Network Modeling Analysis in Health Informatics and Bioinformatics (2024) 13:52
https://doi.org/10.1007/s13721-024-00489-8

REVIEW ARTICLE

Deep learning for breast cancer diagnosis from histopathological images: classification and gene expression: review

Oumeima Thaalbi[1] · **Moulay A. Akhloufi**[1]

Received: 10 June 2024 / Revised: 11 August 2024 / Accepted: 10 September 2024 / Published online: 26 September 2024

Abstract

Histopathology, the microscopic analysis of tissue to study the symptoms of the disease, is used to diagnose breast cancer. Breast cancer is specifically examined using tissue examination. The recent progress in deep learning has reinforced the potential of histopathological analysis by automating diagnostic processes. This review focuses on the integration of deep learning methods into histopathological analyses of breast cancer to categorise different types of cancer and predicting gene expression. We discuss the challenges associated with the use of deep learning models, such as the use of CNN variants to classify histopathological images into benign or malignant categories and specifically to identify different subtypes. We also observe the use of hybrid methods, as well as other approaches such as GNN, vision transformer etc. In addition, we investigate the capacity of deep learning to contribute to the interpretation of gene expression data that facilitates the prediction of breast cancer genes from the whole slide images, with the purpose of supporting personalised medicine. The majority of the selected studies are based on publicly available datasets and use techniques such as noise removal, image normalization, etc. Finally, our work looks to the future of deep learning in histopathology, exploring its role in therapeutic decision-making, predicting treatment outcomes and integrating histopathological and genetic data to increase results. The review can help researchers discover which deep learning techniques are most effective for a specific dataset and which features are significant for breast cancer detection.

Keywords Histopathology · Breast cancer · Deep learning · Classification · Gene expression

Abbreviations

A	Adenosis
ANN	Artificial neural network
ASO	Atom search optimization
B	Benign
BCMT	Breast cancer with molecular Typing
BHI	Breast histopathology images
CNAs	Copy number alterations
CPTAC	Clinical proteomic tumor analysis consortium
CNN	Convolutional neural network
DT	Decision tree
DC	Ductal carcinoma
EO	Equilibrium optimizer
F	Fibroadenoma
GDC	Genomic data commons
GNN	Graph neural network
H[NONSPACE	&E] Hematoxylin and eosin (staining)
IDC	Invasive ductal carcinoma
ITH	Intra-tumour heterogeneity
KNN	K-Nearest Neighbors
LC	Lobular carcinoma
MF	Magnification
M	Malignant
MC	Mucinous carcinoma
PC	Papillary carcinoma
PSO	Particle swarm optimization
PT	Phyllodes tumor
ST	Spatial transcriptomics
SVM	Support vector machine
QCNN	Quantum convolutional neural network

Oumeima Thaalbi and Moulay A. Akhloufi have contributed equally to this work.

✉ Oumeima Thaalbi
eot1123@umoncton.ca

Moulay A. Akhloufi
moulay.akhloufi@umoncton.ca

[1] Perception, Robotics and Intelligent Machines Lab (PRIME), Department of Computer Science, Université de Moncton, Moncton, NB E1C 3E9, Canada

TA	Tubular adenoma
TME	Tumor microenvironment
WDBC	Wisconsin diagnostic breast cancer
WXS	Whole exome sequencing
WGS	Whole genome sequencing
WSI	Whole slide image

1 Introduction

Breast cancer is a significant worldwide health concern and increasing treatment results and survival rates require early diagnosis. Traditional methods, such as histopathological examination, have limitations, particularly in terms of detecting cancer in dense breast tissue and its variability. In this work, the focus is on the use of histopathological images because the only method, that is certain to reveal the microscopic structure of the tissues and determine whether a tumor is benign or malignant, is a biopsy followed by histopathological analysis.

The histopathological images, which are often used to diagnose cancer, provide a detailed view of the illness. During a biopsy, a slice of tissue is stained and examined by a pathologist to differentiate benign tumors from malignant ones and, if necessary, grade the tumor (Rashmi et al. 2022). Pathologists' evaluations are less varied and subjective now that they use digital microscopes, but evaluating slides manually still takes a long time and can lead to mistakes. Consequently, there is an increasing need for automated systems using image processing and artificial intelligence to analyze tissue images. AI can analyze large datasets quickly, potentially identifying patterns that are not visible to the human eye.

The histopathological types of breast cancer can be divided into benign and malignant tumors. Benign tumors include fibroadenoma, phyllodes tumor and intraductal papilloma. Malignant tumors fall into two categories: non-invasive (in situ) and invasive carcinomas. Non-invasive carcinomas include intraductal carcinoma and lobular carcinoma in situ. Invasive carcinomas include invasive ductal carcinoma, invasive lobular carcinoma, medullary carcinoma, colloid (mucinous) carcinoma, papillary carcinoma, tubular carcinoma, adenoid cystic carcinoma, secretory carcinoma, inflammatory carcinoma and metaplastic carcinoma (Rashmi et al. 2022). Also, malignant breast cancers are classified into subtypes based on their molecular characteristics: luminal A, luminal B, HER2-positive and triple negative/basal-like (Thakur et al. 2024).

