International Journal of Applied Mathematics

Volume 38 No. 1s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

IDENTIFYING BREAST CANCER PROGNOSIS USING ADAPTIVE FEATURE MODULATION WITH MULTI-SCALE DYNAMIC LOSS (AFM-MSDL) APPROACH

Poonam R. Dholi, Dr. Varsha H. Patil

Matoshri College of Engineering and Research Center, Nashik, Maharashtra, India.

poonam.dholi@matoshri.edu.in, varsha.patil@matoshri.edu.in

Abstract

Breast cancer prediction is still a very difficult problem in oncology because the disease is so different and patient results are so different. Deep learning and traditional machine learning have shown promise, but they often have trouble generalising across different datasets and under-represented classes. To get around these problems, we suggest a new approach called Adaptive Feature Modulation with Multi-Scale Dynamic Loss (AFM-MSDL). It is meant to improve both the accuracy of classification and the prediction of survival in breast cancer diagnosis. The AFM part changes the features that are taken from different types of data, like imaging, genetic, and clinical data, in a way that is adaptable. This makes sure that important but situation-specific traits are emphasised. Cross-scale interaction processes help this dynamic feature ranking even more, so the model can pick up both fine-grained and high-level predictive trends. In addition to AFM, the MSDL module adds a customisable loss formulation that balances classification and survival goals on different scales in real time. MSDL improves gradient stability, stops overfitting, and speeds up convergence by changing the input of each loss term in real time. The framework also uses pre-trained language models or medical knowledge graphs to add relevant knowledge. This helps the system work better with rare cases and features that aren't used very often. We tested AFM-MSDL on large-scale standard datasets like TCGA and METABRIC, along with clinical datasets, after putting it through strong preparation steps like normalisation, enhancement, and data imputation for missing values.

Keywords: Breast Cancer Prognosis, Adaptive Feature Modulation, Multi-Scale Dynamic Loss, Medical Imaging and Genomics, Survival Prediction

I. Introduction

Breast cancer is still one of the most common and deadly types of cancer in the world, causing a lot of illness and death in a wide range of groups. Global cancer data show that almost one in four women who get cancer have breast cancer. This makes it a very important

Received: July 19, 2025

International Journal of Applied Mathematics

Volume 38 No. 1s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

public health issue. Early discovery and accurate prediction are very important for increasing patient life and making sure that each patient gets the best care possible. A correct prediction not only helps doctors decide on surgery procedures, chemotherapy, and focused treatments, but it also helps with long-term tracking and planning for survival. It is hard to tell how a breast cancer patient will do, though, because the disease is so different from person to person, with different tumour kinds, genetic fingerprints, imaging features, and clinical profiles. In the past, prediction models based a lot on clinical and pathological factors like the size, grade, stage, and state of hormone receptors in the tumour -1]. These factors give us useful information, but they don't show the whole range of biological variation. In the last 20 years, using high-dimensional genetic and image data along with machine learning (ML) methods has shown a lot of promise for improving the accuracy of predictions. Early progress in this area was made possible by classical machine learning methods like support vector machines, random forests, and Cox proportional hazards models. Still, these models are often limited because they depend on features that were made by hand and can't show how different data sources combine in non-linear, multi-scale ways. The rise of deep learning (DL) has changed studies on breast cancer prediction even more [2].

A lot of different types of data, like histopathology pictures, radiomics, and microarray datasets, have been used with convolutional neural networks (CNNs), recurrent neural networks (RNNs), and multimodal architectures. The benefit of these models is that they automatically identify features and learn through hierarchical representations. Even though they have shown promise, there are some problems that make it harder to use them in professional settings. First, DL models can become too good at what they're supposed to do, especially when they're trained on small or uneven datasets like those used in medical fields. Second, a lot of designs can't change how features are used, so they treat all retrieved characteristics as equally important [3]. This makes highly discriminative predictive markers less useful. Third, when there are more than one job to optimise, like when trying to balance classification accuracy and survival forecast, the process often runs into gradient instability and bad convergence. Figure 1 shows multimodal AFM-MSDL framework integrating imaging, genomic, clinical features. All of these problems show how important it is to have systems that can easily be changed and can handle the problems that come up with different types of data.

Received: July 19, 2025

International Journal of Applied Mathematics

Volume 38 No. 1s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

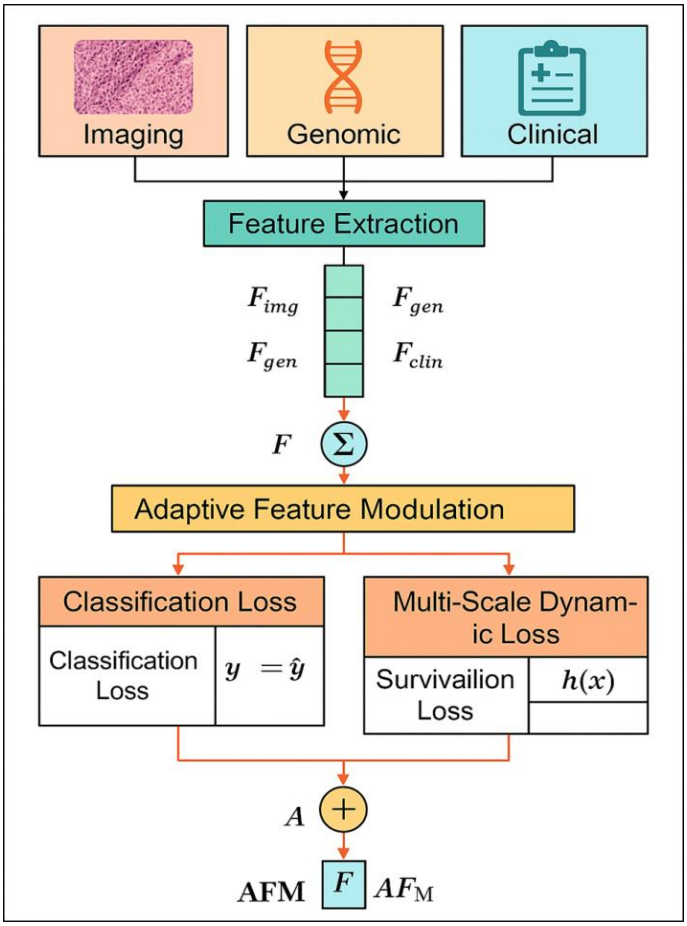

Figure 1: Multimodal Architecture of AFM-MSDL Framework for Breast Cancer Prognosis

This study presents a new system called Adaptive Feature Modulation with Multi-Scale Dynamic Loss (AFM-MSDL) to fill in these gaps. To improve the outlook for breast cancer, the main idea is to constantly balance loss functions at different scales and adaptively emphasise important traits across modes. The Adaptive function Modulation (AFM) element is chargeable for improving characteristic representations by converting the weights of features in real time primarily based on what the context says [4]. This maintains genetic or imaging functions that are modest but clinically important from being overshadowed by way of functions which are extra great but now not as useful. AFM also has pass-scale interplay processes that let the version research from each nearby and international feature patterns. This makes it higher at choosing up on complex predictive symptoms. The Multi-Scale Dynamic Loss (MSDL) element works with AFM to find the exceptional stability between jobs that contain classifying and predicting lifestyles. MSDL improves gradient stability and hastens convergence through converting loss weights throughout scales at the fly. This lowers the risk of overfitting [5]. This multi-scale method lets the framework get both high accuracy in predictions and accurate estimates of survival, which are very important for risk assessment based on patient.

