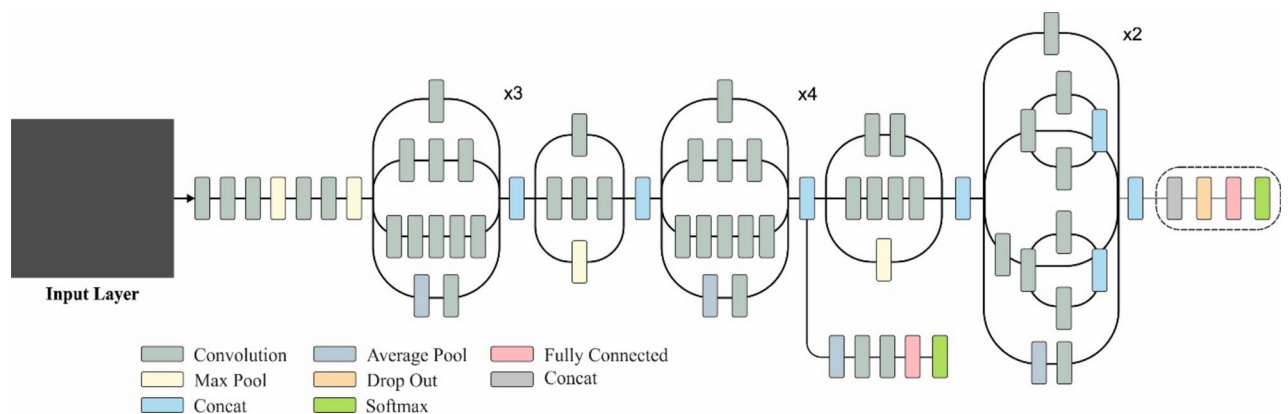
Fig. 3. Framework of DenseNet201.

Fig. 4. Architecture of InceptionNetV3 model.

global mean pooling that is tracked by a completely connected layer, and the activation function of $SoftMax$ is utilized for multiclass classification. Figure 3 illustrates the structure of DenseNet201.

InceptionV3
InceptionV3 is an exact neural network structure aimed at constantly detecting images. The network uses many inception components integrating parallel convolutional processes of diverse kernel dimensions, assisting the effectual capturing of features at numerous spatial measures. The inception component includes convolutional layers with pooling operations and numerous filter sizes, permitting the method to remove a vast range of distinct features from input imageries. Furthermore, InceptionV3 integrates support classifiers to tackle the problem of vanishing gradient in training. Furthermore, it uses BN and factorized 7×7 convolution (Conv) to reduce computational complexity while attaining higher performance. The structure is highly considered for its ability to operate numerous visual detection tasks with exceptional precision effectively. Figure 4 signifies the architecture of the InceptionV3 model.

MobileNetV2
MobileNetV2 is a very effectual CNN design that uses linear constraints and depth-wise segmented Conv to equalize computation complexity and accuracy. The structure includes numerous reversed residual blocks; every one contains convolutional layers, $ReLU$, and BN. Additionally, MobileNetV2 employs shortcut connections and development to increase the flow of data and feature reprocessing among layers. MobileNetV2 is appropriate for present object detection and image identification tasks on gadgets with restricted sources, lightweight operations, and effectual design ethics. Figure 5 demonstrates the architecture of the MobileNetV2 technique.

Classification using the SAE model
For the classification of BC diagnosis, the SAE model is exploited. The SAE network is an extensively applied DL architecture that contains two key modules of stack encoders (recognizing network) and decoders (generating network)[38]. This technique is selected for classification due to its capability to learn hierarchical representations of data, which is specifically beneficial for intrinsic datasets. The model captures increasingly abstract features by stacking various layers of autoencoders, allowing for enhanced feature extraction and dimensionality reduction. This capability improves the model's performance in tasks where conventional methodologies might face difficulty detecting relevant patterns. Furthermore, SAEs are robust to noise and can efficiently handle high-dimensional data, making them appropriate for medical imaging applications. Compared to other classification methods, namely SVMs or decision trees (DTs), SAEs can utilize unsupervised learning to pre-
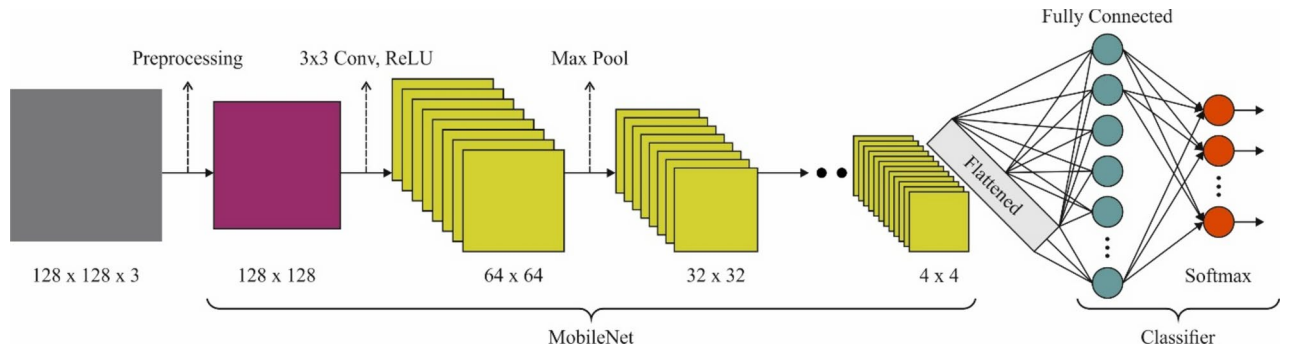
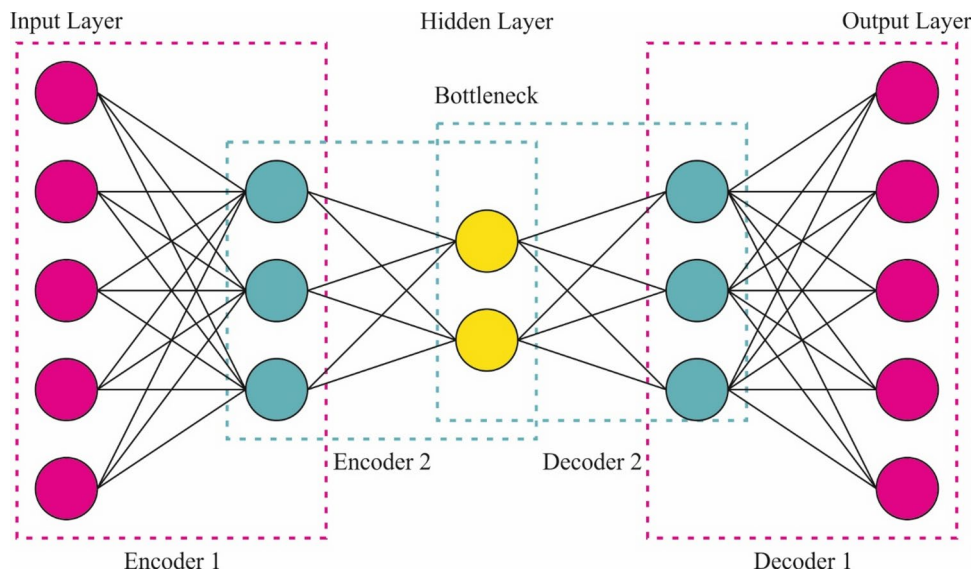**Fig. 5**. Structure of MobileNetV2 model.

Fig. 6. SAE framework.

train on unlabeled data, which can be advantageous in scenarios with limited labelled examples. This results in an improved generalization and improved accuracy in classification tasks. Figure 6 shows the architecture of the SAE methodology.

It portrays a proportioned structure with equivalent hidden neuron counts in the encoder and decoder layers. The encoding layers intellectualize the new input signals (now, multiple-element data concentration) into deeply featured codings that measure numerous top-level illustrations of multivariant signals (where high-level representations of geochemical background populations). These decoding layers utilize but are deeply featured codings for generating the output signals such that they are equivalent to the input signal's structure. The encoding and decoding non-linear functions for an experimental feature vector χ are described as

$$h_k = f_k(x) = \phi_{k,e}(W_{k,e}x_{k,i} + \beta_{k,e}),\tag{8}$$

$$\widehat{h}_l = g_1(x) = \phi_{l,d}(W_{l,d}f(x_{l,i}) + \beta_{l,d}),\tag{9}$$

whereas $f_k(x)$ and $g_l(x)$ represent functions of encoding and decoding in kth hidden layer (HL) (h_k) and lth HL ($\widehat{h}_l$), correspondingly. The $W_{k,e}$ and $W_{l,d}$ signify weighted matrices in h_k and h_l, whereas $\beta_{k,e}$ and $\beta_{1,d}$ showed biased vectors in h_k and h_1, individually. The formulations ($\phi_e(\cdot)$ and ($\phi_d(\cdot)$ symbolize activation functions of h_k and $\widehat{h}_l$ appropriately, and they might be adjusted through tasks like a hyperbolical tangent, linear, Maxout unit, sigmoid, step, $ramp$, and ReLU. Throughout fine-tuning in BP, the error results are utilized to upgrade the network-wide parameter over gradient background techniques to attain an experienced network of SAE.

HHOA-based parameter tuning

Eventually, the HHOA approach is applied to optimally select parameters involved in the SAE method. The HHOA model is employed for optimal parameter selection in the SAE method due to its effectualness in exploring intrinsic search spaces and finding optimal solutions. The HHOA methodology is inspired by horses' natural

behaviour, allowing it to balance exploration and exploitation effectually, which is significant for parameter tuning in DL methods. Compared to conventional optimization approaches such as grid or random search, HHOA is faster and often more efficient in converging to optimal or near-optimal solutions. Its population-based approach enables it to avert local minima, improving the likelihood of attaining enhanced overall performance. Furthermore, the adaptability of the HHOA model makes it appropriate for dynamic environments, allowing for fine-tuning in real-time as new data becomes available. This results in an enhanced accuracy and robustness of the SAE model in classification tasks. Figure 7 depicts the overall structure of the HHOA model.