Recent advances in deep learning offer promising solutions for diagnosing breast cancer using histopathological images. Deep learning algorithms can analyze complex breast cancer data, improving diagnostic consistency. In this study, we highlight the application of deep learning on histopathological images, including WSI, to classify breast cancer and predict gene expression.

Effective data pre-processing is important for training deep learning models. Techniques such as image rotation and flipping, as well as contrast enhancement, can be used to create high-quality and diverse datasets for model training. For example, Shankar et al. (2022) discussed the use of rotations and flips to augment image data, introducing variability that imitates different patient conditions. Bagchi et al. (2022) highlighted the importance of image pre-processing techniques, such as stain normalization and contrast enhancement, to improve image quality.

Various deep learning models, including hybrid architectures that integrate ensemble learning, have been explored. Hybrid models, which combine several deep learning and traditional machine learning architectures, have proved particularly effective in dealing with the complexities of breast cancer images for binary classification (cancerous vs non-cancerous) and multi-class classification (identifying specific types of breast tumors) tasks (Bhowal et al. 2022; Bagchi et al. 2022). These models integrate features of different deep learning networks and machine leaning networks as ensemble learning to improve performance in classifying breast cancer from histopathological images. For instance, Clement et al. (2022) combined feature extractors from ResNet-50, EfficientNetB0 and Inception-v3, followed by SVM classification. Hassan et al. (2023) integrated features from DenseNet201, ResNet152, ResNet101, ResNet50, and InceptionV3 with various SVM classifiers. Karthik et al. (2022) combined CSAResnet (ResNet-101 with channel/spatial attention) and DAMCNN (DenseNet-201 with EfficientNet-B0) to improve performance, demonstrating the effectiveness of combined architectures in this domain. Another hybrid model was proposed by Zerouaoui and Idri (2022) using DenseNet201 as a feature extractor combined with an MLP classifier. Joseph et al. (2022) utilized handcrafted feature extraction followed by a DNN classifier, showing a simple yet effective hybrid approach. Al-Jabbar et al. (2023) employed a more complex hybrid model by integrating AlexNet feature extraction with handcrafted features, PCA for dimensionality reduction, and an ANN classifier. Other approaches have been used, such as dual-stream networks (Zou et al. 2022), several pathway CNNs (Ahmed and Islam 2023), attention mechanisms (Yu et al. 2023), lightweight separable convolution networks (Nneji et al. 2023), and YOLO architecture for cell nuclei localization (Bhausaheb and Kashyap 2023) which will be discussed later.

Besides simple image classification, deep learning is being used to predict gene expression patterns from WSI (He et al. 2020; Gao et al. 2023; Rahaman et al. 2023), which is indispensable for determining breast cancer molecular subtypes (Luminal A, Luminal B, HER2-enriched, and

Basal-like) (Phan et al. 2021; Liu et al. 2022) and guiding therapeutic decisions. Models such as GNNs have been used to analyze spatial relationships and patterns in tissue samples that correlate with gene activity, providing information without direct genetic testing (Phan et al. 2021; Gao et al. 2023). Moreover the vision transformers were held in the studies of Pang et al. (2021), Zeng et al. (2022), Yang et al. (2023).

Recent reviews have examined the application of artificial intelligence in breast cancer diagnosis, highlighting several methodologies. A common theme in these reviews is the potential of AI to significantly increase breast cancer diagnostic and prognostic capabilities, particularly in pathology, by automating and improving the accuracy of traditional methods. Tafavvoghi et al. (2024) focused on the identification of publicly available datasets containing H&E WSIs of breast histopathology. The idea is to provide an overview of these datasets, detailing their characteristics to help researchers select appropriate datasets for specific tasks in computational breast cancer pathology. Soliman et al. (2024) analysed the impact of the artificial intelligence on breast cancer detection and treatment in pathology. It highlighted its effectiveness in identifying invasive breast tumors and lymph node metastases, improving assessment of hormonal status, breast cancer grading and mitosis counts and showed how it helps assess tumor-infiltrating lymphocytes to predict triple-negative breast cancer outcomes and anticipate responses to neoadjuvant chemotherapy. Unger and Kather (2024) looked at how deep learning could transform precision oncology by automating the review of histopathological slides and genomic information. Deep learning models have demonstrated important results in detecting tumor tissue, predicting cancer subtypes, and identifying genetic alterations in histopathological images. Abo-El-Rejal et al. (2024) examined various imaging techniques, such as mammography, ultrasound, MRI, and histopathology, used to diagnose breast cancer. They addressed the limitations of conventional segmentation methods and discussed the transition between traditional approaches and AI-based techniques in medical imaging.