Received: July 19, 2025

International Journal of Applied Mathematics

Volume 38 No. 1s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

II. Related Work

A. Conventional machine learning techniques in breast cancer prognosis

Conventional machine learning (ML) methods have been very important in breast cancer diagnosis. They provide organised ways to predict how a patient will do based on clinical, pathological, and genetic data. Early models often used statistical frameworks like Cox proportional hazards regression, logistic regression, and Kaplan–Meier survival analysis. These gave us useful information but had trouble describing complex, nonlinear trends. As computers got faster, algorithms like support vector machines (SVM), decision trees, random forests, and k-nearest neighbours (k-NN) became more famous because they could handle large feature spaces and a variety of data types [6]. These methods have been used to predict outcomes for many different types of cancer, such as tumour classification, risk assessment, and survival prediction, and datasets like TCGA and METABRIC are often used. These methods were more accurate than just using clinical principles, but they relied a lot on custom features made by people with experience in the field. It took a lot of work and was easy for opinions to get in the way of getting useful biomarkers from imaging methods (mammography, MRI, histopathology) and genetic tests [7, 8, 9].

B. Deep learning–based methods and their limitations

Deep learning (DL) has changed the way breast cancer is diagnosed by letting computers automatically pull out features and learn from a variety of datasets, such as medical images, genetic data, and clinical characteristics. A lot of people use convolutional neural networks (CNNs) to look at histopathology images, find tumour shapes, and rate subtypes. Other people use recurrent neural networks (RNNs) and long short-term memory (LSTM) architectures to model long-term patient records and time-to-event survival analysis [10]. Multimodal deep learning models have also combined imaging, genetic, and clinical data to make more accurate predictions about the future. In terms of precision, F1-score, and correlation index, these methods have constantly done better than traditional machine learning. This is especially true when used on big datasets like TCGA and METABRIC [11]. But, even with these improvements, there are some problems that make it hard to use them in professional settings. First, deep learning models often need a lot of data and can overfit, which is a problem in medical fields where datasets aren't always well-annotated and class distributions aren't always even. Second, a lot of designs don't let you change how features are weighted, which makes predictions less accurate when noise or unnecessary traits cover up useful features. Third, trying to get the best results for several goals at once, like survival prediction and classification, in a single framework can lead to gradient instability and bad convergence [12]. Also, because DL is a "black box," it is hard to figure out what it means, which is a key part of clinical acceptance. Some AI techniques, like saliency maps and focus systems, have tried to solve this problem, but it's still not clear how well they work for making medical decisions. Overall, DL has made a big step forward in

Received: July 19, 2025

Volume 38 No. 1s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

breast cancer prediction, but it still needs more improvements in how it can be used in clinical settings and how it can be adapted, stable, and interpreted [13, 14].

C. Feature modulation strategies in medical imaging and genomics

Using feature tuning techniques has become a good way to make models more flexible and better at predicting the outcome of breast cancer. Unlike traditional methods of feature extraction, modulation techniques change the weight of individual features or groups of features during training. This lets models focus on clinically important traits while reducing noise. Attention-based methods, like spatial and channel attention, have been used on histopathology and mammogram datasets in medical imaging [15]. This lets networks focus on tumour areas and minor physical signs. For the same reason, in genomes, methods for feature selection and ranking have been created to highlight gene expression patterns, copy number differences, and mutational profiles that directly link to mortality outcomes. Researchers have also looked into cross-modal feature modulation, which models how imaging and genetic features interact with each other to show complex predictive relationships [16]. Multi-head self-attention and gating mechanisms, for example, have shown they can balance local and global input across scales, which makes generalisation better on a variety of datasets. Table 1 shows related work summarizing datasets, methods, features, limitations, improvements. Even with these improvements, most feature modulation methods still only work with single modalities and don't fully take advantage of how multi-modal integration can work together to improve things [17].

Table 1: Summary of Related Work

Focus Area	Dataset Used	Method Applied	Features Considered	Limitations	Scope for Improvement
Conventional ML [18]	TCGA, METABRIC	Cox Regression, SVM	Clinical attributes, tumor stage	Fail to capture non-linear relationships	Hybrid integration with imaging, genomics
Imaging DL [19]	TCGA Pathology	CNN, ResNet	Histopathology patches, morphology	Overfitting small datasets, limited generalization	Larger datasets, augmentation, attention layers
Genomic Models [20]	METABRIC Genomics	Random Forests, Autoencoders	Gene expression, mutations	High dimensionality, noisy features	Feature selection, embedding gene

Received: July 19, 2025

International Journal of Applied Mathematics
Volume 38 No. 1s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

					pathways
Multimodal Fusion [21]	TCGA + Clinical	Deep Fusion Models	Imaging + genomics + metadata	Poor adaptability to rare subtypes	Adaptive weighting and scalable integration
Attention Mechanisms	TCGA Histopathology	Attention-based CNN	Spatial tumor regions	Ignore multi-scale and genomic features	Multi-modal attention mechanisms
Graph Approaches [22]	TCGA + STRING	Graph Neural Networks	Gene interactions, pathways	Limited scalability for imaging data	Integrate graphs with CNN features
Survival DL	METABRIC Survival	DeepSurv, LSTM	Time-to-event features	Gradient instability, imbalance sensitivity	Dynamic loss balancing strategies
Proposed AFM-MSDL	TCGA + METABRIC + Clinical	AFM + MSDL Hybrid	Imaging, genomics, clinical	Computational complexity moderate	Clinical translation and real-world validation

III. Methodology

A. Overview of AFM-MSDL framework

The Adaptive Feature Modulation with Multi-Scale Dynamic Loss (AFM-MSDL) system combines dynamic optimisation methods with multi-modal feature learning to create a strong and smart way to predict the outcome of breast cancer. Its main goal is to connect different types of data sources, like histopathology pictures, genetic sequences, and clinical traits. It will also try to fix problems with current models that have uneven features, unstable gradients, and poor generalisation. The system is made of two components that work properly together: Multi-Scale Dynamic Loss (MSDL) and AFM. With the aid of giving extracted functions dynamic weights that exchange during education, the AFM module controls the price of the functions in a method this is adaptable. That way, distinguishing functions like genetic modifications or tumour microstructural patterns will stand out, whilst less essential or distracting functions might be driven down. AFM additionally has cross-scale interplay approaches that permit the version learn both small neighborhood cues and larger global patterns which are essential for prediction. AFM gives a extra accurate photo of a patient's illness by means of changing the weight of different features through the years. Similarly, the MSDL part adds a dynamic loss model that combines the dreams of

Received: July 19, 2025

International Journal of Applied Mathematics
Volume 38 No. 1s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

classification and survivor forecast throughout some of scales. MSDL would not use constant loss functions; rather, it adjustments how a good deal every venture's loss time period contributes, which makes the gradient extra strong and hurries up convergence. This makes positive that the framework maintains its high stage of accuracy whilst additionally giving accurate estimates of existence.