The HHOA approach was advanced depending on the horses behaviour in their natural world at different ages[39]. Generally, the behaviour of horses is split into six types: Hierarchy, Grazing, Defense Mechanism, Sociability, Roaming, and Imitation. Each horse transfers based on the subsequent Eq. in every iteration:

$$X_i^{iter,\text{AGE}} = V_i^{iter,AGE} + X_i^{(iter-1),AGE}, AGE = \alpha, \beta, \gamma, \delta \tag{10}$$

Here, $X_i^{iter,AGE}$ establishes ith horse location, $V_i^{iter,AGE}$ offerings i_{th} horse velocity, AGE specifies the age limit of every horse, and $iter$ displays the present iteration. Every horse's velocity has been computed depending on its age range and its behavioural patterns inside an iteration.

$$V_i^{iter,\alpha} = G_i^{iter,\alpha} + D_i^{iter,\alpha}$$
$$V_i^{iter,\beta} = G_i^{iter,\beta} + H_i^{iter,\beta} + S_i^{iter,\beta} + D_i^{iter,\beta} \tag{11}$$
$$V_i^{iter,\gamma} = G_i^{iter,\gamma} + H_i^{iter,\gamma} + S_i^{iter,\gamma} + I_i^{iter,\gamma} + D_i^{iter,\gamma} + R_i^{iter,\gamma}$$
$$V_i^{iter,\delta} = G_i^{iter,\delta} + I_i^{iter,\delta} + R_i^{iter,\delta}$$

After this, individual and social intellect work for horses is considered.

Grazing (G)

Grazing is a normal horse behavioural pattern. Horses pasture approximately 70% on time in the day and 50% at night. People of all ages use roughly 16 to 20 h of daily pasture on pastures. The pasture area of every horse can be displayed using the HHOA method using the coefficient g Eqs. (12) and (13) define graze mathematics.
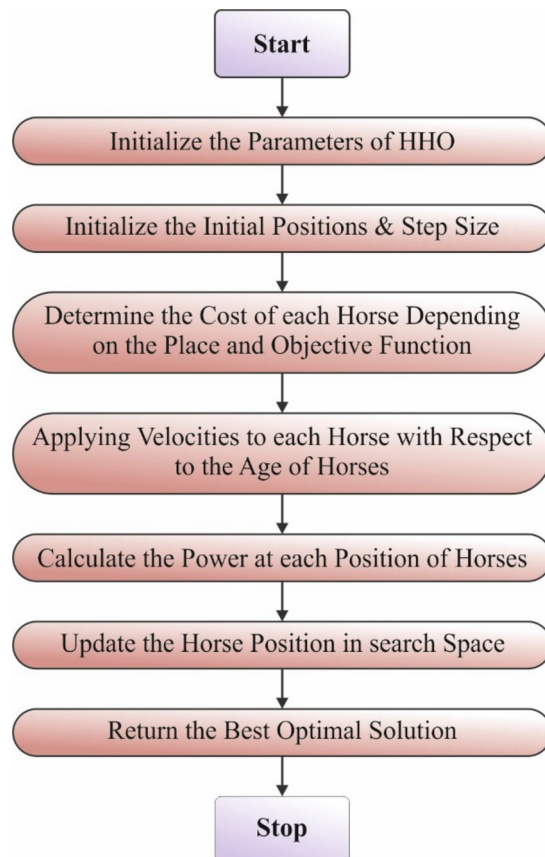

Fig. 7. Structure of HHOA model.

$$G_i^{iter,AGE} = g_{iter}\left(\breve{u} + \mathcal{P}\breve{1}\right)\left[X_i^{(iter-1)}\right], AGE = \alpha, \beta, \gamma, \delta \tag{12}$$

$$g_i^{iter,AGE} = g_i^{(iter-1),AGE} \times \omega_g \tag{13}$$

Now, $G_i^{iter,AGE}$ calculates the trend of the ith horse to grazing, which reduces linearly by ω_g at every iteration. Furthermore, $\breve{1}$ and $\breve{u}$ specify the lower and upper boundaries of the pasture field correspondingly, whereas $\mathcal{P}$ represents a random number among 0 and 1.

Hierarchy (H)

It is possible to find a hierarchy amongst horses every time since there are farm animals. The hierarchy safeguards horses but also permits them admittance to improved feeding lands. In the HHOA model, the coefficient h characterizes the horse's trend to track the horse with the best strength and knowledge. Various studies have presented that within the Middle Ages, β and γ, horses will have a behavioural hierarchy. Equations (14) and (15) define hierarchical behaviours as given below:

$$H_i^{iter,AGE} = h_i^{iter,AGE}\left[X_*^{(iter-1)} - X_i^{(iter-1)}\right], AGE = \alpha, \beta, \gamma \tag{14}$$

$$h_i^{iter,AGE} = h_i^{(iter-1),AGE} \times \omega_h \tag{15}$$

Here $X_*^{(iter-1)}$ establishes the best location for horses, and $H_i^{iter,AGE}$ illustrates how the best location affects the velocity parameter.

Sociability (S)

Social behaviour is a horse's characteristic. Sociability intends to improve predator-protecting methods and reduce scanning time. Additionally, sociability raises intra-group conflicts and competition, transmission of diseases, and the threat of attracted predators. Since flight represents the best-protected method for horses, it is fundamental to detect unseen predators immediately. Hence, the capacity for maintaining group awareness, acting as a cluster in the flight, and communicating efficiently about these activities will be of major significance for every horse of a flock. Factor s determines the horse's social behaviour, which can be designated in the subsequent Eqs. (16) and (17).

$$S_i^{iter,AGE} = s_i^{iter,AGE}\left[\left(\frac{1}{N}\sum_{j=1}^{N}X_j^{(iter-1)}\right) - X_i^{(iter-1)}\right], AGE = \beta, \gamma \tag{16}$$

$$S_i^{iter,AGE} = S_i^{(iter-1),AGE} \times \omega_s \tag{17}$$

$S_i^{iter,AGE}$ represents the social movement vector horse i, which reduces per iteration in the ω_s factor. Moreover, N denotes total horse counts. AGE states the age range of horses. Still, the average location has been gained in the following:

$$Mean\_Position = \left(\frac{1}{N}\sum_{j=1}^{N}X_j^{(iter-1)}\right) \tag{18}$$

Imitation (I)

As social animals, horses will learn from one another about fine and bad behaviour, like finding appropriate ground. Younger horses tend to mimic new horses during the age limit of 0-5 years. Additionally, factor i in the HHOA method shows this characteristic of the behaviour of horses. Simulated in horse herds is described below:

$$I_i^{iter,AGE} = i_i^{iter,AGE}\left[\left(\frac{1}{pN}\sum_{j=1}^{pN}\widehat{X}_j^{(iter-1)}\right) - X^{(iter-1)}\right], AGE = \gamma \tag{19}$$

$$i_i^{iter,\,AGE} = i_i^{(iter-1),AGE} \times \omega i \tag{20}$$

$I_i^{iter,AGE}$ determines ith horse movement near the mean location of the finest horses positioned at $\widehat{X}$. pN specifies horse counts within the finest area. The mentioned p-value is 10% of the total horses. Besides, ω_i is stated as a reducing factor for every iteration, as revealed earlier. Also, the finest position can be gained in the succeeding Eq. (21).

$$Best\text{-}Position = \left(\frac{1}{pN}\sum_{j=1}^{pN}\widehat{X}_j^{(iter-1)}\right) \tag{21}$$

Defense mechanism (D)

Horses have dual main protection methods: fight and flight. Once challenged with a critical condition, horses typically flee, and fighting helps as a subordinate protective method. Factor d defines the survival procedure of horses. Equations (22) and (23) with negative coefficients signify the survival mechanisms of horses, which stop them from reaching unsuitable locations.

$$D_i^{iter,}AGE =_- d_i^{iter,\mathrm{AGE}} \left[\left(\frac{1}{qN} \sum\nolimits_{j=1}^{qN} \widehat{X}_j^{(iter-1)} \right) - X^{(iter-1)} \right], AGE = \alpha, \beta, \gamma \qquad (22)$$

$$d_i^{iter,}AGE = d_i^{(iter-1),AGE} \times \omega_d \qquad (23)$$

$D_i^{iter,AGE}$ symbolizes the fleeing vector of the horse i, depending on the horse's mean in the poor parts. The horse counts with the poorest positions are also exposed to qN. q equates to 20% of each horse. As previously specified, ω_d characterizes the reducing factor for every iteration. Still, the poorest location has been considered in the following:

$$Worst - Position = \left(\frac{1}{qN} \sum\nolimits_{j=1}^{qN} \widehat{X}_j^{(iter-1)} \right) \qquad (24)$$

Roaming (R)

By the age limit of 5 to 15 years, young horses also tended to transfer from grazing to pasture in search of nutrition. Horses were interested animals who searched for novel grazing where they could. As horses attain adulthood, they reduce their behaviour by roaming. The factor r displays this horse's behaviour for randomized motion in Eqs. (25) and (26).

$$R_i^{iter,AGE} = r_i^{iter,AGE} \mathcal{P} X^{(iter-1)}, AGE = \gamma, \delta \qquad (25)$$

$$r_i^{iter,AGE} = r_i^{(iter-1),AGE} \times \omega_r \qquad (26)$$

In such a case, $Ri iter, AGE can denote it h$ horse's randomly generated velocity vector, and its reducing factor is designated by ω_r. To compute the total velocity, the sociability, grazing, defence method, imitation, hierarchy, and roaming were replaced in Eq. (11).