Our contribution can be seen in its unique structure in three main sections: breast cancer classification, prediction of gene expression from histopathological images and the databases corresponding to each task. This structure allows our study to be distinguished by focusing specifically on the analysis of histopathological images and WSIs. We systematically review studies relating to breast cancer classification. We look at the different classes and subtypes, providing detailed data collection and analysis of these studies. We focus on studies predicting gene expression. We examine in detail how these studies collect and analyze data, highlighting the characteristics used. Our article compiles the latest research on this specific topic, offering an insightful view into both breast cancer classification and the prediction of gene expression, making it a valuable resource for understanding recent advances in the field (see Fig. 1).

This review is structured to cover the integration of deep learning in breast cancer diagnostics. We highlight:

- Breast cancer histopathological images classification datasets: We investigate the examination of the diverse datasets used for training models in image classification and their details.
- Breast cancer genes expression datasets: We discuss datasets that are crucial for gene expression patterns associated with breast cancer.
- Deep learning methods for binary classification: We focus on models that classify histopathological images into binary categories, simplifying the problem to a cancerous versus non-cancerous class.
- Deep learning methods for multi-class classification: We study models that distinguish between benign and malignant samples as well as those that further classify tumor subtypes.
- Deep learning methods for genes expression prediction from WSIs and gene expression data: We analyse the state-of-the-art methodologies that use deep learning to predict gene expression from WSIs, offering insights into tumor biology and recurrence risks.

The studies and models evaluated in this work are derived from a review of various works. Our approach is designed to provide readers with a complete insight into the current capabilities and prospective developments of deep learning technologies in the fight against breast cancer. We define the abbreviations used in this paper in Table 1.

2 Methodology

2.1 Selection criteria

- **Time frame:** Initially, articles on breast cancer diagnosis using artificial intelligence published between 2020 and 2023 were selected. The time frame of articles on the use of histopathological images of breast cancer was limited to 2022–2023 on the classification task, while articles on gene expression and cancer biomarkers prediction from 2020 to 2023 were retained. This decision was based on the discovery of comprehensive article reviews that covered the topic of classification and this distinction also reflects the different stages of research maturity between the two areas, with classification being more established and gene expression prediction still evolving.

