INDERSCIENCE PUBLISHERS
Linking academia, business and industry through research

International Journal of Data Mining and Bioinformatics

ISSN online: 1748-5681 - ISSN print: 1748-5673
https://www.inderscience.com/ijdmb

Establishment of artificial intelligence pathological feature diagnosis model and molecular mechanism

Yanping Zhang, Yuan Zhang, Haimiao Xu, Yuxi Wang

DOI: 10.1504/IJDMB.2026.10077478

Article History:

Received:	18 October 2025
Last revised:	22 November 2025
Accepted:	15 January 2026
Published online:	29 May 2026

Establishment of artificial intelligence pathological feature diagnosis model and molecular mechanism

Yanping Zhang

Department of Pathology,
Hangzhou Lin'an Traditional Chinese Medicine Hospital, China
and
Affiliated Hospital, Hangzhou City University,
Hangzhou, 311300, Zhejiang, China
Email: yan1126zhang@163.com

Yuan Zhang

Second Clinical Medical College,
Zhejiang Chinese Medical University,
Hangzhou, 311300, Zhejiang, China
Email: zyuan0516@163.com

Haimiao Xu

Pathology Department,
Zhejiang Cancer Hospital,
Hangzhou, 311300, Zhejiang, China
Email: xuhaimiao@126.com

Yuxi Wang*

Department of Research and Education,
Hangzhou Lin'an Traditional Chinese Medicine Hospital, China
and
Affiliated Hospital, Hangzhou City University,
Hangzhou, 311300, Zhejiang, China
Email: 18906816231@163.com
*Corresponding author

Abstract: In the current diagnosis of cancer, the analysis of pathological section images and molecular markers (such as HER2, hormone receptor status, etc.) is usually performed separately, which can easily lead to difficulties in early identification, deviations in subtype classification, and limitations in personalized treatment decisions. This research solves this problem by establishing a breast cancer diagnosis model based on visual converter (ViT) and full connected neural network (FCNN). The experimental results show that the diagnostic model established in this study performs the best in terms of accuracy (0.963), recall rate (0.947), precision (0.952), and F1 score (0.950). In addition, the model shows high accuracy in classifying eight breast cancer subtypes in the cancer histopathological image dataset. The diagnostic model

established in this study is helpful in promoting the development of precision medicine for cancer, improving the efficiency of clinical treatment, and has important practical value in reducing cancer mortality.

Keywords: breast cancer diagnosis; vision transformer; ViT; fully connected neural network; FCNN; molecular markers; prognosis prediction.

Reference to this paper should be made as follows: Zhang, Y., Zhang, Y., Xu, H. and Wang, Y. (2026) 'Establishment of artificial intelligence pathological feature diagnosis model and molecular mechanism', *Int. J. Data Mining and Bioinformatics*, Vol. 30, No. 6, pp.1–20.

Biographical notes: Yanping Zhang graduated from Wenzhou Medical University, majoring in Clinical Medicine. She is currently working in the Department of Pathology, Lin'an Traditional Chinese Medicine Hospital in Hangzhou. Her research interests include pathological diagnosis, digital pathology, mechanism research on artificial intelligence, etc.

Yuan Zhang received his Bachelor's degree from Taizhou University, P.R. China. Currently, he studies in Zhejiang Chinese Medical University. His research interest is cytopathology.

Haimiao Xu received his Bachelor's degree from Zhejiang University, P.R. China. His research interest is cytopathology.

Yuxi Wang received his Master's degree from Heilongjiang University of Chinese Medicine. He is currently working in the Research and Education Department of Lin'an Traditional Chinese Medicine Hospital in Hangzhou. His research interests include artificial intelligence, pathological diagnosis and treatment, and research on the molecular mechanisms of respiratory diseases.

1 Introduction

Early diagnosis and personalised treatment of breast cancer are important to improve the survival of patients. Individualised treatment involves developing a targeted treatment strategy based on the patient's genetic background, molecular profile, pathological type, and clinical condition. In breast cancer, individualised treatment analyses molecular markers and pathological features to select the most effective surgery, chemotherapy, radiotherapy, or targeted drug regimen to improve efficacy and minimise side effects. With the development of artificial intelligence (AI) technology (Chan et al., 2023; Doma et al., 2022), especially the rise of deep learning (Mahoro and Akhloufi, 2022; Allugunti, 2022), AI has shown great potential in the diagnosis and treatment of breast cancer. However, some challenges persist in the current diagnosis of breast cancer, especially that pathological images and molecular marker analyses are often performed separately. Pathological image analysis relies on the experience and manual annotation of pathologists, while the detection of molecular markers (such as HER2, hormone receptor status, etc.) often relies on separate genetic testing or immunohistochemistry methods. This separate analysis approach may result in decisions on early tumour identification (Ng et al., 2023; Zou et al., 2022), subtype classification, and personalised treatment

(Passaro et al., 2024; Zhang et al., 2024) that fail to take into account molecular information fully.

In recent years, many studies have attempted to apply AI technology to the diagnosis and prognosis of breast cancer. Xie et al. (2022) used an AI technology to enhance the Raman spectroscopy strategy for label-free serum exosome spectroscopy to obtain accurate breast cancer diagnoses and evaluate surgical outcomes. The results showed that the Raman spectroscopy analysis of serum exosomes had a 100% predictive accuracy rate for human patients with different breast cancer subtypes who had not undergone surgery (Xie et al., 2022). Ahmed and Mohammed (2023) used a neural network to process a database of 55,890 images by dividing the images into five categories: non-cancer, benign calcification, malignant calcification, benign mass, and malignant mass; they used a machine learning algorithm to classify and identify breast cancer. Additionally, O'Connell et al. (2022) asked AI and ten radiologists to diagnose a set of 299 breast ultrasound images jointly. They compared the recognition results of AI with those of radiologists and found that the diagnosis results were consistent with the readings of the ten radiologists in multiple indicators (O'Connell et al., 2022; Wu et al., 2024). Meanwhile, Liu et al. proposed a novel breast pathology classification framework. The model was trained using the ImageNet dataset and fine-tuned using an enhanced dataset. An improved cross-entropy loss function was also designed for fitting. Experimental results showed that the model they proposed performed better than other methods at different magnifications (Liu et al., 2022a; Shu, 2025). Although these methods have achieved remarkable results in image analysis, they still have shortcomings. Notably, the method to combine image features with molecular marker data is not close enough, thus resulting in the failure to utilise the molecular-level information of the tumour fully for comprehensive analysis.