By using a sorting method across a global matrix, HHOA utilizes a suitable methodology to boost the speediness of problem resolution whereas moreover evading local optimum entrapment. Equations (27) and (28) represent the global matrix that can be made by juxtaposing locations (X) and the value of cost for each location $(C(X))$.

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & x_{2,3} & \cdots & x_{2,d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_{m,1} & x_{m,2} & x_{m,3} & \cdots & x_{m,d} \end{bmatrix}, C(X) = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} \qquad (27)$$

$$Global\ Matrix = [XC(X)] = \begin{bmatrix} x_{1,1} & x_{1,2} & x_{1,3} & \cdots & x_{1,d} & c_1 \\ x_{2,1} & x_{2,2} & x_{2,3} & \cdots & x_{2,d} & c_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{m,1} & x_{m,2} & x_{m,3} & \cdots & x_{m,d} & c_m \end{bmatrix} \qquad (28)$$

The HHOA method develops a fitness function (FF) to improve classification performance. It identifies a positive integer to signify the enhanced performance of the candidate solutions. In this research, the minimization of the classification rate of error can be identified as an FF, as specified in Eq. (29).

$$fitness(x_i) = Classifier Error Rate(x_i) = \frac{number\ of\ misclassified\ samples}{Total\ number\ of\ samples} \times 100 \qquad (29)$$

Experimental validation

In this sector, the CVFBJTL-BCD model's experimental validation is verified using the histopathological datasets[40] and the Ultrasound dataset[41]. The histopathological dataset comprises 9,109 microscopic images of breast tumour tissue from 82 patients, categorized into 2,480 benign and 5,429 malignant samples. The images are collected utilizing the SOB approach (excisional biopsy) at diverse magnifications (40X, 100X, 200X, and 400X) in PNG format, measuring 700×460 pixels with 8-bit RGB depth. The ultrasound dataset, accumulated over a year at Baheya Hospital in DICOM format, initially encompasses 1,100 grayscale images but was refined to 780 after preprocessing to remove duplicates and irrelevant data. Scans were performed by employing LOGIQ E9 ultrasound systems, giving 1280×1024 pixel resolution images. Preprocessing encompasses annotation corrections and conversion to PNG, with ground truth boundaries established using MATLAB through freehand segmentation. The datasets are organized into folders by cancer type, with filenames indicating class and including a "_mask" suffix for segmented images. Figure 8 denotes the sample images of two datasets. The histopathological 200X dataset contains 2013 images under 2 classes as shown in Table 1. The suggested technique is simulated using the Python 3.6.5 tool on PC i5-8600 k, 250 GB SSD, GeForce 1050Ti 4 GB, 16 GB RAM, and 1 TB HDD. The parameter settings are provided: learning rate: 0.01, activation: ReLU, epoch count: 50, dropout: 0.5, and batch size: 5.

Figure 9 represents the classification outcomes of the CVFBJTL-BCD process on the Histopathological dataset. Figure 9a-9b displays the confusion matrix with precise identification and classification of all class labels
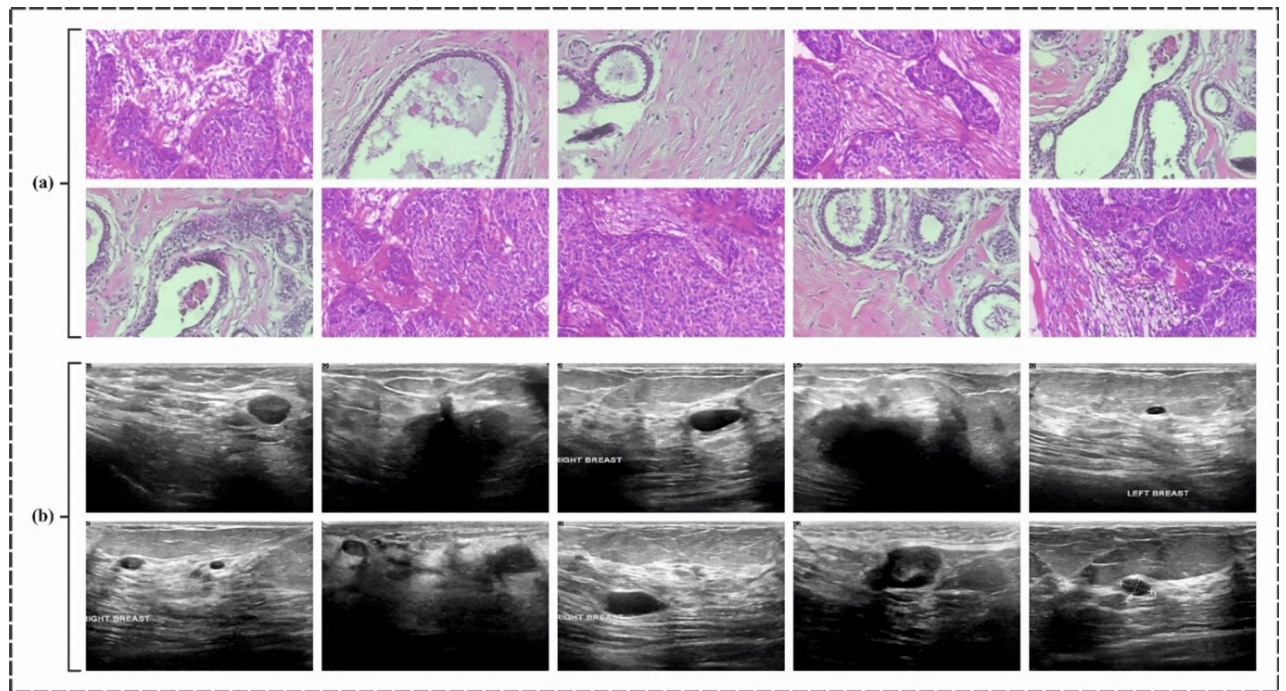
Fig. 8. Sample Images (**a**) Histopathological Images (**b**) Ultrasound Images.

Histopathological dataset 200X	
Class	**No. of images**
Benign	623
Malignant	1390
Total images	2013

Table 1. Details on Histopathological Dataset.

on a 70:30 TRAP/TESP. Figure 9c demonstrates the PR analysis, indicating maximum performance across all classes. Lastly, Fig. 9d represents the ROC analysis, showing efficient results with high ROC values for distinct class labels.

In Table 2 and Fig. 10, the BC detection outcomes of the CVFBJTL-BCD model are represented on the Histopathological dataset. The results reported that the CVFBJTL-BCD process discriminated adequately between benign and malignant samples. On 70%TRAP, the CVFBJTL-BCD technique offers an average $accu_y$ of 97.37%, $prec_n$ of 97.13%, $sens_y$ of 96.70%, $spec_y of$ 96.70% and $F1_{score}$ of 96.91%. Moreover, on 30%TESP, the CVFBJTL-BCD approach provides an average $accu_y$ of 98.18%, $prec_n$ of 98.38%, $sens_y$ of 97.37%, $spec_y of$ 97.37%, and $F1_{score}$ of 97.85%.

Figure 11 illustrates the training (TRA) and validation (VLA) accuracy outcomes of the CVFBJTL-BCD process on the Histopathological dataset. The accuracy values are computed throughout 0–50 epochs. The figure emphasized that the TRA and VLA accuracy values show a rising trend, which indicates the capability of the CVFBJTL-BCD method with improved performance over several iterations. Additionally, the training and validation accuracy remain closer over the epochs, indicating low minimal overfitting and displaying the enhanced performance of the CVFBJTL-BCD technique, guaranteeing consistent prediction on unseen samples.

Figure 12 presents the TRA and VLA loss graph of the CVFBJTL-BCD approach on the Histopathological dataset. The loss values are computed for 0–50 epochs. The TRA and VLA accuracy values show a reduced trend, notifying the CVFBJTL-BCD approach's ability to balance a trade-off between data fitting and generalization. The continual reduction in loss values guarantees the enhanced performance of the CVFBJTL-BCD methodology and tunes the prediction outcomes over time.