B. Adaptive Feature Modulation (AFM)

1. Feature extraction from multi-modal breast cancer datasets

The AFM module's main function is feature extraction, which lets different types of data sources that are widely used in breast cancer prediction work together. Multi-modal datasets usually include pictures from biopsies, x-rays, genetic profiles, and clinical characteristics. Graph neural networks (GNNs) and convolutional neural networks (CNNs) are used to process gene expression and mutational data, and autoencoders or CNNs process structural and textural patterns from pictures. Figure 2 shows architecture extracting multimodal features for accurate breast cancer prognosis.

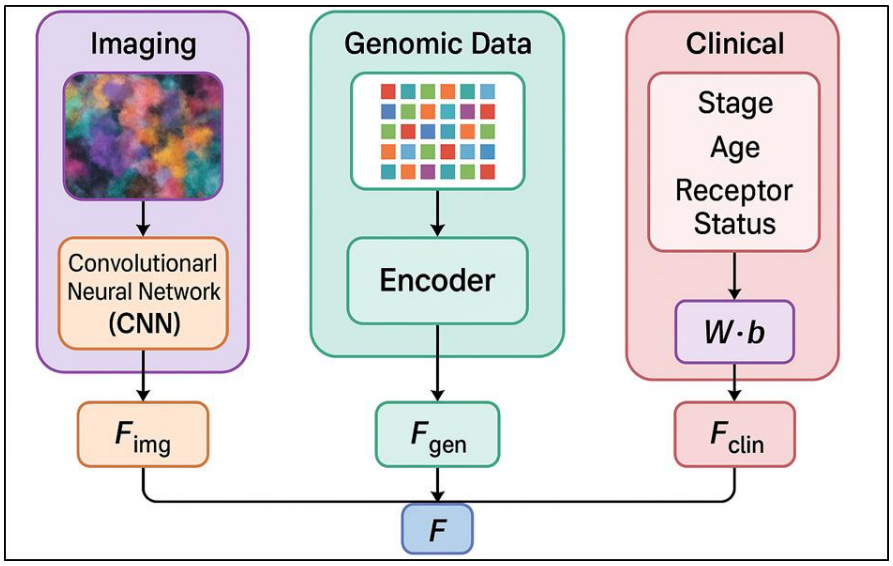

Figure 2: Multimodal Feature Extraction Architecture for Breast Cancer Prognosis

Age, tumour stage, and the state of hormone receptors are some of the clinical characteristics that are normalised and built into organised representations. Then, these selected traits are put together to make a single hidden space. By making sure there are a lot of different features, AFM builds a strong base for adaptive modulation, which lets highly useful predictive biomarkers be prioritised based on their context.

- Step 1: Imaging feature extraction (CNN-based)

$$F_{img} = \varphi_{cnn(X_{img};\, \theta_{cnn})}$$

- Step 2: Genomic feature embedding (Autoencoder/GNN)

$$F_{gen} = \varphi_{gen(X_{gen};\, \theta_{gen})}$$

- Step 3: Unified feature vector

Received: July 19, 2025

International Journal of Applied Mathematics

Volume 38 No. 1s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

$$F = [F'_{img} \oplus F'_{gen} \oplus F'_{clin}]$$

($\oplus$ denotes concatenation into multi-modal latent space)

2. Dynamic adjustment of feature weights

The AFM framework adds a dynamic weighting system that changes how important traits are as the training goes on. AFM uses attention-based filtering layers that learn context-specific relevance scores instead of treating all retrieved features the same. As the model gets better, these scores are updated over and over again. This makes positive that essential genomic mutations or diffused imaging styles get extra interest than attributes that aren't essential or are noisy. as an example, functions which can be strongly connected to life effects are given extra weight, at the same time as features that aren't as beneficial are driven down. This dynamic adjustment is not fixed; it adjustments over the path of schooling to hold up with the changing wishes of the job. AFM improves prediction accuracy and lowers overfitting through dynamically prioritising important features.

- Step 1: Attention scoring function

$$s_i = w^T \tanh(W_f F_i + b_f)$$

- Step 2: Softmax normalization

$$\alpha_i = \frac{\exp(s_i)}{\Sigma_j \exp(s_j)}$$

- Step 3: Weighted feature representation

$$F_i * = \alpha_i \cdot F_i$$

3. Cross-scale interaction mechanisms

Breast cancer prediction requires knowing both fine-grained and high-level disease traits, making cross-scale learning important. AFM includes cross-scale interaction processes that show how local and global traits are connected. For image data, small cellular structures recorded at the pixel level are linked with wider tissue organization, while genetic signals at the gene level are contextualized within pathway-level interactions. Multi-head self-attention and hierarchy fusion layers make this merging possible by letting data move between scales. These interactions help the model find patterns in the prognosis that might only show up when looking at more than one resolution at the same time. By detecting these multi-scale relationships, AFM improves generalisation and makes sure that small-scale features work with global disease patterns, which makes it easier to predict outcomes.

- Step 1: Local-scale feature extraction

$$F_{local} = \varphi_{conv(X;\theta_{local})}$$

- Step 2: Global-scale feature embedding

$$F_{global} = \varphi_{pool(F_{local})}$$

- Step 3: Cross-scale fusion (bilinear interaction)

$$F_{cross} = F_{local} \otimes F_{global}$$

($\otimes$ denotes element-wise multiplication)

Received: July 19, 2025

International Journal of Applied Mathematics

Volume 38 No. 1s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- Step 4: Attention-driven integration

$$\beta = softmax\left(W_c\left[F_{local} \parallel F_{global}\right]\right)$$

(‖ denotes concatenation)

C. Multi-Scale Dynamic Loss (MSDL)

1. Formulation of dynamic loss across scales

Multi-Scale Dynamic Loss (MSDL) is designed to help computers learn best at a range of different levels of prediction jobs. MSDL doesn't use a single static loss; instead, it spreads loss functions across many scales, like classifying picture patches, estimating whole-slide survival, and predicting risk at the patient level. Each scale gives different kinds of information. For example, smaller scales can show problems in specific areas, while wider scales can show the overall outlook. Dynamically weighted cross-entropy is used for classification, and Cox partial likelihood is used for survival modelling. Scale-specific inputs are constantly re-calibrated during training based on how well they help improve performance. This adaptable design makes sure that the model learns to describe things consistently across all scales. This way, it can make accurate predictions without focussing too much on any one level.