An increasing number of studies have attempted to overcome this shortcoming by fusing different types of data to improve the accuracy of breast cancer diagnosis and prognosis prediction. Liu et al. (2022b) proposed a hybrid deep learning model based on multimodal data, which combines the genetic modality data and image modality data of breast cancer patients to construct a multimodal fusion framework. Based on the idea of weighted linear aggregation and the fusion of the outputs of the two feature grids, the fused features are used to predict the subtype of breast cancer (Liu et al., 2022b; Luo et al., 2024). Furthermore, Atrey et al. (2024) proposed a new semiautomatic breast tumour multimodal classification system that combines the features of mammograms and ultrasound images. The relevant features are finally determined to classify the tumour subtypes by extracting 42 greyscale features and performing statistical significance analysis (Atrey et al., 2024; Ali and Simmou, 2025). Yuan and Xu (2024) proposed a deep multimodal fusion network that predicts the five-year survival rate of breast cancer patients by integrating clinical data, gene expression, and mutation data; they verified the effectiveness of the network by comparing it with other methods (Rizvi and Khalid, 2025). However, these methods continue to face issues regarding high model complexity and incomplete fusion of data processing. In particular, an efficient fusion method between high-dimensional pathological image features and molecular features is still difficult to find. Current research in the diagnosis of breast cancer pathology is fragmented between image recognition and molecular testing, thus causing difficulty in achieving multidimensional information complementarity (Wu, 2024). Existing AI models often focus on pathology image classification, which overlooks the critical impact of molecular markers on tumour biological behaviour. This shortcoming results in limited

accuracy in subtype identification and prognosis prediction. Some multimodal fusion methods suffer from insufficient feature integration and poor model interpretability, thus limiting their application value in personalised clinical care.

In recent years, the vision transformer (ViT) model has demonstrated tremendous potential in medical image analysis because of its powerful global feature modelling capabilities. Unlike traditional convolutional neural networks (CNNs), which are limited to a local receptive field, ViT uses a self-attention mechanism to capture dependencies between distant pixels in an image, thereby making it highly suitable for identifying complex tissue structures and heterogeneous lesions in pathological sections. This study constructs an intelligent model for breast cancer diagnosis that fuses pathological images with molecular markers. It proposes a multimodal fusion framework based on the ViT and a fully connected neural network (FCNN). ViT extracts deep morphological features from pathological slides, while FCNN processes molecular marker data. This model effectively fuses these features at the feature level, thereby overcoming the disconnect between image and molecular information in traditional approaches. It not only improves the accuracy of breast cancer diagnosis and subtype classification but also enables a reliable prognostic prediction based on clinical data, thus providing a valuable reference for precision medicine in breast cancer. When establishing a breast cancer diagnosis model, the ViT model is used to identify pathological slice images in blocks. Each small block is used as the smallest processing unit. After multihead attention processing, the average pooling method is used to aggregate the features of each small block to form the overall features of the entire pathological slice image. The molecular annotations are processed using the FCNN model. Then, the features at the molecular level are extracted. Pathological slice image features and molecular annotation features are fused via the fully connected layer. Classification probabilities are output through the activation function to diagnose breast cancer. Based on the diagnosis, the patient's clinical information is also introduced for prognosis prediction, which can estimate the patient's survival time and provide support for the patient's personalised clinical treatment. The experimental verification confirms that the diagnostic model proposed in this study outperforms other compared models across four evaluation metrics, thus demonstrating superior diagnostic capability. The model in this research can classify the eight subtypes on the breast cancer histopathological image (BreakHis) dataset and identify most of the subtypes effectively. Notably, the lowest recognition accuracy is 90.4%. Combining the patient's clinical data for survival prediction reveals that the model can accurately capture the patient's survival trend and provide strong support for clinical decision-making. Finally, an adversarial attack test demonstrated that the model can maintain high stability and accuracy in the face of small perturbations.

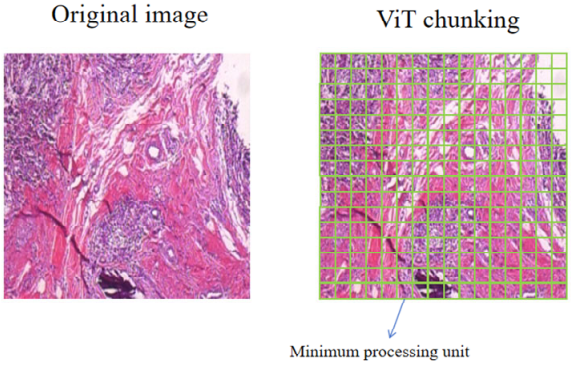
2 Pathological section image feature extraction

In the diagnosis of breast cancer, the most important visual data source is pathological section image data. The extraction of pathological section image features is crucial for subsequent disease prediction and subtype classification. If people only rely on artificial features, capturing the subtle changes of lesions is often difficult. However, automatic feature extraction based on deep learning, such as the ViT (Zhang et al., 2023; Khan et al., 2023) used in this study, effectively learns the global and local features of

pathological slice images through the self-attention mechanism, thereby improving the diagnostic accuracy of breast cancer pathology images.

ViT is a transformer-based image analysis and processing model that demonstrates excellent performance when processing large-scale image data. Unlike traditional CNNs, ViT does not perform convolution operations; instead, it divides the input image into multiple small blocks, treats each small block as the smallest unit to be processed independently, and learns through a multilayer self-attention mechanism (Houssein et al., 2024; Kumar and Solanki, 2023). Figure 1 presents a schematic diagram of ViT's block recognition of pathological slice images.

Figure 1 Pathological slice images divided by ViT (see online version for colours)



2.1 Image segmentation and linear embedding

The pathological slice image is divided into multiple small blocks to input the image into ViT. Each small block is linearly embedded. The specific steps are as follows:

1 Image segmentation: The pathological image is divided into N small blocks. Each small block represents a two-dimensional matrix. Different small blocks are labelled with different numbers.

2 Linear embedding: Each small block is transformed into a high-dimensional vector through linear embedding. The purpose of the transformation is to map each small block into a space of fixed dimension. The mapping formula is as follows:

$$z_i = W_{embed} \cdot X_i + b_i, \tag{1}$$

where W_{embed} represents the learned weight matrix.

3 Position encoding: ViT adds position encoding to the embedding vector of each small block to preserve the spatial position information of each small block in the pathological slice image. Then, it adds a fixed encoding value to the embedding vector of each small block:

$$z_i^{pos} = z_i + pos_i, \tag{2}$$

where pos_i represents the position information of the i^{th} small block.

2.2 Self-attention mechanism

Through the self-attention mechanism, ViT can dynamically focus on the connections between small blocks in the pathological slice image. For the embedded representation of each small block, ViT calculates self-attention to learn the correlation between different small blocks.

The query, key, and value vectors for the embedded representation of each small block are calculated using the learned weight matrix:

$$Q_i = W_Q \cdot z_i^{pos}, \tag{3}$$

$$K_i = W_K \cdot z_i^{pos}, \tag{4}$$

$$V_i = W_V \cdot z_i^{pos}. \tag{5}$$

The attention score between each small block is calculated using the dot product:

$$A_{ij} = \frac{Q_i \cdot K_j^T}{\sqrt{d}}. \tag{6}$$

The value vector is weighted summed using the calculated attention score to obtain the final representation of each patch:

$$Z_i = \sum_{j=1}^{N} A_{ij} \cdot V_j \tag{7}$$

2.3 Multihead self-attention mechanism

At the same time, a multihead self-attention mechanism is introduced to capture the features of different levels and aspects in pathological slice images (Reddi et al., 2024; Li et al., 2024). The query vector, key vector, and value vector are transformed into multiple heads through different linear transformations. Each head can perform independent self-attention calculations. Finally, the outputs of multiple heads are spliced:

$$Z_{multi} = Concat\left(Z_1, Z_2, ..., Z_h\right) \cdot W_{out} \tag{8}$$

where h represents the number of attention heads.