Table 3 examines the comparison outcomes of the CVFBJTL-BCD approach on the Histopathological dataset with existing models[42,43]. Figure 13 illustrates the comparative outcomes of the CVFBJTL-BCD approach on the Histopathological dataset with existing models in terms of $accu_y$ and $prec_n$. The outcomes emphasized that the IncepceptionV3, VGG16, and ResNet-50 models have stated worse performance. In the meantime, DTLRO-HCBC, IncepceptionV3 -LSTM, and IncepceptionV3-BiLSTM methodology have got nearer outcomes. At the
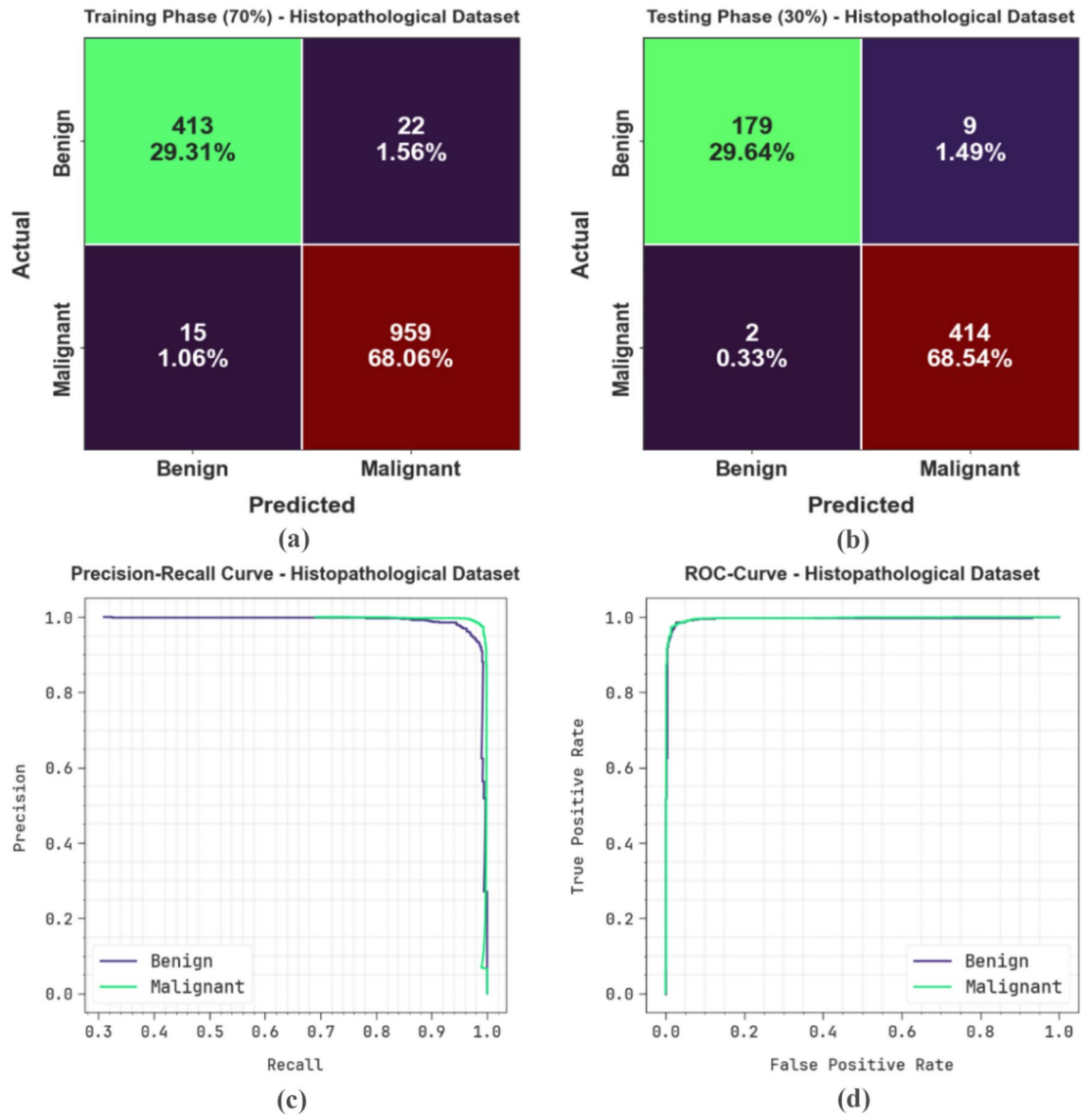
Fig. 9. Histopathological dataset (**a-b**) Confusion matrices and (**c-d**) PR and ROC curves.

Class	$Accu_y$	$Prec_n$	$Sens_y$	$Spec_y$	$F1_{score}$
TRAP (70%)					
Benign	97.37	96.50	94.94	98.46	95.71
Malignant	97.37	97.76	98.46	94.94	98.11
Average	**97.37**	**97.13**	**96.70**	**96.70**	**96.91**
TESP (30%)					
Benign	98.18	98.90	95.21	99.52	97.02
Malignant	98.18	97.87	99.52	95.21	98.69
Average	**98.18**	**98.38**	**97.37**	**97.37**	**97.85**

Table 2. BC detection outcomes of CVFBJTL-BCD technique on Histopathological dataset.
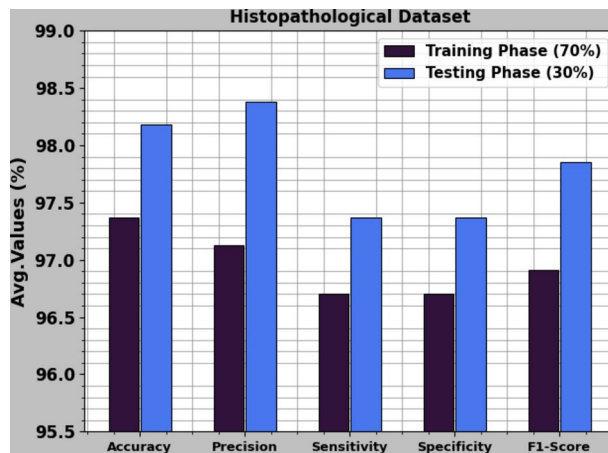
Fig. 10. Average of CVFBJTL-BCD technique on Histopathological dataset.
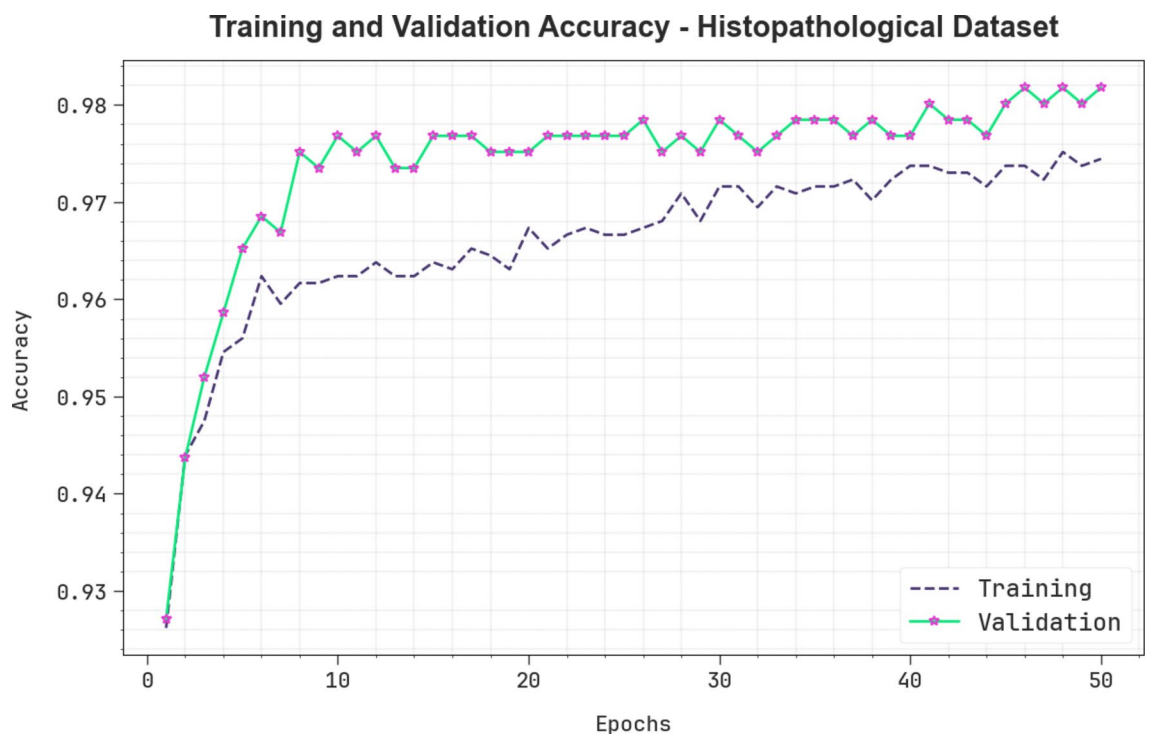
Fig. 11. $Accu_y$ curve of CVFBJTL-BCD technique on Histopathological dataset.

same time, the AOADL-HBCC process got a higher $accu_y$ of 96.95% and $prec_n$ of 95.62%. Also, the CVFBJTL-BCD method stated improved performance with the greatest $accu_y$ of 98.18% and $prec_n$ of 98.38%.

Figure 14 displays the comparative outcomes of the CVFBJTL-BCD technique on the Histopathological dataset with existing models in terms of $sens_y$ and $spec_y$. The results highlighted that the IncepceptionV3, VGG16, and ResNet-50 techniques have stated worse performance. In the meantime, DTLRO-HCBC, IncepceptionV3-LSTM, and IncepceptionV3-BiLSTM techniques have got nearer outcomes. At the same time, the AOADL-HBCC technique got closer $sens_y$ of 96.03% and $spec_y$ of 96.51%. Also, the CVFBJTL-BCD method stated improved performance with the greatest $sens_y$ of 97.37% and $spec_y$ of 97.37%.

Table 4 and Fig. 15 inspect the computational time (CT) analysis of the CVFBJTL-BCD approach on the Histopathological dataset with existing models. The histopathological dataset analysis reveals diverse computational times for various methods. The CVFBJTL-BCD method performed the fastest at 1.54 s, followed by the VGG16 Classifier at 3.17 s and DTLRO-HCBC at 3.66 s. AOADL-HBCC required 5.23 s, while ResNet-50 took 4.94 s. InceptionV3 had a longer CT of 8.22 s, with its LSTM variant at 6.87 s and the BiLSTM variant at 4.31 s.