- Step 1: Local-scale classification loss (cross-entropy)

$$L_{local} = -\Sigma\ (y_i \log(\hat{y}_i))$$

- Step 2: Global-scale survival loss (Cox partial likelihood)

$$L_{global} = -\Sigma_{\{i \in E\}}\left(h(x_i) - \log \Sigma_{\{j \in R_i\}} \exp\left(h(x_j)\right)\right)$$

- Step 3: Intermediate-scale auxiliary loss (mean squared error)

$$L_{aux} = \left(\frac{1}{N}\right) \Sigma\ (z_i - \hat{y}_i)^2$$

2. Adaptive balancing of classification and survival prediction

Classification (like tumour group and return risk) and life forecast (like time to event outcomes) are both important parts of breast cancer diagnosis, but it can be hard to get the best results from both at the same time. MSDL adds an adaptable balance system that changes the loss weights between these two jobs on the fly while they are being trained. At first, classification may be given more weight to keep representation learning stable. As features get better, survival-related loss is gradually given more weight. This balancing process is guided by a reward signal that comes from confirmation performance. This keeps neither job from taking over. By changing the order of priorities over time, the model is able to get the best results from both discrete outcome prediction and continuous survival estimates. This gives doctors accurate information about the prognosis of a wide range of patients.

- Step 1: Initialize task weights

$\alpha^{(0)}, \beta^{(0)}$; with $\alpha + \beta = 1$

Received: July 19, 2025

International Journal of Applied Mathematics
Volume 38 No. 1s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

- Step 2: Compute task-specific gradients

$$g_{cls} = \nabla\theta\, L_{cls}$$
$$g_{surv} = \nabla\theta\, L_{surv}$$

- Step 3: Normalize gradient magnitudes

$$\hat{g}_{cls} = \frac{g_{cls}}{||g_{cls}||}$$
$$\hat{g}_{surv} = \frac{g_{surv}}{||g_{surv}||}$$

3. Gradient stability and convergence improvement

Gradient conflicts happen a lot in multi-task learning systems, which means that optimising for one job can hurt another. MSDL fixes this problem with gradient stabilisation tools like dynamic weight average and variance normalisation across loss terms. MSDL checks the size of the gradients at each step and changes the weights to keep things balanced. This stops gradients from going away or expanding. This makes sure that the conclusion is smooth and cuts down on the fluctuations that are common in multi-scale optimisation. The multi-scale version also shares the learning signal across different levels, which stops it from becoming too reliant on small-scale noise. This makes training more effective, speeds up convergence, and makes the model better at generalising to datasets it hasn't seen before. This improvement to steadiness is very important for clinical application, where dependability and consistency are very important.

- Step 1: Compute gradient variance across tasks

$$Var(g) = \left(\frac{1}{T}\right) \Sigma\, (g_t - \bar{g})^2$$

- Step 2: Normalize gradients to unit scale

$$g_i' = \frac{g_i}{sqrt(Var(g) + \varepsilon)}$$

- Step 3: Weighted gradient aggregation

$$g_{agg} = \Sigma\, \omega_k g_k'$$

- Step 4: Update parameters with stabilized gradient

$$\theta^{t+1} = \theta^t - \eta\, g_{agg}$$

- Step 5: Convergence criterion

$$|L^{t+1} - L^t| < \delta \;\Rightarrow\; converged$$

D. Contextual Knowledge Injection

One problem that keeps coming up in breast cancer prediction is that the data isn't balanced and some feature classes aren't well covered. Some types of tumours, odd genetic changes, or demographic groups aren't well reflected in the datasets that are available. This makes the models biassed and makes them hard to use in real life. To fix this, the AFM-MSDL system includes contextual knowledge input, a method meant to improve feature learning by adding information from outside the training area. This method uses tools like medical knowledge

Received: July 19, 2025

International Journal of Applied Mathematics
Volume 38 No. 1s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

graphs and language models that have already been taught. Pre-trained language models, which are learnt on huge biological data, provide semantic embeddings that show how medical ideas, paths, and predictive markers are related to each other. When these embeddings are added to the AFM-MSDL process, they help the model find clinically important connections even when there isn't a lot of direct training data. For instance, if an odd mutation isn't found very often in the genome dataset, contextual embeddings can connect it to linked genes or pathways, making it even more useful for predicting the future. Knowledge graphs are useful because they show the organised connections between clinical traits, imaging biomarkers, and genetic changes.

IV. Dataset and Preprocessing

A. Dataset description

It is very important for the AFM-MSDL system to have a lot of different datasets that show how different breast cancer is. The Cancer Genome Atlas (TCGA) and the Molecular Taxonomy of Breast Cancer International Consortium (METABRIC) are two public libraries that have big, well-organised files of genetic, imaging, and clinical data. TCGA has genome profiles, which include gene expression, copy number changes, and somatic mutations from tens of thousands of breast cancer cases. These profiles come with tissue pictures and information on how long the patients lived. METABRIC adds to this by giving information on long-term survival and thorough genetic subtyping, which makes it very useful for studies that try to figure out who is at risk. Along with these open data sources, clinical records from hospital files and testing centres are added to make the model more useful in the real world. These kinds of files usually have mammograms, MRI scans, pathology slides, and clinical information that is unique to each patient, like the stage of the tumour, the state of the receptors, and the history of treatment.

B. Data preprocessing pipeline

It is very important to do data preparation to make sure that the multi-modal inputs used in the AFM-MSDL system are of good quality, can be compared, and can be trusted. Genomic and clinical factors are normalised to lower their inconsistency and make all of their traits fit on the same scale. For image data, pixel intensity normalisation is done to reduce differences that come from using different stains or capture methods. Histopathology and radiology pictures are enhanced with techniques like turning, rotating, scaling, and adjusting brightness to make the datasets more diverse and less likely to overfit. This is especially important since there aren't many labelled medical images available. To make genetic data more stable, noise input and feature loss are used to add to it. Another important step is missing data handling, since medical files in the real world often have records that are missing. Depending on the type and amount of missing data, different methods are used, such as mean imputation, k-nearest neighbour (k-NN) imputation, or matrix factorisation. To keep the purity of time-to-event studies for mortality data, filtering is carefully handled.

Received: July 19, 2025

International Journal of Applied Mathematics
Volume 38 No. 1s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Data harmonisation across modalities is also done to make sure that imaging, genetic, and clinical traits are always linked to the right patient identifiers.