2.4 Feature aggregation

ViT can extract the long-distance dependencies of each small block in the pathological slice image and learn the global image features. After being processed by multihead self-attention, all processed small block features are aggregated. The aggregation method used is global average pooling:

$$z_{final} = GAP\left(Z_{multi}\right) \tag{9}$$

Through ViT, the feature extraction of the slice images of breast cancer pathology can deeply explore the detailed information in pathology slice images especially those in the capture of morphology and subtle lesions. ViT leverages the self-attention mechanism

and multihead self-attention to capture the relationship between distant pixels in the image effectively by utilising the global context of pathology slice images. This process enables ViT to exhibit enhanced performance and flexibility in breast cancer diagnosis.

3 Molecular marker data processing

The analysis of molecular markers provides valuable biological information for diagnosing and prognosticating breast cancer. This study employs FCNN (Sun et al., 2023; Prajapati and Kwon, 2022) to process molecular marker data from breast cancer patients. Moreover, it integrates the processed data with features from pathological slice images to enhance the diagnostic accuracy and subclassification capabilities of breast cancer.

3.1 Molecular marker data preprocessing

Molecular annotation data consists of gene expression data, mutation status, and protein expression. The data is high-dimensional and usually contains noise. Therefore, it needs to be properly preprocessed. In breast cancer, common molecular markers include the estrogen receptor (ER), progesterone receptor (PR), HER2, and Ki-67, which are used to determine the biological behaviour of tumours, guide targeted therapy, and evaluate prognosis.

The Z-score method (Utami and Hardana, 2022; Destriwanti et al., 2022) is used to standardise molecular annotations:

$$X_{norm} = \frac{X - \mu}{\sigma}. \tag{10}$$

The K nearest neighbor interpolation method is used to fill the missing values:

$$X_{fill} = \frac{\sum_{k=1}^{K} w_k X_{i,k}}{\sum_{k=1}^{K} w_k}. \tag{11}$$

Among them, $X_{i,k}$ represents the k^{th} neighbour value of the missing data. Meanwhile, w_k represents the weight of the k^{th} neighbour. This weight is calculated by the distance between the missing value and the neighbour value. A close distance entails great weight. The distance calculation formula is as follows:

$$d\left(X_i, X_j\right) = \sqrt{\sum_{k=1}^{m} \left(X_{ik} - X_{jk}\right)^2}. \tag{12}$$

3.2 Design of FCNN

In the process of processing the molecular marker data of breast cancer in this study, FCNN is mainly used to learn and extract the potential rules in the data to provide reliable feature representation for subsequent breast cancer diagnosis, subtype classification, and prognosis prediction (Yang et al., 2023; Kim et al., 2023). FCNN

consists of several layers of fully connected neurons. Neurons within each layer are entirely connected to neurons in the preceding layer, thus allowing each neuron to receive signals from all the neurons in that layer.

The core calculation of FCNN is the weighted sum operation between each fully connected layer, where the calculation process of the l^{th} layer can be expressed as follows:

$$y^{(l)} = f\left(W^{(l)} y^{(l-1)} + b^{(l)}\right) \tag{13}$$

Finally, FCNN outputs a high-dimensional feature vector of molecular annotation data, which is represented as $y^{(L)}$.

4 Diagnostic model

In the diagnostic model established in this study, the features extracted by FCNN and ViT are fused to obtain a rich representation to help diagnose breast cancer. The two feature vectors are fused by weighted summation:

$$V_{fused} = \alpha \cdot z_{final} + \beta \cdot y^{(L)} \tag{14}$$

where α, β represent the weight coefficients of the high-dimensional feature vector extracted by FCNN and the pathological slice image features extracted by ViT. These two coefficients are used to control the influence weight of the feature source during fusion.

The diagnosis result is output through a fully connected network layer. The activation function used is the Softmax (Wang et al., 2023; Rajasekaran et al., 2024) function:

$$p = Soft\max\left(W \cdot V_{fused} + b\right) \tag{15}$$

The output of the activation function is a diagnostic category probability distribution:

$$p_i = \frac{e^{z_i}}{\sum_{i=1}^{C} e^{z_i}}, \tag{16}$$

where p_i represents the category probability of different subtypes. This probability can be used to determine whether the patient is ill and the subtype of the patient.

For the diagnosis of breast cancer, FCNN is optimised by minimising the cross-entropy loss function (Hong and Ling, 2024; Shin and Chung, 2023):

$$\Gamma = -\sum_{i=1}^{C} y_i \log\left(p_i\right). \tag{17}$$

In FCNN, the back propagation algorithm (Napoleon and Kalaiarasi, 2022; Chakraborty et al., 2024) is used to calculate the gradient and update the weights:

$$W^l = W^l - \eta \cdot \frac{\partial \Gamma}{\partial W^l}. \tag{18}$$

Through multiple iterations of training, the weights of FCNN are gradually updated, thus enabling the network to make accurate diagnoses on given data. Given that pathology images contain tumour morphology information and molecular markers reflect intrinsic biological characteristics, ViT can effectively capture global contextual features of images. Meanwhile, FCNN excels at processing structured molecular data. A dual-branch network extracts and fuses these two types of features, thus preserving the integrity of each and achieving complementary advantages for heterogeneous data. The model structure is shown in Figure 2.

Figure 2 Diagnostic model structure (see online version for colours)

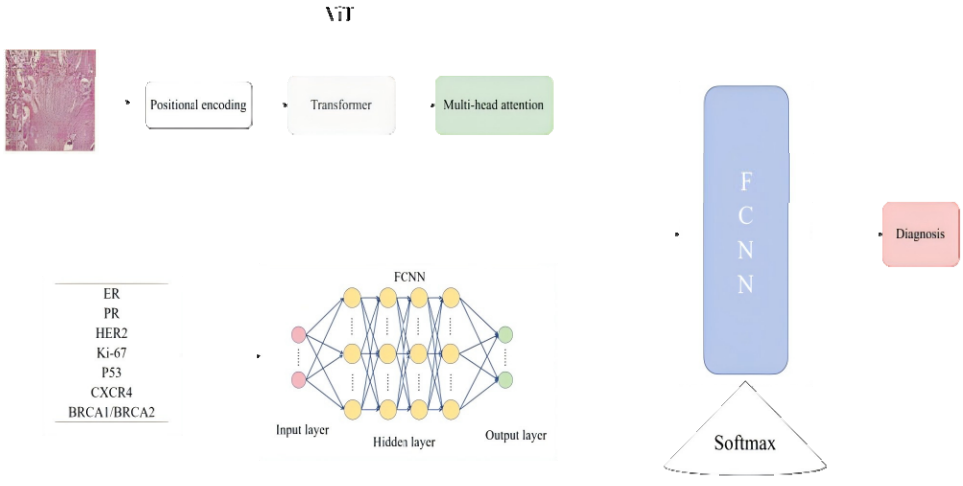


Figure 2 is the diagnostic model designed in this study. This model consists of two main branches: the image branch uses ViT to segment pathology slide images into image blocks. After linear embedding and positional encoding, it uses a multihead self-attention mechanism to extract global and local morphological features. Ultimately, it outputs a high-dimensional feature vector for the image. The molecular branch uses FCNN to perform nonlinear transformations on standardised molecular marker data (such as ER, PR, and HER2) to extract molecular-level features. The features from both branches are weighted and fused before being fed into a fully connected layer. Combined with the Softmax activation function, the model outputs the breast cancer diagnosis result and subtype classification probability, thereby enabling collaborative decision-making based on multimodal information. The model first inputs the pathological slice image into the ViT model for processing. Specifically, the pathological slice image can be divided into multiple small blocks. Each small block is used as the smallest processing unit. After position encoding, the spatial information is embedded into the image. Then, each small block is processed by multihead attention to capture the correlation between each small block and extract the deep features of the pathological slice image. At the same time, molecular marker data are processed and feature extracted through FCNN, which can extract key information at the molecular level, including the expression level and degree of variation of molecular markers. Features from ViT and FCNN are sent to an FCNN for fusion, which efficiently fuses the information of pathological slice images and molecular annotations to obtain comprehensive and accurate diagnostic information. Finally, the

diagnosis result of breast cancer is output through the Softmax activation function, including whether breast cancer exists and the probability of possible subtype categories.