The ultrasound dataset contains 780 images under 3 classes, as exposed in Table 5.
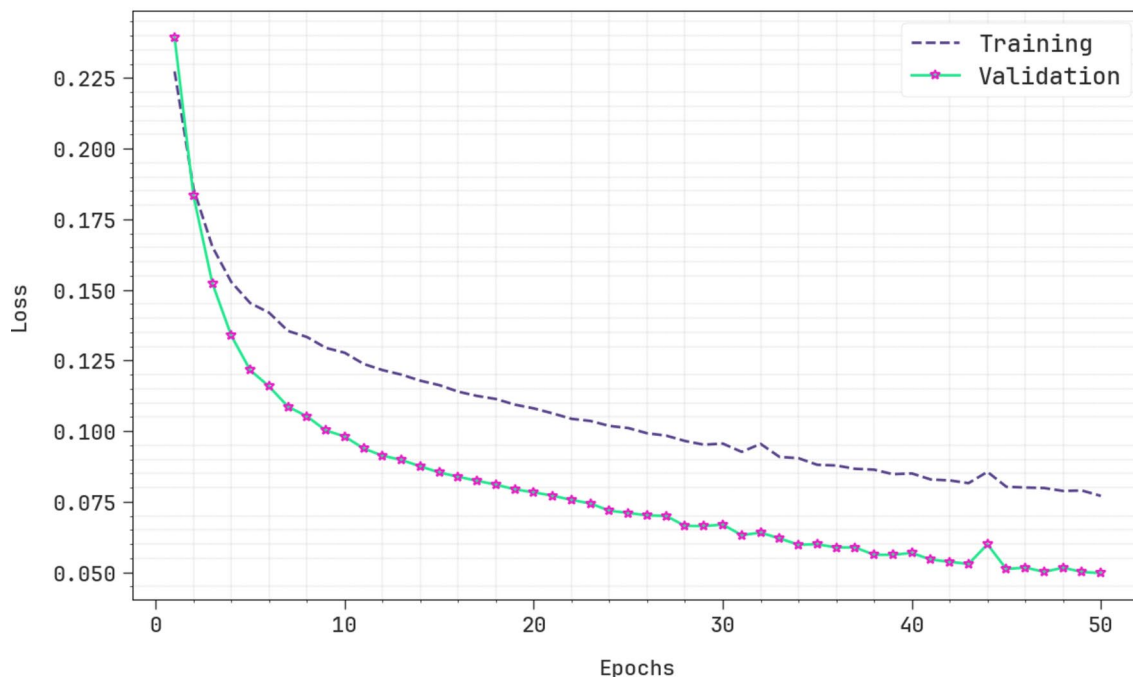
Fig. 12. Loss curve of CVFBJTL-BCD technique on Histopathological dataset.

Histopathological dataset				
Methods	$Accu_y$	$Prec_n$	$Sens_y$	$Spec_y$
CVFBJTL-BCD	98.18	98.38	97.37	97.37
AOADL–HBCC	96.95	95.62	96.03	96.51
DTLRO–HCBC	93.68	90.57	94.84	93.52
IncepceptionV3	81.84	86.56	90.74	91.05
IncepceptionV3 -LSTM	91.61	92.70	92.55	92.60
IncepceptionV3 -BiLSTM	92.23	92.85	96.90	91.12
VGG16 Classifier	80.32	89.36	89.92	90.76
ResNet-50	82.33	83.46	91.06	93.40

Table 3. Comparative analysis of CVFBJTL-BCD technique on Histopathological dataset with existing models[42,43].

Figure 16 represents the classification results of the CVFBJTL-BCD process on the Ultrasound dataset. Figure 16a-16b displays the confusion matrices with accurate identification and classification of all class labels on a 70:30 TRAP/TESP. Figure 16c represents the PR analysis, indicating maximum performance across all classes. Lastly, Fig. 16d exhibits the ROC analysis, representing efficient outcomes with high ROC values for distinct class labels.

In Table 6 and Fig. 17, the BC detection results of the CVFBJTL-BCD approach are denoted on the Ultrasound dataset. The outcomes reported that the CVFBJTL-BCD approach discriminated adequately between benign and malignant samples. On 70%TRAP, the CVFBJTL-BCD approach provides an average $accu_y$ of 98.66%, $prec_n$ of 97.61%, $sens_y$ of 97.31%, $spec_y$ of 98.84% and $F1_{score}$ of 97.44%. Also, on 30%TESP, the CVFBJTL-BCD method offers an average $accu_y$ of 99.15%, $prec_n$ of 97.50%, $sens_y$ of 98.55%, $spec_y$ of 99.49%, and $F1_{score}$ of 97.96%.

Figure 18 shows the TRA and VLA accuracy outcomes of the CVFBJTL-BCD methodology on the ultrasound dataset. The accuracy values are computed for 0–50 epochs. The figure emphasizes that the TRA and VLA accuracy values are rising, indicating the CVFBJTL-BCD methodology's capability with improved performance over several iterations. Additionally, the training and validation accuracy remain closer over the epochs, indicating minimal overfitting and enhancing the CVFBJTL-BCD approach's performance, guaranteeing consistent prediction on unseen samples.

Figure 19 demonstrates the TRA and VLA loss graph of the CVFBJTL-BCD approach on the Ultrasound dataset. The loss values are computed for 0–50 epochs. The TRA and VLA accuracy values show a lower trend,
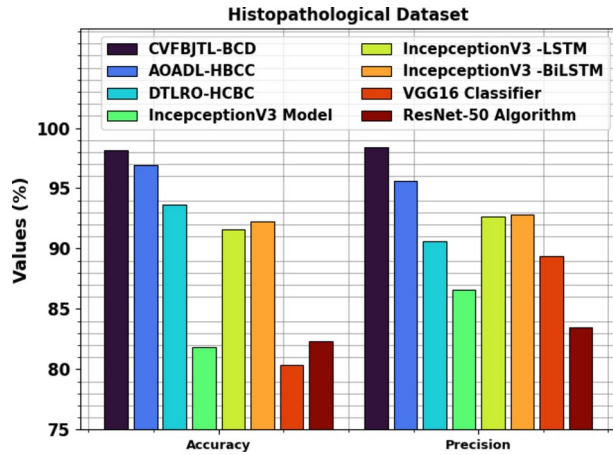
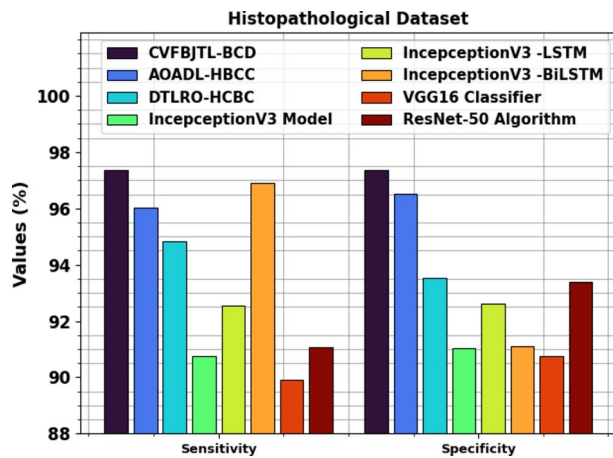**Fig. 13**. $Accu_y$ and $Prec_n$ analysis of CVFBJTL-BCD technique on Histopathological dataset.

Fig. 14. $Sens_y$ and $Spec_y$ analysis of CVFBJTL-BCD technique on Histopathological dataset.

Histopathological dataset	
Methods	**CT (sec)**
CVFBJTL-BCD	1.54
AOADL-HBCC	5.23
DTLRO-HCBC	3.66
IncepceptionV3	8.22
IncepceptionV3 -LSTM	6.87
IncepceptionV3 -BiLSTM	4.31
VGG16 Classifier	3.17
ResNet-50	4.94

Table 4. CT analysis of CVFBJTL-BCD technique on Histopathological dataset with existing models.

notifying the CVFBJTL-BCD approach's ability to balance a trade-off between data fitting and generalization. The continual reduction in loss values guarantees the enhanced performance of the CVFBJTL-BCD method and tunes the prediction results over time.

Table 7 examines the comparison results of the CVFBJTL-BCD methodology on the Ultrasound dataset with existing models. Figure 20 exhibits the comparative outcomes of the CVFBJTL-BCD approach on the Ultrasound dataset with existing models in terms of $accu_y$ and $prec_n$. The results highlighted that the CNN-VGG16, CNN-VGG19, CNN-Resnet10, CNN-Densnet121, and CNN-Densnet161 methods have stated worse performance. Meanwhile, the EDLCDS-BCDC and ODTLD-CABCC methodologies have yielded nearer outcomes. Also, the CVFBJTL-BCD model stated improved performance with the greatest $accu_y$ of 99.15% and $prec_n$ of 97.50%.
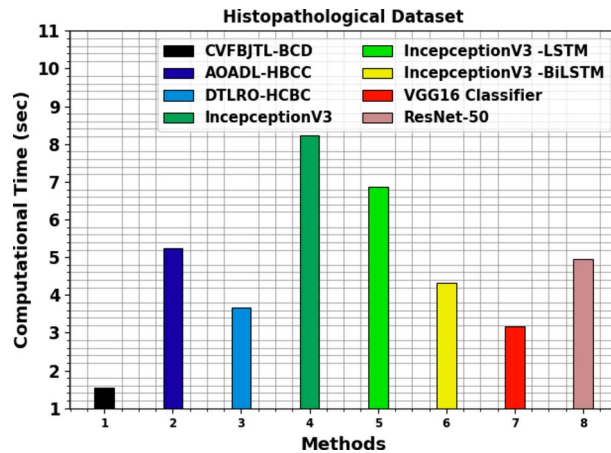
Fig. 15. CT analysis of CVFBJTL-BCD technique on Histopathological dataset with existing models.

Ultrasound dataset	
Class	Image count
Normal	133
Benign	437
Malignant	210
Total images	780

Table 5. Details on dataset.

Figure 21 represents the comparative results of the CVFBJTL-BCD method on the Ultrasound dataset with existing models in terms of $sens_y$ and $spec_y$. The results emphasized that the CNN-VGG16, CNN-VGG19, CNN-Resnet10, CNN-Densnet121, and CNN-Densnet161 processes have stated worse performance. Meanwhile, the EDLCDS-BCDC and ODTLD-CABCC approaches have yielded nearer outcomes. Also, the CVFBJTL-BCD methodology indicated improved performance with the greatest $sens_y$ of 98.55% and $spec_y$ of 99.49%.