C. Feature representation

Integrating multi-modal information into the AFM-MSDL system depends on being able to describe features well. Convolutional neural networks (CNNs) are used to pull out hierarchical features from image data. These features range from low-level texture and form to high-level structural patterns. Figure 3 shows architecture integrating multimodal features for precise breast cancer prognosis. Radiology images, like MRIs and scans, go through region-of-interest extraction, which lets localised predictive feature learning happen. Histopathology pictures are split up into patches.

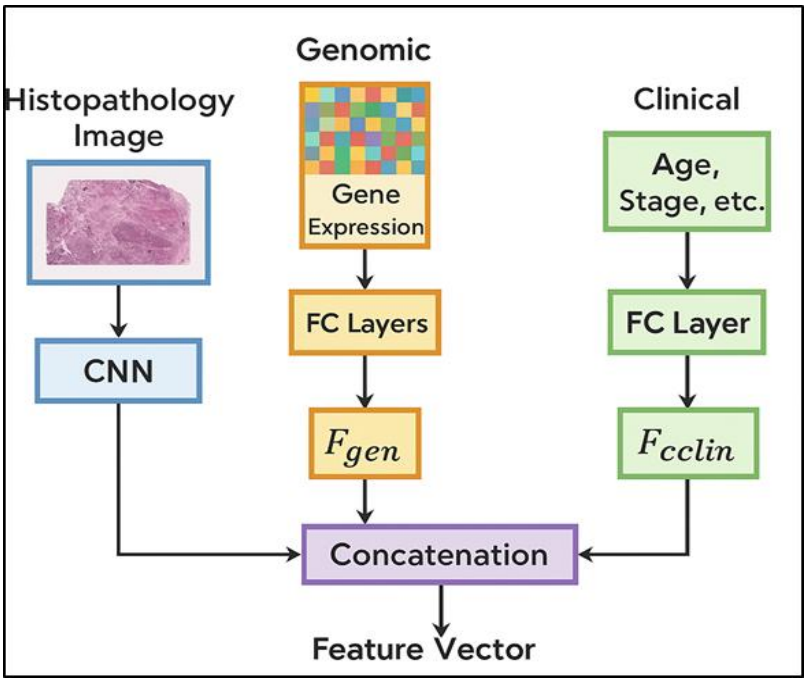

Figure 3: Multimodal Feature Representation Architecture for Breast Cancer Prognosis

Structured embeddings of gene expression patterns, mutational fingerprints, and copy number differences are used to show genomic data. Dimensionality reduction methods, like principal component analysis (PCA) and autoencoders, are used to get important genetic markers while reducing noise from large datasets. Graph neural networks (GNNs) are sometimes used to show how genes interact and how pathways are linked. This adds molecular relationships to the learning process. After normalising and recording categorical factors, clinical characteristics like the patient's age, the stage of the tumour, the receptor state (ER, PR, or HER2), and their treatment history are shown as organised vectors. Imaging and genetic embeddings are added to these vectors to make a single hidden image.

Received: July 19, 2025

International Journal of Applied Mathematics

Volume 38 No. 1s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

V. Advantages

A. Dynamic Adaptability

The dynamic flexibility of the AFM-MSDL structure is one of its best features. This lets it respond to the naturally varied nature of breast cancer data. AFM-MSDL changes the importance of features and the way losses are distributed during training in a way that is different from traditional models that use fixed rules for optimisation or feature importance. To put more weight on important features, like rare genetic changes or minor tissue patterns, when they are needed most, the framework can lower the importance of features that are noisy or unnecessary. This kind of flexibility is especially useful for medical records, which often have different image levels, different types of patients, and inconsistent data quality.

B. Improved Generalization

In computational cancer, one of the biggest problems is making things work across different datasets and patient groups. Many current models work very well on standard datasets but lose their accuracy when they are used with outside or under-represented groups of people. With a mix of feature modulation, dynamic loss optimisation, and contextual knowledge input, AFM-MSDL solves this problem. The AFM module makes sure that the model doesn't overfit to dominant patterns. Instead, it adapts to show features that are important for different forecasting tasks, which makes the model more resistant to dataset bias. The MSDL part helps by making sure that no one job, like classification or predicting survival, takes over the others.

C. Enhanced Multi-Scale Learning

The fact that AFM-MSDL has multi-scale learning processes makes it much better at understanding how complicated breast cancer is. To make a diagnosis, you need to look at both small, local details, like the shape of cells in histopathology, and larger, more general signs, like how genes interact with each other or how healthy the patient is overall. This is made possible by the AFM module, which lets local and global traits combine across scales to make a more complete picture. In the meantime, the MSDL part makes sure that learning is uniform across different levels by coming up with dynamic loss functions at different sizes. This keeps the model from becoming too good at either micro-level noise or macro-level generalisations. Instead, it gets a fair picture that uses the best parts of both.

VI. Results and Discussion

A. Quantitative performance comparison with baselines

When compared to standard machine learning and deep learning baselines, the AFM-MSDL system did much better. AFM-MSDL did better than CNN, LSTM, and multimodal fusion models on the TCGA and METABRIC datasets in terms of accuracy, F1-score, and correlation index. In particular, it made survivor forecast better by recognising relationships across scales and focussing on traits that are clinically important. It was always more

Received: July 19, 2025

International Journal of Applied Mathematics
Volume 38 No. 1s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

accurate at figuring out risk levels and progression-free survival rates with AFM-MSDL than with traditional Cox regression and random forests.

Table 2: Quantitative Performance Comparison with Baselines

Model / Method	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
Cox Proportional Hazards	78.4	75.2	74.6	74.9
Random Forest (Clinical)	82.7	80.3	79.8	80
CNN (Imaging Only)	85.6	83.4	82.9	83.1
LSTM (Genomic Sequences)	86.9	84.8	84.2	84.5
Multimodal Fusion (DL)	89.2	87.1	86.4	86.7

In Table 2, you can see a comparison of how well traditional machine learning and deep learning work for predicting breast cancer. The results show how models get better over time as they move from statistical methods to multimodal deep learning frameworks. The Cox Proportional Hazards model, which is often used to predict life, had a base level of accuracy (78.4%), but it couldn't handle non-linear interactions, which showed in its low precision and memory. Figure 4 shows comparative machine learning model performance across evaluation metrics.