5 Prognosis prediction

After diagnosis, if the patient is confirmed to have breast cancer, prognosis prediction is conducted using clinical data. The primary objective at this stage is to forecast survival, recurrence risk, and other prognostic indicators based on the patient's clinical data, diagnosis outcomes, and relevant molecular marker information.

The input of prognosis prediction includes two parts:

1 The diagnosis of breast cancer includes the diagnosed breast cancer subtype and the molecular markers of the tumour.

2 Clinical data include information such as the patient's age, gender, tumour size, and lymph node metastasis status.

In prognosis prediction, an FCNN is used to process the input data. The input data are expressed in the form of $X = [X_{cancer}, X_{clinical}]$, where X_{cancer}, $X_{clinical}$ represent the diagnostic information of breast cancer and the clinical data of the patient, respectively. The output layer outputs the prognosis prediction results:

$$\hat{y} = W_s \cdot X + b_s. \tag{19}$$

The final output is the predicted result, which includes survival and recurrence risk.

The mean square error function (Subbarao et al., 2023; Nguyen et al., 2023) is used to optimise prognosis prediction:

$$\Gamma_r = \frac{1}{N}\sum_{i=1}^{N}(\hat{y} - y)^2. \tag{20}$$

6 Experimental design

6.1 Dataset preparation

Images of breast cancer pathological sections, molecular marker data, and clinical data for individual patients are essential to diagnose and predict the prognosis of breast cancer. This study utilises the BreakHis, which consists of 7,909 images. These images are categorised into benign and malignant tumours with the malignant tumours further divided into eight subtypes. Figure 3 shows part of the BreakHis dataset.

Table 1 presents information on molecular markers linked to breast cancer, which play a significant role in its occurrence, development, and prognosis. ER and PR, two key hormone receptors in breast cancer, promote the growth and proliferation of cancer cells by binding with estrogen and progesterone. HER2 is a cell membrane receptor that is involved in the growth and division of cells. In breast cancer cells with overexpression of HER2, this overexpression accelerates the growth and spread of tumours. Ki-67 is a marker of cell proliferation. Its expression level is closely related to the proliferation activity of tumour cells. In the diagnosis and prognosis of breast cancer, high expression

of Ki-67 usually shows that tumour cells have a strong proliferation ability, which means that the tumour is malignant and has a high risk of metastasis. P53 is a well-known tumour suppressor that can maintain the genetic stability of cells by regulating the cell cycle and repairing damaged DNA. As a chemokine receptor, CXCR4 is involved in the migration process of cells. Its expression is commonly seen in metastatic tumours of breast cancer.

Figure 3 Part of the BreakHis dataset (see online version for colours)

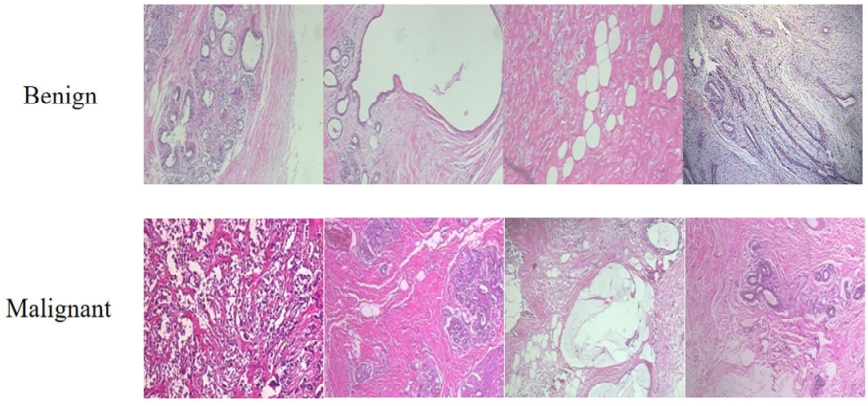


Table 1 Molecular annotation information

Molecular markers	*Full name*	*Biological function*
ER	Estrogen receptor	Involved in the growth and differentiation of breast cancer cells
PR	Progesterone receptor	Involved in breast cancer cell proliferation and survival
HER2	Human epidermal growth factor receptor 2	Receptors that drive tumour cell growth and spread
Ki-67	Proliferating cell nuclear antigen Ki-67	Cell proliferation markers
P53	Tumour suppressor gene p53	Responsible for cell cycle regulation and DNA repair
CXCR4	CXC chemokine receptor 4	Involved in cell migration, especially chemotaxis of leukocytes
BRCA1/BRCA2	Breast cancer susceptibility gene 1/2	Participate in DNA repair and inhibit tumour formation

In data preprocessing, slide images of breast cancer pathology are uniformly resized to 224 × 224 pixels, normalised using ImageNet standardisation parameters, and subjected to data augmentation, such as rotation and flipping, to improve generalisation ability. Molecular marker data are Z-score normalised to eliminate dimensionality differences. Then, missing values are filled using K-nearest neighbour interpolation. Continuous variables in clinical data are standardised. Next, categorical variables are one-hot encoded. All data are divided into training, validation, and test sets in a 7:1:2 ratio to ensure balanced distribution across categories and avoid bias.

6.2 Hyperparameter tuning

This study mainly uses the ViT model and FCNN. Table 2 shows the relevant hyperparameter design during model training.

Table 2 Hyperparameter settings

Hyperparameters	*Values*	*Description*
Learning rate	0.0001	Control the step size of weight updates
Transformer layers	8	Number of Transformer layers in ViT
Attention heads	12	Number of heads in self-attention mechanism
Hidden dimension	256	Hidden layer dimensions of Transformer in ViT and number of neurons in fully connected layers in FCNN
FCNN layers	4	The number of hidden layers in the FCNN part
Dropout rate	0.3	Preventing Overfitting
Optimiser	Adam	Adam is suitable for large-scale data
Batch size	64	The number of samples per training
Epochs	100	The number of training set iterations

The learning rate is set to 0.0001 to balance convergence speed and stability. The Adam optimiser is used for adaptive adjustment. ViT sets an eight-layer transformer module and 12 attention heads to capture the deep features of the image fully. The hidden dimension is unified to 256. The FCNN has four layers. Moreover, a dropout rate of 0.3 is used to prevent overfitting. The batch size is set to 64 to balance memory efficiency and gradient estimation accuracy. The training cycle is set to 100 rounds to ensure the full convergence of the model. Selecting a small learning rate (0.0001) can ensure stable convergence of the model. The design of an eight-layer transformer and 12 attention heads can improve the expressive power of the ViT model and the feature extraction effect of pathological segmentation images. Setting the hidden dimension to 256 can effectively enhance the expressiveness of the ViT and FCNN models. In addition, setting the hidden layer of FCNN to four layers can ensure sufficient performance while avoiding overfitting. Using the Adam optimiser can improve training efficiency through adaptive learning rate. The batch size of 64 can achieve an excellent balance between memory consumption and training speed. Having a total of 100 training cycles ensures that the model can be fully trained.