Table 8 and Fig. 22 investigate the CT analysis of the CVFBJTL-BCD methodology on the Ultrasound dataset with existing models. The ultrasound dataset analysis exhibits varying CTs for diverse approaches: CNN-VGG16 took 11.75 s, CNN-VGG19 was faster at 6.52 s, while CNN-Resnet101 required 12.58 s. Other methods comprise CNN-Densnet121 at 14.21 s, CNN-Densnet161 at 13.14 s, EDLCDS-BCDC at 10.33 s, and ODTLD-CABCC at 10.49 s. On the contrary, the CVFBJTL-BCD method attained the shortest CT of 4.96 s.

Conclusion

This research proposes a new CVFBJTL-BCD technique using multi-modality medical imaging. The presented CVFBJTL-BCD technique gets the perceptions of feature fusion and DL models for effectively detecting and identifying BC diagnosis. Primarily, the presented CVFBJTL-BCD technique applies GF for the noise removal process. Next, the CVFBJTL-BCD technique uses a fusion-based joint TL process comprising three models, namely DenseNet201, InceptionV3, and MobileNetV2. For the classification of BC diagnosis, the SAE model is exploited. Eventually, the HHOA is utilized to select parameters involved in the SAE model optimally. To demonstrate the improved results of the CVFBJTL-BCD methodology, a wide series of experimentations are performed on two benchmark datasets. The comparative analysis of the CVFBJTL-BCD technique portrayed a superior accuracy value of 98.18% and 99.15% over existing methods under Histopathological and Ultrasound datasets. The limitations of the CVFBJTL-BCD technique encompass the reliance on specific datasets that may not fully represent the diversity of breast tumour types and imaging conditions, which can hinder the generalizability of outcomes. Furthermore, discrepancies in imaging equipment and protocols may present biases affecting diagnostic accuracy. Future work may expand the datasets to include a broader range of tumour types and imaging modalities, incorporating ultrasound with MRI or histopathological data to improve diagnostic precision. Exploring advanced DL and optimization models could enhance feature extraction and classification performance. Finally, developing real-time analysis capabilities and user-friendly interfaces will facilitate clinical adoption and support improved decision-making in BC diagnosis.
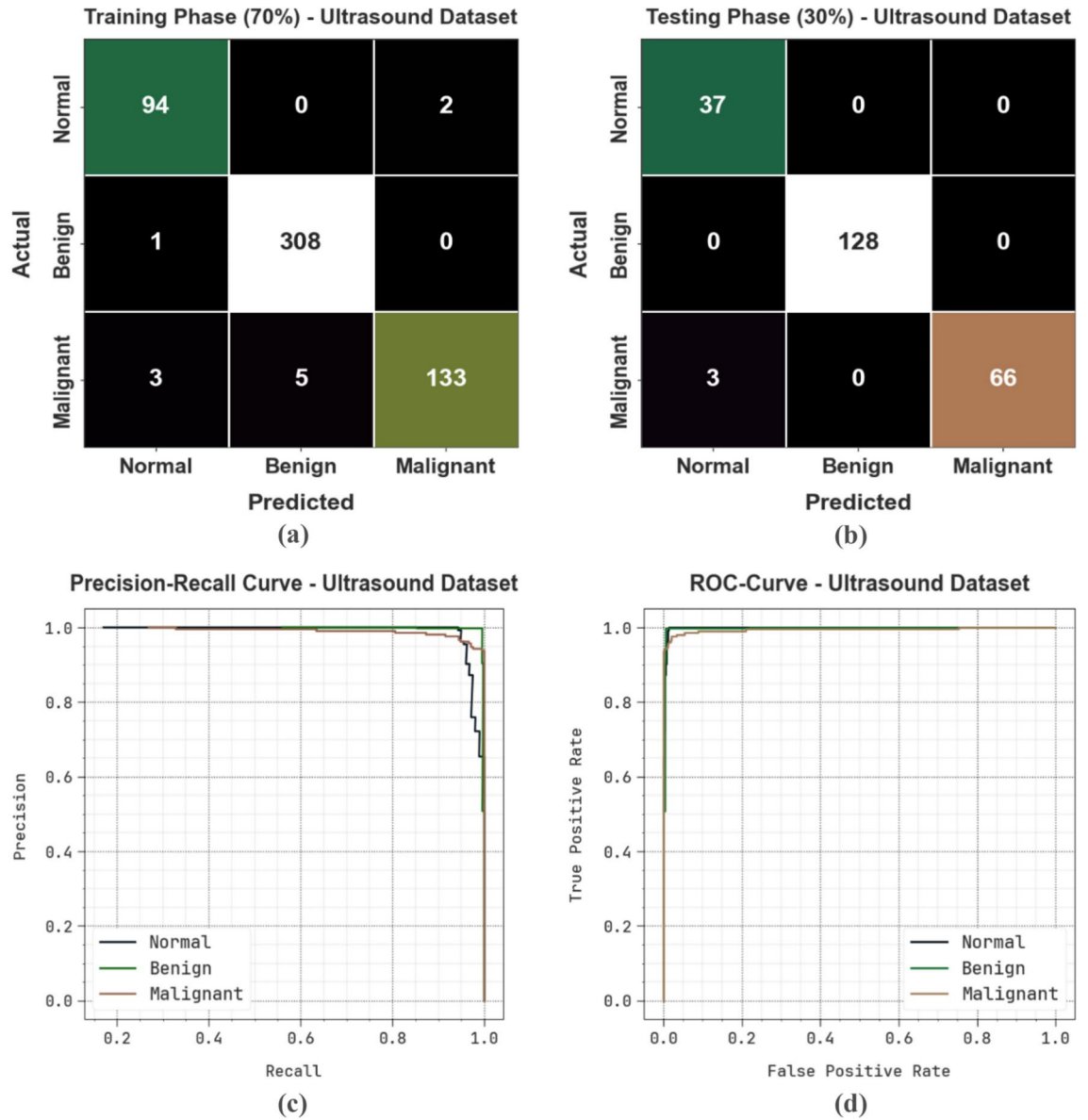
Fig. 16. Ultrasound Dataset (**a-b**) Confusion matrices and (**c-d**) PR and ROC curves.

Class	$Accu_y$	$Prec_n$	$Sens_y$	$Spec_y$	$F1_{score}$
TRAP (70%)					
Normal	98.90	95.92	97.92	99.11	96.91
Benign	98.90	98.40	99.68	97.89	99.04
Malignant	98.17	98.52	94.33	99.51	96.38
Average	**98.66**	**97.61**	**97.31**	**98.84**	**97.44**
TESP (30%)					
Normal	98.72	92.50	100.00	98.48	96.10
Benign	100.00	100.00	100.00	100.00	100.00
Malignant	98.72	100.00	95.65	100.00	97.78
Average	**99.15**	**97.50**	**98.55**	**99.49**	**97.96**