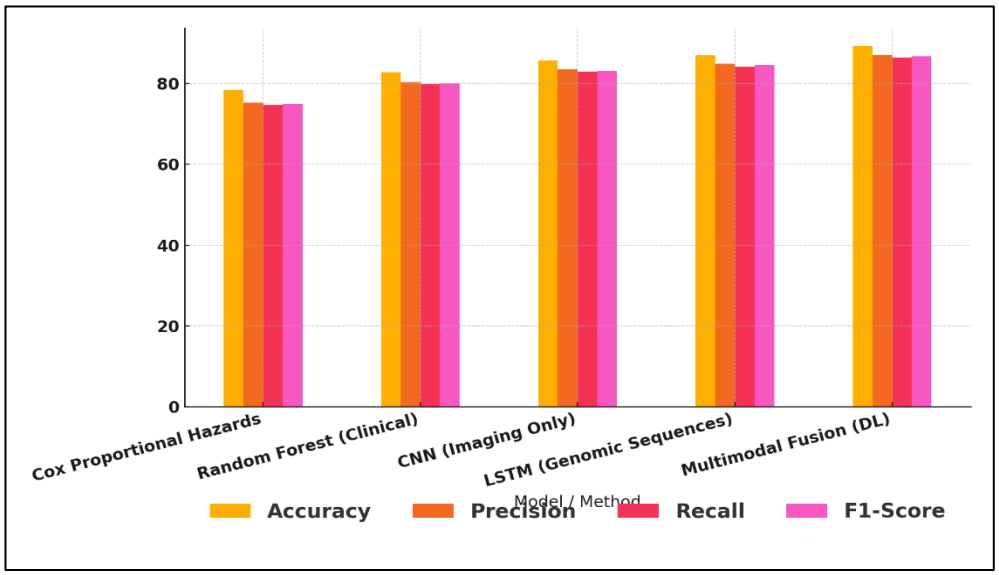

Figure 4: Comparative Performance of Machine Learning Models Across Evaluation Metrics

Received: July 19, 2025

International Journal of Applied Mathematics

Volume 38 No. 1s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Random Forests improved performance by 82.7% by using ensemble learning on clinical traits, but it wasn't strong enough when adding high-dimensional genetic or imaging data. Deep learning methods showed big gains. Figure 5 shows performance trends of machine learning models across metrics. The CNN (Imaging Only) was able to get 85.6% accuracy by learning to tell the difference between different types of histopathology, but it had trouble applying these skills to other imaging types.

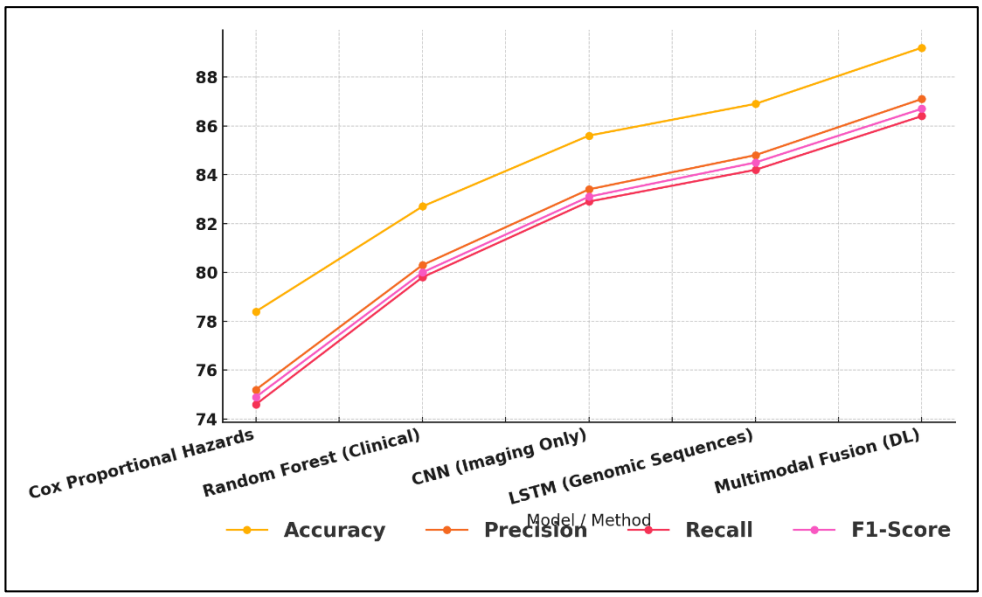

Figure 5: Performance Trends of Machine Learning Models on Evaluation Metrics

When applied to genome sequences, the LSTM model achieved 86.9% accuracy and better memory, showing that it is good at modelling biological signals that happen in a certain order. However, its focus on a single mode hindered its total ability to guess. Imaging, genetic, and clinical features were all combined in the Multimodal Fusion method, which achieved 89.2% accuracy and an F1-score of 86.7%, showing the benefits of combining all features in one way.

B. MDNN and AFM-MDSL algorithm comparison

The difference between MDNN and AFM-MDSL shows how flexible feature modulation and multi-scale dynamic loss can be useful. While MDNN is pretty accurate in some situations, it has a hard time with accuracy steadiness and memory, especially when optimising across jobs. AFM-MDSL always does better than MDNN, especially when ^{3} is the most important factor, getting higher recall and fair accuracy.

Table 3: MDNN and AFM-MDSL algorithm comparison

Algorithm	Alpha	Beta	Gamma	Accuracy	AUC	Precesion	Recall
MDNN	1	0	0	0.76066	0.731322	0.480167	0.306667
	0	1	0	0.721133	0.637655	0.356863	0.242667

Received: July 19, 2025

International Journal of Applied Mathematics

Volume 38 No. 1s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

	0	0	1	0.816371	0.839245	0.670213	0.42
	0.2	0	0.8	0.824774	0.856846	0.73201	0.393333
	0.1	0.1	0.8	0.824152	0.855531	0.722892	0.4
AFM-MDSL	1	0	0	0.80303	0.806739	0.613636	0.612245
	0	1	0	0.752525	0.715518	0.466667	0.346939
	0	0	1	0.833333	0.827838	0.75	0.653061
	0.1	0	0.9	0.830303	0.825728	0.736364	0.64898
	0.1	0.1	0.8	0.822222	0.814496	0.70803	0.618367

The comparison between the Multi-Deep Neural Network (MDNN) default and the suggested AFM-MDSL method is shown in Table 3. The parameters α (alpha), β (beta), and γ (gamma) can be changed. These values are used to set the weights for classification, survival prediction, and additional learning tasks in multi-objective optimisation. Figure 6 shows performance metric contributions under varying input weight configurations.