7 Result analysis

7.1 Diagnostic performance analysis

The diagnostic model based on ViT + FCNN established in this study is compared with some existing models. The comparison indicators include accuracy, precision, recall rate, and F1 score. The compared models are Swin Transformer (Tanimola et al., 2024; Zhuang et al., 2024), DenseNet (Samudrala and Mohan, 2024; Chandana Mani and Kamalakannan, 2024), DeepLabV3+ (Sufri et al., 2024; He et al., 2024) and Attention U-net (Chen et al., 2022). The aim of this comparative analysis of diagnostic accuracy indicators is to assess the performance of the mentioned models in breast cancer

diagnosis tasks thoroughly, thus ensuring that the diagnostic models achieve an optimal balance between accuracy and reliability. This analysis can objectively measure the ability of each model in tumour identification and provide a scientific basis for clinical practice by comparing these indicators related to diagnostic accuracy.

Table 3 Comparison of diagnostic performance

Model	*Precision*	*Recall*	*Accuracy*	*F1-score*
ViT + FCNN	0.952	0.947	0.963	0.950
Swin transformer	0.941	0.924	0.949	0.933
DenseNet	0.936	0.915	0.938	0.926
DeepLabV3+	0.921	0.906	0.916	0.914
Attention U-Net	0.938	0.923	0.943	0.931

According to the comparison results in Table 3, the diagnostic model in this study performs best in precision (0.952), accuracy (0.963), recall (0.947), and F1 score (0.950), which is significantly better than he performance of other models. The advantages of ViT + FCNN in precision and recall show that the diagnostic model has high accuracy and a low missed diagnosis rate in the early identification of breast cancer and the prediction of positive cases. In comparison, other models, such as Swin Transformer, have a precision and recall of 0.941 and 0.924, respectively, which is good but not as good as ViT + FCNN. Attention U-net is also slightly insufficient in these indicators with a precision and recall of 0.938 and 0.923, respectively. The performance of DenseNet and DeepLabV3+ is relatively weak. Specifically, the accuracy of DeepLabV3+ is 0.921, and the recall rate is 0.906, which limits the efficiency of the model in clinical application. Overall, the ViT + FCNN diagnostic model in this study has the advantage of integrating the image features of breast cancer pathology sections and molecular marker information data. It shows high diagnostic accuracy and robustness in breast cancer diagnosis and has strong clinical application potential.

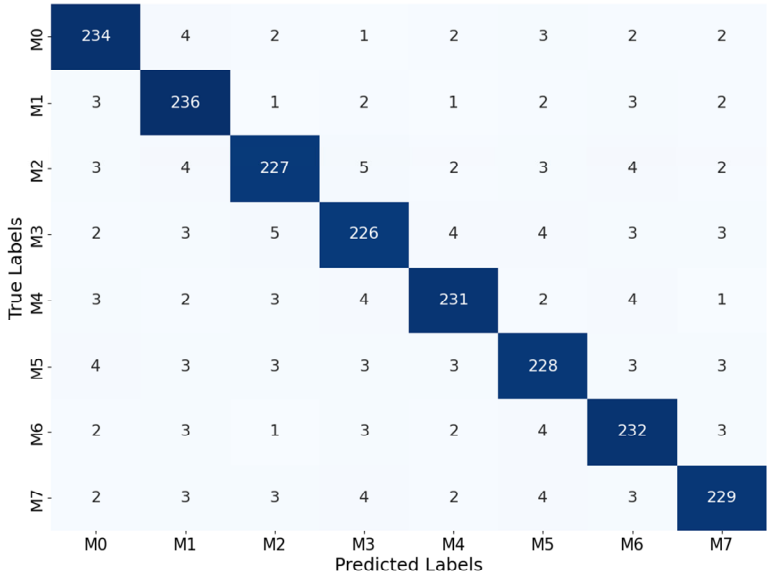
7.2 Subtype classification

After the diagnostic model is trained, the 2,000 malignant tumour pathology section images in the dataset are classified into subtypes. The dataset has eight subtypes: M0, M1, M2, M3, M4, M5, M6, and M7. Each subtype selected has 250 corresponding pathology section images. The classification results are presented in Figure 4. Accurate subtype classification can help select the appropriate treatment plan, reduce the risk of misdiagnosis and missed diagnosis, and improve the effect of personalised treatment.

According to the results in Figure 4, the diagnostic model established in this study performs well in the classification task of the eight subtypes of breast cancer. In particular, the classification accuracy is high in the M0 (accurate classification of 234), M1 (accurate classification of 236), M4 (accurate classification of 231), and M6 (accurate classification of 232) subtypes. This result verifies ViT's strong recognition ability for typical pathological features and the promotion of multimodal fusion to stable subtype discrimination. It indicates that the model is suitable for the accurate classification of common breast cancer subtypes and has the potential to assist clinical pathological diagnosis especially in the discrimination of typical cases. However, the model has some errors in some similar subtypes, such as five misclassifications between M2 and M3 and

four misclassifications between M3 and M4. This outcome indicates that some morphological similarities exist between these subtypes, which are difficult for the model to distinguish. The diagnostic model established in this research can effectively identify most subtypes. Even the M3 subtype with the least number of accurate classifications has a classification accuracy of 90.4%. Therefore, the diagnostic model has shown good accuracy in the classification of the eight subtypes of breast cancer, which can provide effective support for the early diagnosis and personalised treatment of breast cancer.

Figure 4 Confusion matrix of 8 subtypes classification (see online version for colours)