Table 6. BC detection outcomes of CVFBJTL-BCD technique on Ultrasound dataset.
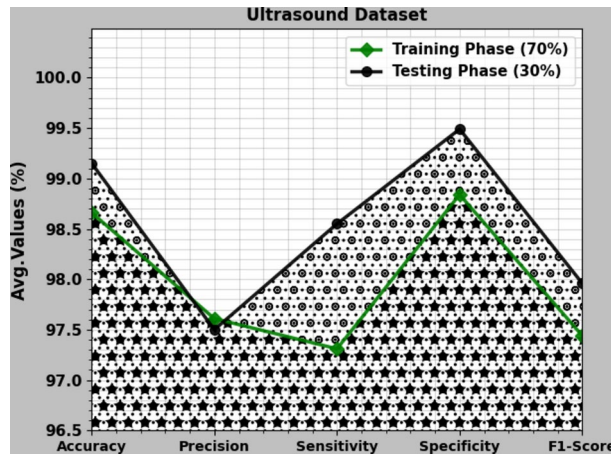
Fig. 17. Average of CVFBJTL-BCD technique on Ultrasound dataset.
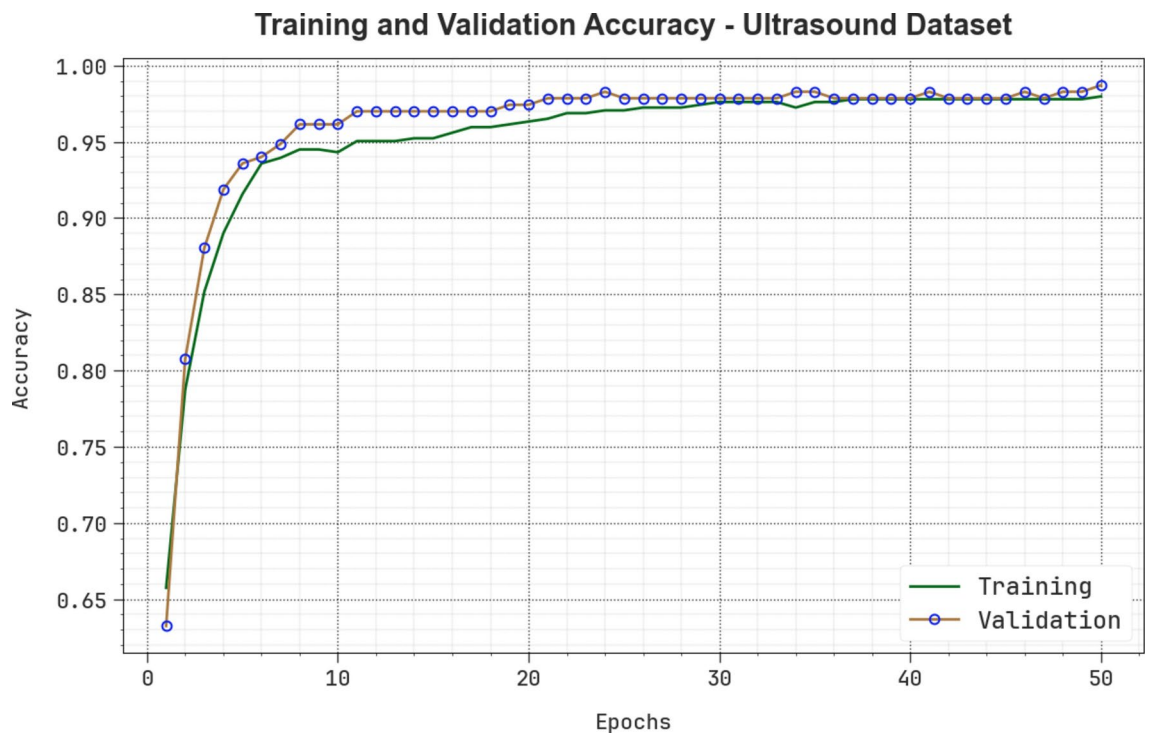
Fig. 18. $Accu_y$ curve of CVFBJTL-BCD technique on Ultrasound dataset.
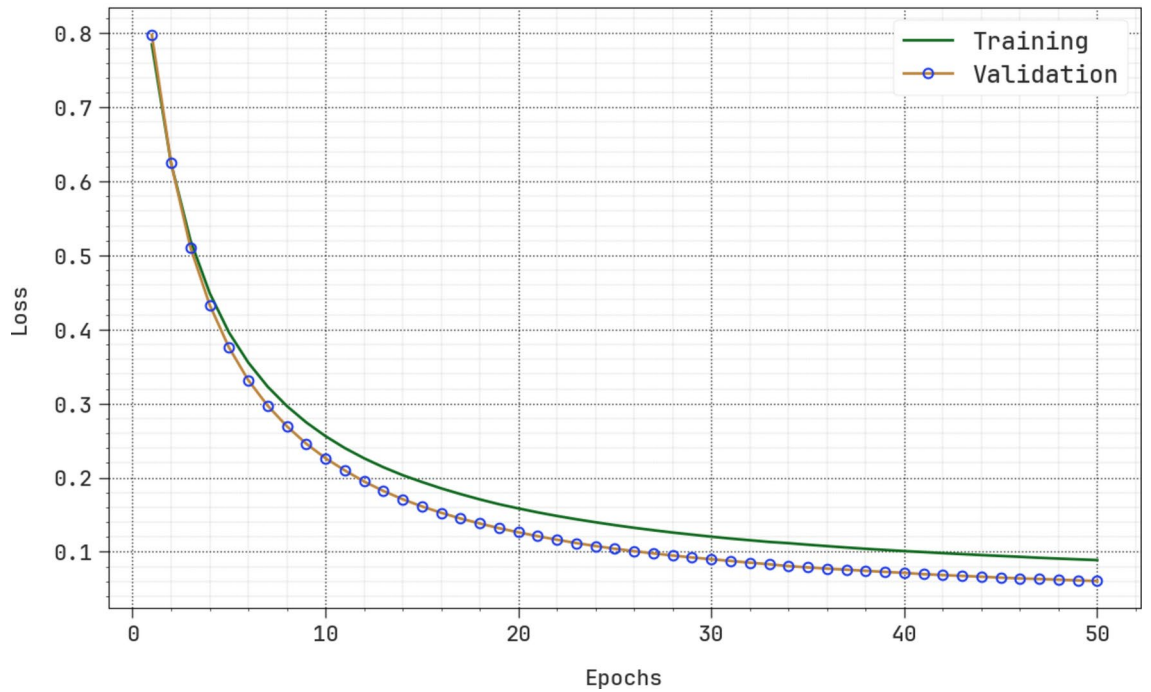
Fig. 19. Loss curve of CVFBJTL-BCD technique on Ultrasound dataset.

Ultrasound dataset				
Methods	$Accu_y$	$Sens_y$	$Spec_y$	$Prec_n$
CNN-VGG16	91.15	85.72	90.42	87.32
CNN-VGG19	92.80	88.54	95.51	94.50
CNN-Resnet101	92.06	91.32	96.75	93.72
CNN-Densnet121	91.10	89.51	94.73	94.39
CNN-Densnet161	88.49	83.82	96.42	95.59
EDLCDS-BCDC	97.40	95.99	98.69	95.39
ODTLD-CABCC	98.96	98.49	99.24	97.36
CVFBJTL-BCD	99.15	98.55	99.49	97.50

Table 7. Comparative analysis of CVFBJTL-BCD technique on Ultrasound dataset with existing models.
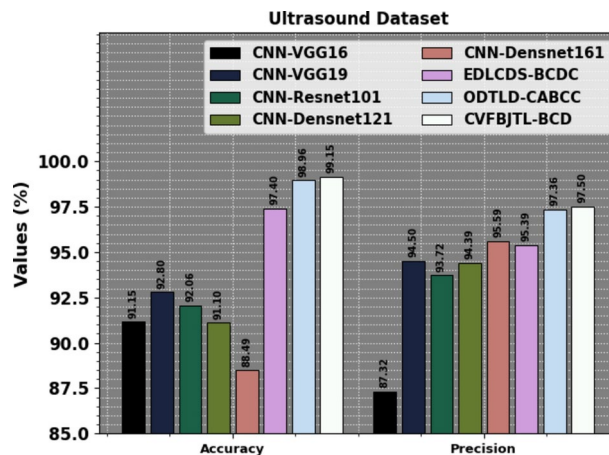
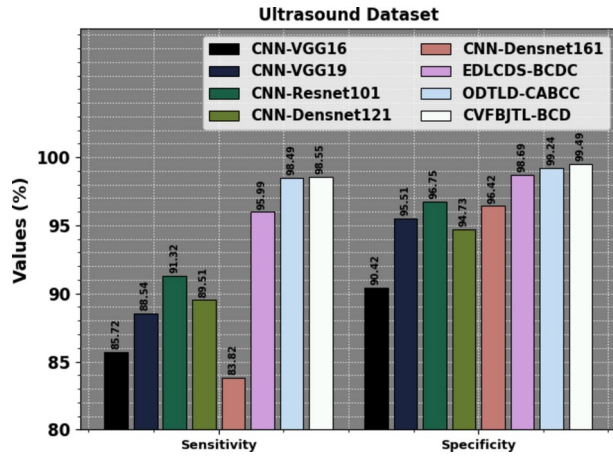**Fig. 20**. $Accu_y$ and $Prec_n$ analysis of CVFBJTL-BCD technique on Ultrasound dataset.

Fig. 21. $Sens_y$ and $Spec_y$ analysis of CVFBJTL-BCD technique on Ultrasound dataset.

Ultrasound dataset	
Methods	**Computational time (sec)**
CNN-VGG16	11.75
CNN-VGG19	6.52
CNN-Resnet101	12.58
CNN-Densnet121	14.21
CNN-Densnet161	13.14
EDLCDS-BCDC	10.33
ODTLD-CABCC	10.49
CVFBJTL-BCD	4.96

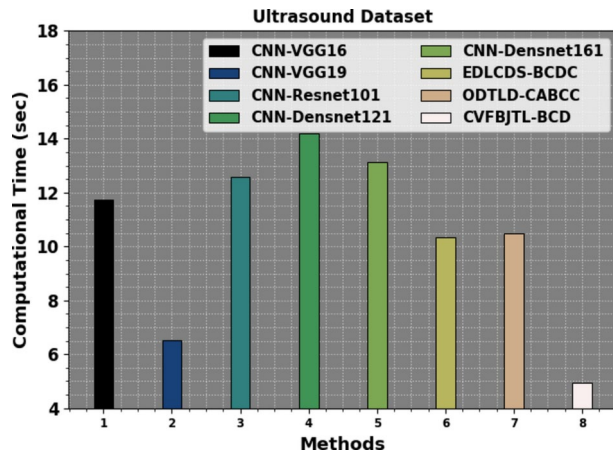**Table 8**. CT evaluation of CVFBJTL-BCD approach on Ultrasound dataset with existing methods.

Fig. 22. CT evaluation of the CVFBJTL-BCD technique on Ultrasound dataset with existing methods.

Data availability

The data that support the findings of this study are openly available in the Kaggle repository at https://web.inf.ufpr.br/vri/databases/breast-cancer-histopathological-database-breakhis/ [40].