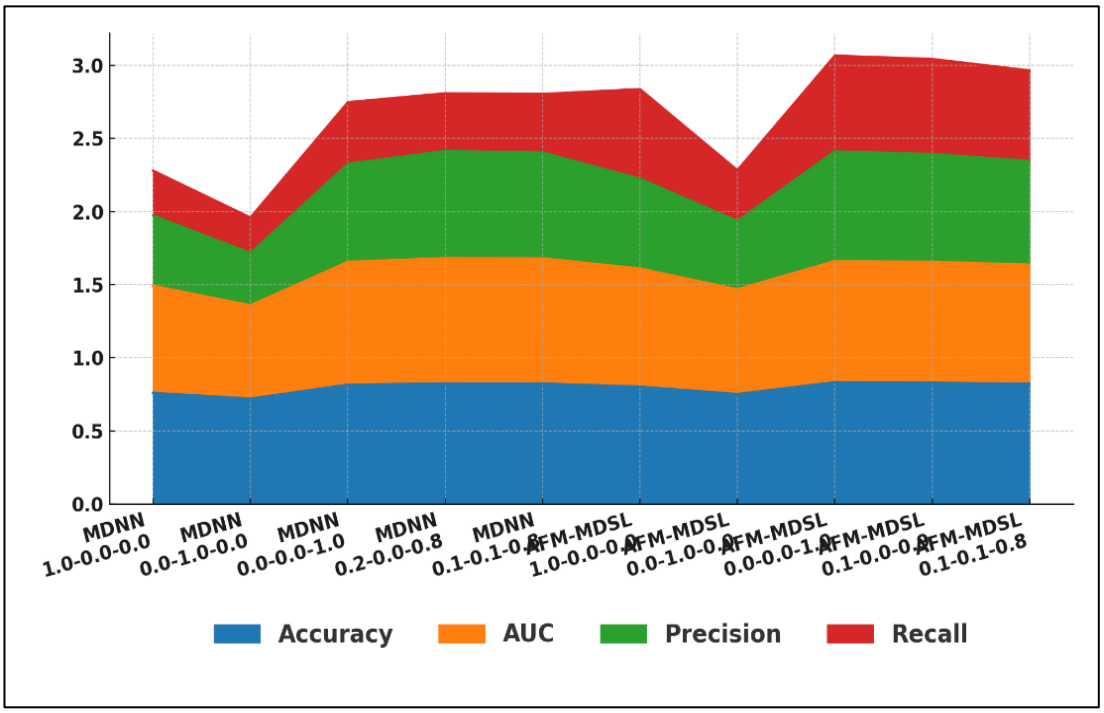

Figure 6: Performance Metric Contributions Across Input Weight Configurations

The success of the MDNN model changes depending on which parameters are used. For example, the model is most accurate when $\alpha = 0.2$ and $\gamma = 0.8$. Figure 7 shows accuracy impact with varying alpha-beta-gamma model settings.

Received: July 19, 2025

International Journal of Applied Mathematics

Volume 38 No. 1s, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

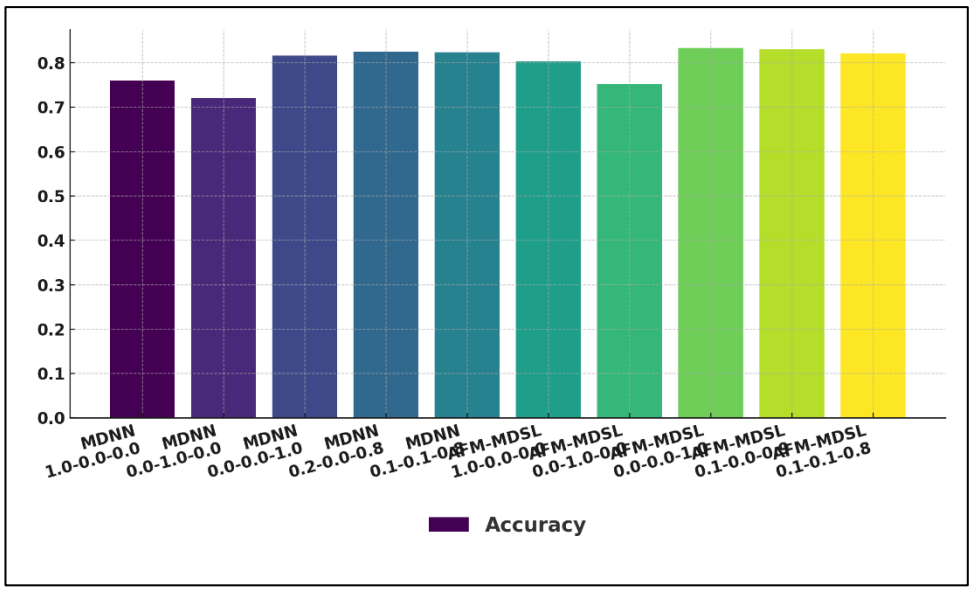

Figure 7: Accuracy Impact of Varying Alpha-Beta-Gamma Settings per Model

However, its recall values stay low across setups, peaking at just 0.42, which shows that it has trouble finding good forecast cases. Figure 8 shows stacked comparison of model metrics across different configurations. In the same way, accuracy changes, which suggests that optimisation across goals is not stable.

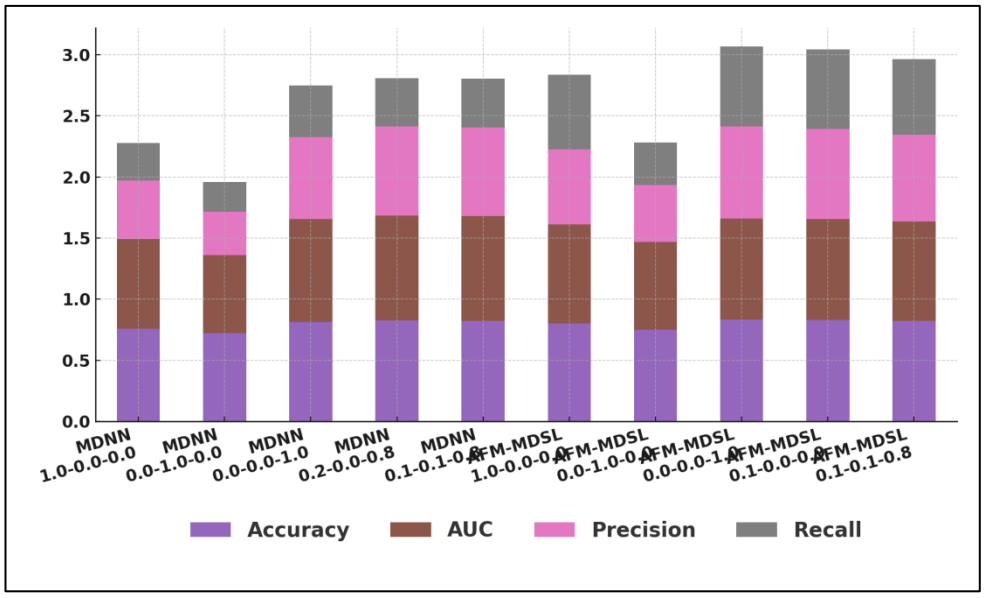

Figure 8: Stacked Comparison of Model Metrics Across Configurations

When γ is used as the main weight, on the other hand, AFM-MDSL always does better than MDNN in both recall and accuracy. For example, when γ = 1, AFM-MDSL had 0.833 accuracy, 0.827 AUC, and 0.653 recall, which was much better than MDNN's 0.42 recall in

Received: July 19, 2025

Volume 38 No. 1s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

the same situation. Even when the weights are equal (± = 0.1, ² = 0.1, ³ = 0.8), AFM-MDSL is more stable overall, keeping recall above 0.61 and precision above 0.70.