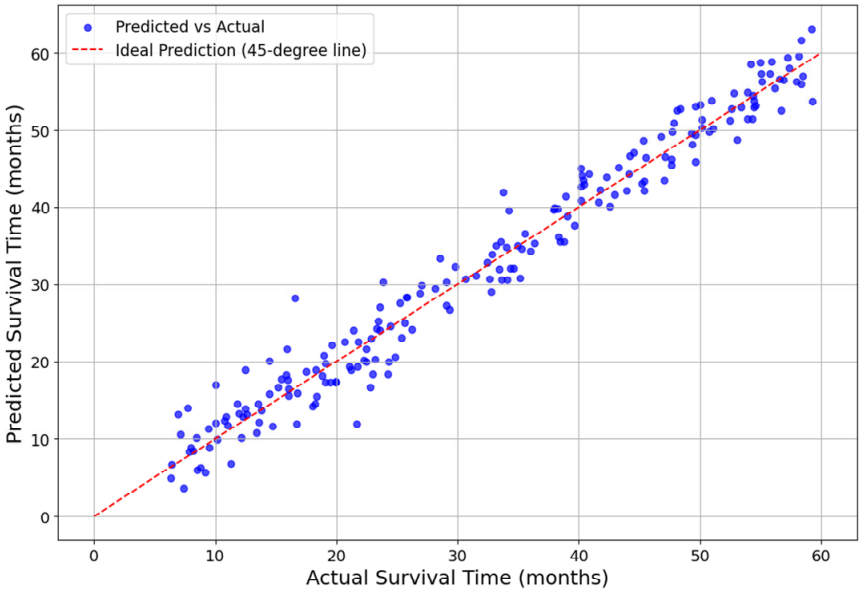
7.3 Survival time prediction analysis

The survival time of 200 deceased patients is predicted by combining the clinical data of the last examination with the patients' pathological section images and molecular marker information. The predicted survival time is compared with the actual survival time. A scatter plot is generated, as shown in Figure 5. Points that are close to the diagonal line indicate that the predicted survival time is near the actual survival time, thus reflecting an accurate prognosis prediction. A comparative analysis of the actual and predicted survival times of patients provides an intuitive evaluation of the model's prediction accuracy and reliability in breast cancer prognosis. The model's performance in breast cancer prognosis prediction is assessed by examining the proximity of the predicted values to the actual ones with a particular focus on the accuracy of survival prediction. This analysis can provide clinicians with accurate survival predictions, thereby providing a reference for the formulation of personalised treatment plans and patient risk assessment.

As can be seen from Figure 5, most of the points are closely distributed near the diagonal line. The figure shows that the survival time predicted by the model is very close to the actual survival time of the patient. Moreover, the model shows high accuracy in predicting the survival time of breast cancer patients. When the points are close to the diagonal line, the model's predicted time is also close to the actual survival time. In this

case, the prediction error is small. Although a few points deviate from the diagonal line, the overall deviation is small, thus indicating that the model can provide reliable survival time predictions in most cases. Through this prediction method, the model can accurately capture the patient's survival trend and provide strong support for clinical decision-making.

Figure 5 Comparison of predicted survival time and actual survival time (see online version for colours)



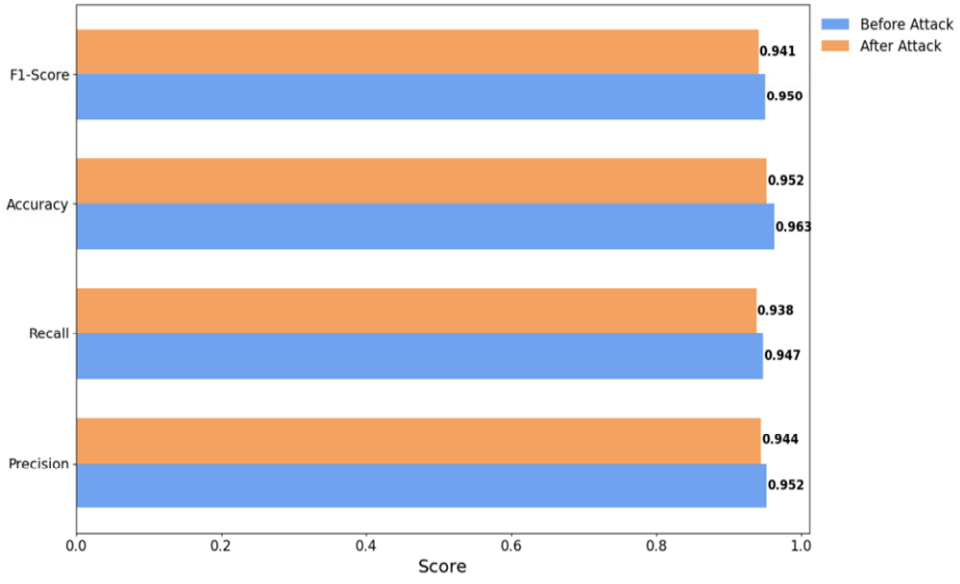
7.4 Model robustness analysis

The actual data of medical pathology slice images often have noise, sample variation, and even forged samples, which may affect the accuracy and reliability of the model. Through adversarial attack testing, the stability of the model in the face of small perturbations can be evaluated, thereby identifying the areas or features of the model that are most vulnerable to attack. Adversarial attack testing can help understand the potential weaknesses of the model and provide guidance for reliability and safety in clinical applications, thus ensuring that the model can work stably in complex real-world environments. The projected gradient descent (Siddiqi et al., 2023), which is an adversarial attack method, is used for testing. The dataset is iteratively perturbed several times. The perturbations are retained within a certain range. The changes in breast cancer diagnosis accuracy of the model before and after the adversarial attack are compared to evaluate the robustness of the model.

All evaluation indicators have dropped to varying degrees in Figure 6, which shows that the model cannot achieve the diagnostic accuracy before the adversarial attack in actual applications. The precision dropped by 0.8%, the recall rate dropped by 0.9%, the accuracy dropped by 1.1%, and the F1 score dropped by 0.9%. These changes show that although the breast cancer diagnosis model established in this study performs well in the

dataset, the performance of the model is significantly affected after the adversarial attack. The reason for the impact is that the adversarial attack changes the local features in the pathological slice image by applying small perturbations, thereby disrupting the model's ability to diagnose the image, thus resulting in an increase in misdiagnosis and missed diagnosis. Although the model is affected numerically, the decline is not very large, which shows that the model can still maintain high stability and accuracy in the face of small disturbances. Therefore, the model has good robustness to a certain extent, can effectively deal with adversarial attacks, and has the potential for practical clinical application.

Figure 6 Adversarial attack test (see online version for colours)



8 Conclusions

This study fuses and analyses the pathological section images and molecular annotation information of breast cancer, establishes a breast cancer diagnosis model to diagnose breast cancer, and classifies its subtypes. Based on the diagnosis, the patient's clinical information is combined to achieve the prognosis prediction of breast cancer patients. The main contribution of this study is as follows. This work combines pathological section images and molecular annotation information and comprehensively considers the characteristics at the pathological and molecular levels in the diagnosis of breast cancer. Moreover, it uses the ViT model to process pathological images in blocks and combines FCNN to analyse the characteristics of molecular annotations to establish a breast cancer diagnosis model successfully, which significantly improves the diagnostic accuracy of breast cancer. Furthermore, it uses the BreakHis dataset to classify breast cancer subtypes. The model shows good classification ability when dealing with different subtypes of breast cancer and has good generalisation. Experimental results demonstrate that the proposed model achieves an accuracy of 0.963 and an F1 score of 0.950 on the

BreakHis dataset, thus significantly outperforming the comparison model and validating the effectiveness of the ViT and FCNN fusion strategy. The lowest accuracy across eight subtypes can reach 90.4%. Meanwhile, the survival prediction results are close to reality, thereby demonstrating its strong classification capabilities and clinical prognostic value. This model also provides an efficient and reliable AI tool for the precision diagnosis and treatment of breast cancer with significant potential for translational applications. The introduction of patients' clinical data further expands the diagnostic function of breast cancer. The model can accurately capture the survival trend of breast cancer patients, provide clinicians with personalised treatment decision support, and improve the comprehensive value of diagnosis. Adversarial attack tests are conducted to verify the robustness and practical application capabilities of the model. The results show that the model maintains high accuracy and stability under adversarial attacks.