References

1. Boumaraf, S., Liu, X., Zheng, Z., Ma, X. & Ferkous, C. A new transfer learning based approach to magnification dependent and independent classification of breast cancer in histopathological images. *Biomed. Signal Process. Control* **63**, 102192 (2021).
2. Bose, S.; Garg, A.; Singh, S.P. Transfer Learning for Classification of Histopathology Images of Invasive Ductal Carcinoma in Breast. In Proceedings of the 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 17–19 August 2022; pp. 1039–1044.
3. Ahmad, N., Asghar, S. & Gillani, S. A. Transfer learning-assisted multi-resolution breast cancer histopathological images classification. *Vis. Comput.* **38**, 2751–2770 (2021).
4. Issarti, A., Consejo, M., Jiménez-García, S., Hershko, C. & Koppen, J. J. Rozema, Computer aided diagnosis for suspect keratoconus detection. *Comput. Biol. Med.* **109**, 33–42. https://doi.org/10.1016/j.compbiomed.2019.04.024 (2019).
5. Habib, M. et al. Detection of microaneurysms in retinal images using an ensemble classifier. *Informat. Med. Unlock.* **9**, 44–57 (2017).
6. Nazir, T. et al. Diabetic retinopathy detection through novel tetragonal local octa patterns and extreme learning machines. *Artif. Intell. Med.* **99**, 101695 (2019).
7. Ahmad, H.M.; Ghuffar, S.; Khurshid, K. Classification of breast cancer histology images using transfer learning. In Proceedings of the 2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST), Islamabad, Pakistan, 8–12 January 2019; pp. 328–332.
8. Alzubaidi, L. et al. Optimizing the performance of breast cancer classification by employing the same domain transfer learning from hybrid deep convolutional neural network model. *Electronics* **9**, 445 (2020).
9. Tsiknakis, N. et al. Deep learning for diabetic retinopathy detection and classification based on fundus images: A review. *Comput. Biol. Med.* **135**, 104599 (2021).
10. Ai, Z. et al. DR-IIXRN : Detection algorithm of diabetic retinopathy based on deep ensemble learning and attention mechanism. *Front. Neuroinform.* **15**, 778552. https://doi.org/10.3389/fninf.2021.778552 (2021).
11. Sushanki, S., Bhandari, A. K. & Singh, A. K. A review of computational methods for breast cancer detection in ultrasound images using multi-image modalities. *Archives of Computational Methods in Engineering* **31**(3), 1277–1296 (2024).
12. Abdullakutty, F., Akbari, Y., Al-Maadeed, S., Bouridane, A. and Hamoudi, R., 2024. Advancing Histopathology-Based Breast Cancer Diagnosis: Insights into Multi-Modality and Explainability. arXiv preprint arXiv:2406.12897.
13. Oyelade, O. N., Irunokhai, E. A. & Wang, H. A twin convolutional neural network with a hybrid binary optimizer for multi-modal breast cancer digital image classification. *Scientific Reports* **14**(1), 692 (2024).
14. Hu, H., Liang, W., Zou, X. and Zou, X., 2024, July. Graph Convolutional Networks Based Multi-modal Data Integration for Breast Cancer Survival Prediction. In International Conference on Intelligent Computing (pp. 85–98). Singapore: Springer Nature Singapore.
15. Yang, X. et al. Triple-attention interaction network for breast tumor classification based on multi-modality images. *Pattern Recognition* **139**, 109526 (2023).
16. Guo, D. et al. A multi-modal breast cancer diagnosis method based on Knowledge-Augmented Deep Learning. *Biomedical Signal Processing and Control* **90**, 105843 (2024).
17. Zhang, J., Chen, Q., Zhou, L., Cui, Z., Gao, F., Li, Z., Feng, Q. and Shen, D., 2023, October. MoSID: Modality-specific information disentanglement from multi-parametric MRI for breast tumor segmentation. In MICCAI Workshop on Cancer Prevention through Early Detection (pp. 94–104). Cham: Springer Nature Switzerland.
18. Yan, P., Gong, W., Li, M., Zhang, J., Li, X., Jiang, Y., Luo, H. and Zhou, H., 2024. TDF-Net: Trusted Dynamic Feature Fusion Network for breast cancer diagnosis using incomplete multi-modal ultrasound. Information Fusion, p.102592.
19. Karthik, A. et al. Ensemble-based multi-modal medical imaging fusion for tumor segmentation. *Biomedical Signal Processing and Control* **96**, 106550 (2024).
20. Atrey, K., Singh, B. K. & Bodhey, N. K. Integration of ultrasound and mammogram for multi-modal classification of breast cancer using hybrid residual neural network and machine learning. *Image and Vision Computing* **145**, 104987 (2024).
21. Subaar, C. et al. Investigating the detection of breast cancer with deep transfer learning using ResNet18 and ResNet34. *Biomedical Physics & Engineering Express* **10**(3), 035029 (2024).
22. Chi, J., Chen, J.H., Wu, B., Zhao, J., Wang, K., Yu, X., Zhang, W. and Huang, Y., 2024. A Dual-Branch Cross-Modality-Attention Network for Thyroid Nodule Diagnosis Based on Ultrasound Images and Contrast-Enhanced Ultrasound Videos. *IEEE Journal of Biomedical and Health Informatics*.
23. Li, C. et al. Segmentation prompts classification: A nnUNet-based 3D transfer learning framework with ROI tokenization and cross-task attention for esophageal cancer T-stage diagnosis. *Expert Systems with Applications* **258**, 125067 (2024).
24. Rahman, M. M., Khan, M. S. I. & Babu, H. M. H. BreastMultiNet: A multi-scale feature fusion method using deep neural network to detect breast cancer. *Array* **16**, 100256 (2022).
25. Hasan, M. M. et al. FP-CNN: Fuzzy pooling-based convolutional neural network for lung ultrasound image classification with explainable AI. *Computers in Biology and Medicine* **165**, 107407 (2023).
26. Hossain, M. M. et al. Covid-19 detection from chest ct images using optimized deep features and ensemble classification. *Systems and Soft Computing* **6**, 200077 (2024).
27. Hasan, M. M., Hossain, M. M., Mia, S., Ahammad, M. S. & Rahman, M. M. A combined approach of non-subsampled contourlet transform and convolutional neural network to detect gastrointestinal polyp. *Multimedia Tools and Applications* **81**(7), 9949–9968 (2022).
28. Hossain, M. M., Rahim, M. A., Bahar, A. N. & Rahman, M. M. Automatic malaria disease detection from blood cell images using the variational quantum circuit. *Informatics in Medicine Unlocked* **26**, 100743 (2021).
29. Mostafiz, R., Uddin, M. S., Jabin, I., Hossain, M. M. & Rahman, M. M. Automatic brain tumor detection from MRI using curvelet transform and neural features. *International Journal of Ambient Computing and Intelligence (IJACI)* **13**(1), 1–18 (2022).
30. Ai, Y., Liu, J., Li, Y., Wang, F., Du, X., Jain, R.K., Lin, L. and Chen, Y.W., 2024. SAMA: A Self-and-Mutual Attention Network for Accurate Recurrence Prediction of Non-Small Cell Lung Cancer Using Genetic and CT Data. *IEEE Journal of Biomedical and Health Informatics*.
31. Nakach, F. Z., Zerouaoui, H. & Idri, A. Binary classification of multi-magnification histopathological breast cancer images using late fusion and transfer learning. *Data Technologies and Applications* **57**(5), 668–695 (2023).
32. Mukhlif, A. A., Al-Khateeb, B. & Mohammed, M. Classification of breast cancer images using new transfer learning techniques. *Iraqi Journal For Computer Science and Mathematics* **4**(1), 167–180 (2023).
33. Mukhlif, A. A., Al-Khateeb, B. & Mohammed, M. A. Incorporating a novel dual transfer learning approach for medical images. *Sensors* **23**(2), 570 (2023).
34. Ali, A. M. & Mohammed, M. A. A comprehensive review of artificial intelligence approaches in omics data processing: evaluating progress and challenges. *International Journal of Mathematics, Statistics, and Computer Science* **2**, 114–167 (2024).
35. Kamil, M. Y. Computer-aided diagnosis system for breast cancer based on the Gabor filter technique. *International Journal of Electrical and Computer Engineering (IJECE)* **10**(5), 5235–5242 (2020).
36. Xiao, J., Gan, C., Zhu, Q., Zhu, Y. & Liu, G. CFNet: Facial expression recognition via constraint fusion under multi-task joint learning network. *Applied Soft Computing* **141**, 110312 (2023).
37. Islam, O., Assaduzzaman, M. and Hasan, M.Z., 2024. An explainable AI-based blood cell classification using optimized convolutional neural network. *Journal of Pathology Informatics*, p.100389.

38. Esmaeiloghli, S., Lima, A. and Sadeghi, B., 2024. Lithium exploration targeting through robust variable selection and deep anomaly detection: An integrated application of sparse principal component analysis and stacked autoencoders. *Geochemistry*, p.126111.
39. Mehrabi, N., Haeri Boroujeni, S.P. and Pashaei, E., 2024. An efficient high-dimensional gene selection approach based on the Binary Horse Herd Optimization Algorithm for biologicaldata classification. *Iran Journal of Computer Science*, pp.1–31.
40. https://web.inf.ufpr.br/vri/databases/breast-cancer-histopathological-database-breakhis/
41. Al-Dhabyani, W., Gomaa, M., Khaled, H. & Fahmy, A. Dataset of breast ultrasound images. *Data in Brief.* **28**, 104863. https://doi.org/10.1016/j.dib.2019.104863 (2020Feb).
42. Obayya, M. et al. Hyperparameter optimizer with deep learning-based decision-support systems for histopathological breast cancer diagnosis. *Cancers* **15**(3), 885 (2023).
43. Ragab, M. et al. Optimal deep transfer learning driven computer-aided breast cancer classification using ultrasound images. *Expert Systems* **41**(4), e13515 (2024).

Author contributions
Iniyan S: Conceptualization, methodology development, experiment, formal analysis, investigation, writing. M. Senthil Raja: Formal analysis, investigation, validation, visualization, writing. R. Poonguzhali: Formal analysis, review and editing. A Vikram : Methodology, investigation. Janjhyam Venkata Naga Ramesh: Review and editing. Sachi Nandan Mohanty: Discussion, review and editing. Khasim Vali Dudekula: Conceptualization, methodology development, investigation, supervision, review and editing. All authors have read and agreed to the published version of the manuscript.

Declarations

Competing interests
The authors declare that they have no conflict of interest. The manuscript was written with the contributions of all authors, and all authors have approved the final version.

Ethics approval
This article does not contain any studies with human participants performed by any of the authors.

Consent to participate
Not applicable.

Informed consent
Not applicable.

Additional information
Correspondence and requests for materials should be addressed to K.V.D.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.