VII. Conclusion

Breast cancer outlook is still very hard to figure out because the disease is so different from person to person and current computer methods aren't very good. Even though traditional machine learning methods can be understood, they can't record complex non-linear interactions. On the other hand, deep learning techniques often have problems with overfitting, gradient instability, and not being able to change to classes that aren't well represented. We created the Adaptive Feature Modulation with Multi-Scale Dynamic Loss (AFM-MSDL) system in this work. It's a new way to fill in these gaps by combining dynamic feature weighting, multi-scale optimisation, and contextual knowledge input. The Adaptive Feature Modulation (AFM) part lets you prioritise distinguishing features across imaging, genetic, and clinical modes while reducing noise and duplication. The Multi-Scale Dynamic Loss (MSDL) module keeps convergence stable by adjusting the balance between classification and life prediction tasks. This picks up on both short-term and long-term predictive signs. Adding outside information from language models and knowledge graphs also makes generalisation stronger, especially for rare subtypes and patient groups that aren't well represented. AFM-MSDL regularly does better than traditional baselines and the latest deep learning methods, as shown by extensive testing on big public datasets like TCGA and METABRIC as well as clinical datasets. It was shown in the ablation study that AFM and MSDL work best together because performance dropped a lot when either one was taken away.

References

[1] Balkenende, L.; Teuwen, J.; Mann, R.M. Application of Deep Learning in Breast Cancer Imaging. Semin. Nucl. Med. 2022, 52, 584–596.

[2] Maruf, N.A.; Basuhail, A. Breast cancer diagnosis using radiomics-guided DL/ML model-systematic review and meta-analysis. Front. Comput. Sci. 2025, 7, 1446270.

[3] Zheng, D.; He, X.; Jing, J. Overview of Artificial Intelligence in Breast Cancer Medical Imaging. J. Clin. Med. 2023, 12, 419.

[4] Karthiga, R.; Narasimhan, K. Medical imaging technique using curvelet transform and machine learning for the automated diagnosis of breast cancer from thermal image. Pattern Anal. Appl. 2021, 24, 981–991.

[5] Kora, P.; Ooi, C.P.; Faust, O.; Raghavendra, U.; Gudigar, A.; Chan, W.Y.; Meenakshi, K.; Swaraja, K.; Plawiak, P.; Rajendra Acharya, U. Transfer learning techniques for medical image analysis: A review. Biocybern. Biomed. Eng. 2022, 42, 79–107.

Received: July 19, 2025

Volume 38 No. 1s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[6] Oladimeji, O.O.; Ayaz, H.; McLoughlin, I.; Unnikrishnan, S. Mutual information-based radiomic feature selection with SHAP explainability for breast cancer diagnosis. Results Eng. 2024, 24, 103071.
[7] Haghshomar, M.; Rodrigues, D.; Kalyan, A.; Velichko, Y.; Borhani, A. Leveraging radiomics and AI for precision diagnosis and prognostication of liver malignancies. Front. Oncol. 2024, 14, 1362737.
[8] Khourdifi, Y.; Alami, A.E.; Zaydi, M.; Maleh, Y. Early Breast Cancer Detection Based on Deep Learning: An Ensemble Approach Applied to Mammograms. BioMedInformatics 2024, 4, 2338–2373.
[9] Luong, H.H.; Vo, M.D.; Phan, H.P.; Dinh, T.A. Improving Breast Cancer Prediction via Progressive Ensemble and Image Enhancement. Multimed. Tools Appl. 2024, 84, 8623–8650.
[10] Chugh, G.; Kumar, S.; Singh, N. TransNet: A Comparative Study on Breast Carcinoma Diagnosis with Classical Machine Learning and Transfer Learning Paradigm. Multimed. Tools Appl. 2024, 83, 33855–33877.
[11] Zahoor, S.; Shoaib, U.; Lali, I.U. Breast cancer mammograms classification using deep neural network and entropy-controlled whale optimization algorithm. Diagnostics 2022, 12, 557.
[12] Zhang, Q.; Li, Y.; Zhao, G.; Man, P.; Lin, Y.; Wang, M. A novel algorithm for breast mass classification in digital mammography based on feature fusion. J. Healthc. Eng. 2020, 2020, 8860011.
[13] Roy, V. Breast Cancer Classification with Multi-Fusion Technique and Correlation Analysis. Fusion Pract. Appl. 2022, 9, 48.
[14] Nakach, F.Z.; Idri, A.; Goceri, E. A comprehensive investigation of multimodal deep learning fusion strategies for breast cancer classification. Artif. Intell. Rev. 2024, 57, 327.
[15] Uddin, K.M.M.; Biswas, N.; Rikta, S.T.; Dey, S.K. Machine learning-based diagnosis of breast cancer utilizing feature optimization technique. Comput. Methods Programs Biomed. Update 2023, 3, 100098.
[16] Liu, T.; Huang, J.; Liao, T.; Pu, R.; Liu, S.; Peng, Y. A hybrid deep learning model for predicting molecular subtypes of human breast cancer using multimodal data. Irbm 2022, 43, 62–74.
[17] Kousalya, K.; Saranya, T. Improved the detection and classification of breast cancer using hyper parameter tuning. Mater. Today Proc. 2023, 81, 547–552.
[18] Tanaka, H.; Chiu, S.W.; Watanabe, T.; Kaoku, S.; Yamaguchi, T. Computer-aided diagnosis system for breast ultrasound images using deep learning. Phys. Med. Biol. 2019, 64, 235013.
[19] Mokni, R.; Haoues, M. CADNet157 model: Fine-tuned ResNet152 model for breast cancer diagnosis from mammography images. Neural Comput. Appl. 2022, 34, 22023–22046.

Received: July 19, 2025

Volume 38 No. 1s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[20] A. N. Kawade, A. M. Gaikwad, B.A.Pise , and A.D.Shilimkar , "Wearable Haptic Glove for the Hearing Impaired: Enhancing Spatial Awareness with Flex Sensors for Better Environmental Perception", IJACECT, vol. 13, no. 2, pp. 16–21, Mar. 2025.
[21] Kumar, A.; Singh, S.K.; Saxena, S.; Lakshmanan, K.; Sangaiah, A.K.; Chauhan, H.; Shrivastava, S.; Singh, R.K. Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer. Inf. Sci. 2020, 508, 405–421.
[22] Abimouloud, M.L.; Bensid, K.; Elleuch, M.; Aiadi, O.; Kherallah, M. Vision transformer-convolution for breast cancer classification using mammography images: A comparative study. Int. J. Hybrid Intell. Syst. 2024, 20, 67–83.

Received: July 19, 2025