Although this study has made some contributions and has certain application value in the diagnosis and treatment of breast cancer, some areas need to be improved. Although the BreakHis dataset used in this work contains a variety of breast cancer subtypes, the size of the dataset and the number of samples are limited. In particular, imbalances may occur in the distribution of samples of certain subtypes. Moreover, the pathological images in the dataset may not cover all the variations encountered in actual clinical practice. Although this research experimentally validates the performance of the model on the dataset, it has not yet been extensively experimented in a real clinical setting. In a real clinical setting, the quality of pathological slide images, individual differences in patients, and changes in other medical conditions may affect the performance of the model. The acquisition of molecular marker data depends on the accuracy of the detection technology. Missing data or noise may also interfere with the fusion effect.

Therefore, future research can be expanded to a large and diverse breast cancer dataset that includes data from different hospitals and countries. At the same time, future studies should conduct prospective validation on real clinical data and conduct experiments in diverse data and patient groups with diverse clinical characteristics. Specifically, attention mechanisms can be introduced to optimise cross-modal feature interactions. Self-supervised learning can also be combined with unlabelled data to improve generalisation capabilities. Moreover, lightweight designs can be explored to adapt to real-time clinical needs. Multicentre, large-sample real-world datasets can also be constructed to conduct prospective verification and promote the transformation of models into clinical practical applications. Through such validation, the adaptability and actual effect of the model in complex and changing clinical environments can be further evaluated.

Declarations

The authors declare no conflict of interest.

References

Ahmed, F.M. and Mohammed, B.S. (2023) 'Feasibility of breast cancer detection through a convolutional neural network in mammographs', *Tamjeed Journal of Healthcare Engineering and Science Technology*, Vol. 1, No. 2, pp.36–43.

Ali, N.W. and Simmou, W. (2025) 'Framework for measuring the impact of a firm's artificial intelligence capability on creativity within the organizational and performance of the firm', *Journal of Intelligence Technology and Innovation*, Vol. 3, No. 1, pp.74–92.

Allugunti, V.R. (2022) 'Breast cancer detection based on thermographic images using machine learning and deep learning algorithms', *International Journal of Engineering in Computer Science*, Vol. 4, No. 1, pp.49–56.

Atrey, K., Singh, B.K. and Bodhey, N.K. (2024) 'Multimodal classification of breast cancer using feature level fusion of mammogram and ultrasound images in machine learning paradigm', *Multimedia Tools and Applications*, Vol. 83, No. 7, pp.21347–21368.

Chakraborty, A., Janapatla, P. and Chatterjee, B. (2024) 'Levenberg-Marquardt back-propagation algorithm for a developing unsteady hybrid nanofluid mixed convective flow across a revolving sphere: irreversibility analysis', *The European Physical Journal Plus*, Vol. 139, No. 11, pp.1–23.

Chan, R.C.K., To, C.K.C., Cheng, C.T., Yoshikazu, T., Yan, L.L.A. and Tse, G.M. (2023) 'Artificial intelligence in breast cancer histopathology', *Histopathology*, Vol. 82, No. 1, pp.198–210.

Chandana Mani, R.K. and Kamalakannan, J. (2024) 'Modeling of aquila optimizer with hybrid resnet-densenet enabled breast cancer classification on histopathological images', *Journal of Intelligent & Fuzzy Systems*, Vol. 46, No. 2, pp.5087–5102.

Chen, G., Li, L., Dai, Y., Zhang, J. and Yap, M.H. (2022) 'AAU-net: an adaptive attention U-net for breast lesions segmentation in ultrasound images', *IEEE Transactions on Medical Imaging*, Vol. 42, No. 5, pp.1289–1300.

Destriwanti, O., Sintha, L., Bertuah, E. and Munandar, A. (2022) 'Analyzing the impact of good corporate governance and financial performance on predicting financial distress using the modified Altman Z score model', *American International Journal of Business Management (AIJBM)*, Vol. 5, No. 2, pp.27–36.

Doma, M.K., Padmanandam, K., Sunil, T., Keshav Kumar, K., Abdualgalil, B. and Thakur, R.N. (2022) 'Artificial intelligence-based breast cancer detection using WPSO', *International Journal of Operations Research and Information Systems (IJORIS)*, Vol. 13, No. 2, pp.1–16.

He, S., Xiao, T. and Xia, J. (2024) 'Research on rural road extraction method based on improved DeepLabv3+ model', *Journal of Yunnan University: Natural Sciences Edition*, Vol. 46, No. 3, pp.486–495.

Hong, W. and Ling, S. (2024) 'Neural collapse for unconstrained feature model under cross-entropy loss with imbalanced data', *Journal of Machine Learning Research*, Vol. 25, No. 192, pp.1–48.

Houssein, E.H., Hammad, A., Samee, N.A., Alohali, M.A. and Ali, A.A. (2024) 'TFCNN-BiGRU with self-attention mechanism for automatic human emotion recognition using multi-channel EEG data', *Cluster Computing*, Vol. 27, No. 10, pp.14365–14385.

Khan, A., Rauf, Z., Sohail, A., Khan, A.R., Asif, H., Asif, A. et al. (2023) 'A survey of the vision transformers and their CNN-transformer based variants', *Artificial Intelligence Review*, Vol. 56, Suppl. 3, pp.2917–2970.

Kim, M.H., Kim, G.M., Ahn, J.M., Ryu, W-J., Kim, S-G., Kim, J.H. et al. (2023) 'Copy number aberrations in circulating tumor DNA enables prognosis prediction and molecular characterization of breast cancer', *JNCI: Journal of the National Cancer Institute*, Vol. 115, No. 9, pp.1036–1049.

Kumar, S. and Solanki, A. (2023) 'An abstractive text summarization technique using transformer model with self-attention mechanism', *Neural Computing and Applications*, Vol. 35, No. 25, pp.18603–18622.

Li, S., Xu, Y., Jiang, W., Zhao, K. and Liu, W. (2024) 'A modular fault diagnosis method for rolling bearing based on mask kernel and multi-head self-attention mechanism', *Transactions of the Institute of Measurement and Control*, Vol. 46, No. 5, pp.899–912.

Liu, M., Hu, L., Tang, Y., Wang, Chu., He, Y. and, Zeng, C. (2022a) 'A deep learning method for breast cancer classification in the pathology images', *IEEE Journal of Biomedical and Health Informatics*, Vol. 26, No. 10, pp.5025–5032.

Liu, T., Huang, J., Liao, T., Pu, R., Liu, S. and Peng, Y. (2022b) 'A hybrid deep learning model for predicting molecular subtypes of human breast cancer using multimodal data', *IRBM*, Vol. 43, No. 1, pp.62–74.

Luo, Z., Shao, X. and Ma, X. (2024) 'Enhancing learners' performance in contest through knowledge mapping algorithm: the roles of artificial intelligence and blockchain in scoring and data integrity', *Journal of Organizational and End User Computing*, Vol. 36, No. 1, pp.1–21.

Mahoro, E. and Akhloufi, M.A. (2022) 'Applying deep learning for breast cancer detection in radiology', *Current Oncology*, Vol. 29, No. 11, pp.8767–8793.

Napoleon, D. and Kalaiarasi, I. (2022) 'Classifying lung cancer as benign and malignant nodule using ANN of back-propagation algorithm and GLCM feature extraction on chest X-ray images', *Wireless Personal Communications*, Vol. 126, No. 1, pp.167–195.

Ng, A.Y., Oberije, C.J., Karpati, E., Fox, G., Glocker, B., Morris, E.A. et al. (2023) 'Prospective implementation of AI-assisted screen reading to improve early detection of breast cancer', *Nature Medicine*, Vol. 29, No. 12, pp.3044–3049.

Nguyen, V.A., Shafieezadeh-Abadeh, S., Kuhn, D. and Esfahani, P.M. (2023) 'Bridging Bayesian and minimax mean square error estimation via Wasserstein distributionally robust optimization', *Mathematics of Operations Research*, Vol. 48, No. 1, pp.1–37.

O'Connell, A.M., Bartolotta, T.V., Orlando, A., Jun, S-H., Baek, J. and Parker, K.J. (2022) 'Diagnostic performance of an artificial intelligence system in breast ultrasound', *Journal of Ultrasound in Medicine*, Vol. 41, No. 1, pp.97–105.

Passaro, A., Bakir, M.A., Hamilton, E.G., Wistuba, I.I., Swanton, C., Peters, S. et al. (2024) 'Cancer biomarkers: emerging trends and clinical implications for personalized treatment', *Cell*, Vol. 187, No. 7, pp.1617–1635.

Prajapati, R. and Kwon, G-R. (2022) 'A binary classifier using fully connected neural network for Alzheimer's disease classification', *Journal of Multimedia Information System*, Vol. 9, No. 1, pp.21–32.

Rajasekaran, V.A., Indirajithu, A., Jayalakshmi, P., Nayyar, A. and Balusamy, B. (2024) 'Gradient scaling and segmented SoftMax Regression Federated Learning (GDS-SRFFL): a novel methodology for attack detection in industrial internet of things (IIoT) networks', *The Journal of Supercomputing*, Vol. 80, No. 12, pp.16860–16886.

Reddi, P., Srinivas, G., Reddy, P.V.G.D.P. and Krihsna, D.S. (2024) 'A multi-head self-attention mechanism for improved brain tumor classification using deep learning approaches', *Engineering, Technology & Applied Science Research*, Vol. 14, No. 5, pp.17324–17329.

Rizvi, R.A. and Khalid, W. (2025) 'Unlocking innovation in manufacturing: the impact of data analytics maturity and knowledge-oriented leadership', *Journal of Management Science and Operations*, Vol. 3, No. 2, pp.17–43, DOI: https://doi.org/10.30210/JMSO.202503.002.

Samudrala, S. and Mohan, C.K. (2024) 'Semantic segmentation of breast cancer images using DenseNet with proposed PSPNet', *Multimedia Tools and Applications*, Vol. 83, No. 15, pp.46037–46063.

Shin, J. and Chung, W. (2023) 'Multi-band CNN with band-dependent kernels and amalgamated cross entropy loss for motor imagery classification', *IEEE Journal of Biomedical and Health Informatics*, Vol. 27, No. 9, pp.4466–4477.

Shu, M. (2025) 'Utilizing transfer learning for deep learning based image classification', *Journal of Intelligence Technology and Innovation*, Vol. 3, No. 1, pp.58–73.

Siddiqi, R., Ahraf, S.N. and Kandhro, I.A. (2023) 'Susceptibility of paediatric pneumonia detection model under projected gradient descent adversarial attacks', *International Journal of Electronic Security and Digital Forensics*, Vol. 15, No. 3, pp.322–331.

Subbarao, D., Kumar, A.S., Bikshapthi, S.K., Malathi, N. and Ashraf, Md. (2023) 'Analysis of data compression techniques in smart grids for optimising mean-square-error', *Applied Nanoscience*, Vol. 13, No. 2, pp.1591–1599.

Sufri, N.A.J., Babu, B. and As'ari, M.A. (2024) 'Polyp segmentation from colonoscopy image using DeepLabV3+', *Journal of Human Centered Technology*, Vol. 3, No. 2, pp.54–60.

Sun, J., Li, C., Wang, Z. and Wang, Y. (2023) 'A memristive fully connect neural network and application of medical image encryption based on central diffusion algorithm', *IEEE Transactions on Industrial Informatics*, Vol. 20, No. 3, pp.3778–3788.

Tanimola, O., Shobayo, O., Popoola, O. and Okoyeigbo, O. (2024) 'Breast cancer classification using fine-tuned SWIN transformer model on mammographic images', *Analytics*, Vol. 3, No. 4, pp.461–475.

Utami, T.W. and Hardana, A. (2022) 'Analisis Prediksi Kebangkrutan dengan Menggunakan Metode Altman Z-Score pada PT. Indofood Sukses Makmur, Tbk', *SOSMANIORA: Jurnal Ilmu Sosial Dan Humaniora*, Vol. 1, No. 4, pp.399–404.

Wang, H., Mahmood, T. and Ullah, K. (2023) 'Improved cocoso method based on frank softmax aggregation operators for t-spherical fuzzy multiple attribute group decision-making', *International Journal of Fuzzy Systems*, Vol. 25, No. 3, pp.1275–1310.

Wu, C-H. (2024) 'Eco-economic predictions: applying QPSO-BiLSTM and attention mechanisms for accurate renewable energy forecasting', *Journal of Management Science and Operations*, Vol. 2, No. 4, pp.28–47, DOI: https://doi.org/10.30210/JMSO.202402.011.

Wu, Z., She, Q. and Zhou, C. (2024) 'Intelligent customer service system optimization based on artificial intelligence', *Journal of Organizational and End User Computing*, Vol. 36, No. 1, pp.1–27.

Xie, Y., Su, X., Wen, Y., Zheng, C. and Li, M. (2022) 'Artificial intelligent label-free SERS profiling of serum exosomes for breast cancer diagnosis and postoperative assessment', *Nano Letters*, Vol. 22, No. 19, pp.7910–7918.

Yang, Y., Zhang, L., Tian, W., Li, Y., Qin, Q., Mao, Y. et al. (2023) 'Prognosis prediction and risk factors for triple-negative breast cancer patients with brain metastasis: a population-based study', *Cancer Medicine*, Vol. 12, No. 7, pp.7951–7961.

Yuan, H. and Xu, H. (2024) 'Deep multi-modal fusion network with gated unit for breast cancer survival prediction', *Computer Methods in Biomechanics and Biomedical Engineering*, Vol. 27, No. 7, pp.883–896.

Zhang, Q., Xu, Y., Zhang, J. and Tao, D. (2023) 'Vitaev2: vision transformer advanced by exploring inductive bias for image recognition and beyond', *International Journal of Computer Vision*, Vol. 131, No. 5, pp.1141–1162.

Zhang, X., Beeraka, N.M., Sinelnikov, M.Y., Glazachev, O.S., Ternovoy, K.S., Lu, P. et al. (2024) 'Breast cancer-related lymphedema: recent updates on clinical efficacy of therapies and bioengineering approaches for a personalized therapy', *Current Pharmaceutical Design*, Vol. 30, No. 1, pp.63–70.

Zhuang, J., Wu, X., Meng, D. and Jing, S. (2024) 'A Swin transformer and residual network combined model for breast cancer disease multi-classification using histopathological images', *Instrumentation*, Vol. 11, No. 1, pp.112–120.

Zou, R., Loke, S.Y., Tang, Y.C., Too, H-P., Zhou, L., Lee, A.S.G. et al. (2022) 'Development and validation of a circulating microRNA panel for the early detection of breast cancer', *British Journal of Cancer*, Vol. 126, No. 3, pp.472–481.