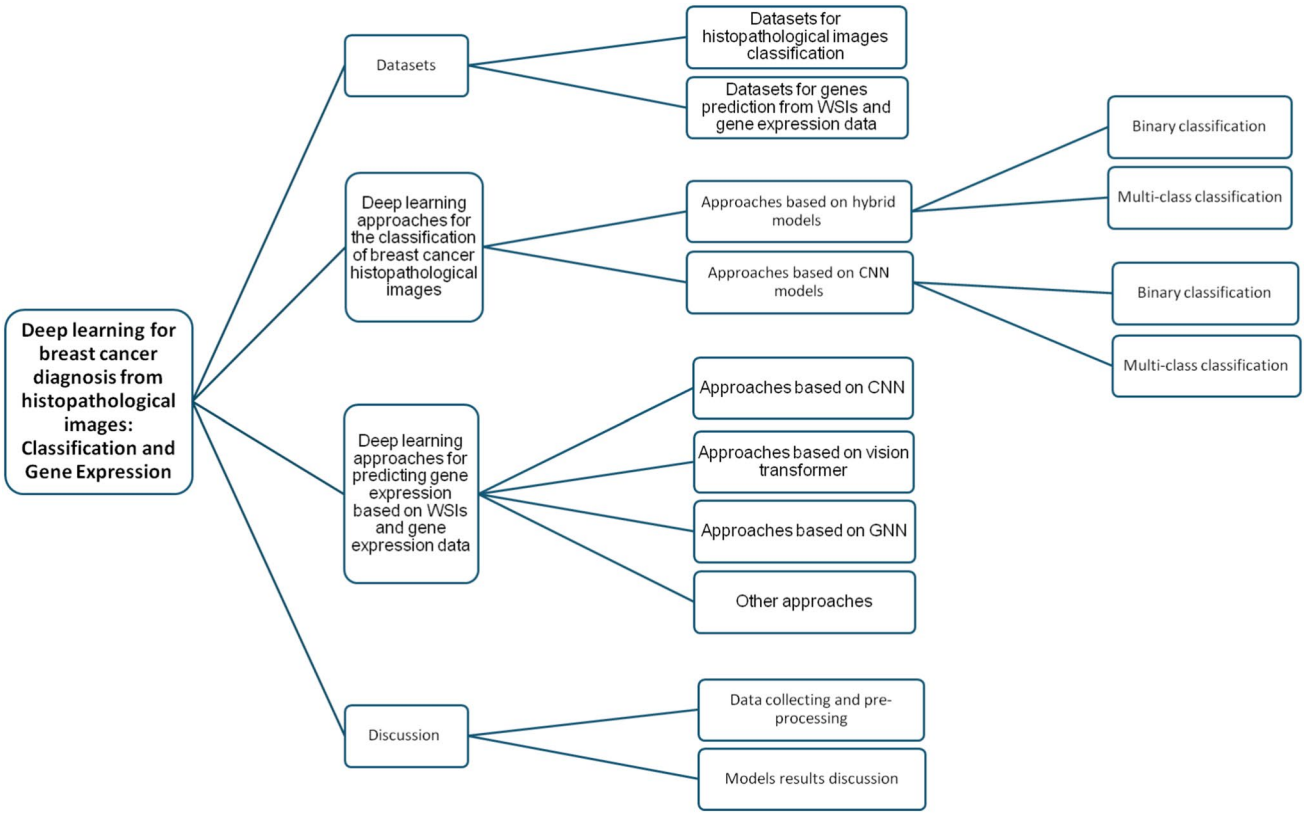

Fig. 1 Overview of the review structure for deep learning-based diagnosis of breast cancer from histopathological images and gene expression. The diagram details the primary components including datasets, deep learning approaches for image classification and gene expression analysis, and the discussions on model results and methodologies

- **Scope:** Focus on studies using deep learning for breast cancer diagnosis, including classification and prediction of gene expression from histopathological images in particular.
- **Database searches:** The Databases Google Scholar, PubMed, IEEE Xplore and ScienceDirect were searched using specific keywords: "breast cancer diagnosis", "deep learning", "classification", "gene expression", "histopathological images", "deep learning in breast histopathology", "breast WSI".

2.2 Selection steps

- Articles were selected by titles and then abstracts were reviewed to identify studies relevant to the research topic.
- Any documents outside of the specified research areas were not considered.
- Articles that did not focus on breast cancer were excluded.
- Only English-language articles published in the specified period were included.
- Duplicates studies were excluded.
- Only articles presented in this work were considered Fig. 2.

3 Datasets

In the deep learning applied to breast cancer imaging, access to diverse and well-annotated datasets is important for developing models that are reliable. In this section, we will detail the datasets used to advance our knowledge of breast cancer through two main perspectives: histopathological image classification in Table 1 and gene expression from WSIs in Table 2. We will first present datasets focused on histopathological images of breast cancer classification, wherein our objective is to describe how these images help distinguish between different pathologies and stages of the disease with the various classes represented in the datasets. Next, we will investigate datasets related to breast cancer gene expression, highlighting their importance in identifying specific genetic profiles. These datasets not only include histopathological images that offer insights into the morphological variations of cancerous tissues but also encompass rich genetic expression data that

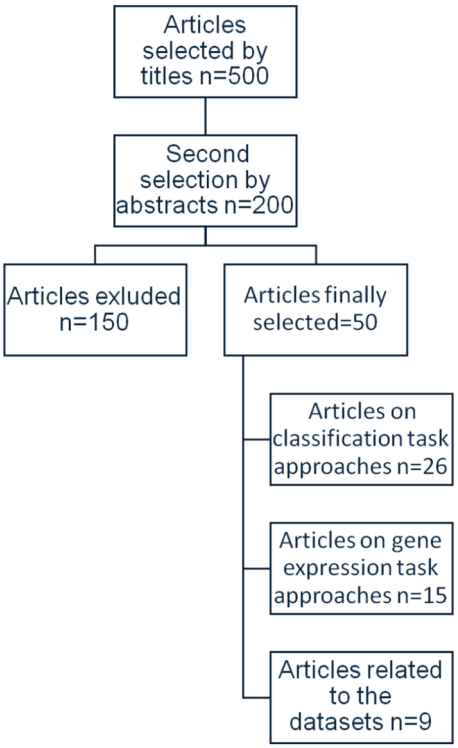

Fig. 2 PRISMA diagram: Our literature review selection process

reveal underlying molecular mechanisms. This section is describing these datasets and their distribution as shown in Tables 1 and 2, providing a strong basis for study in breast cancer research.

3.1 Datasets for histopathological images classification

Histopathological images are used for the classification of breast cancer. A number of databases provide access to histopathological image for the advancement of this area study. In this first part, we will discuss the datasets used in the classification of different types of breast cancer (Table 1).

- ***BreakHis dataset:*** The BreaKHis dataset is publicly accessible and utilized for researching breast cancer classification issues. It was introduced in Spanhol et al. (2015), wherein it comprised 7909 microscopic samples generated from breast tissue biopsy slides. These samples were collected by SOB method from 82 patients using varying magnifications of 40×, 100×, 200× and 400×. Each image in the dataset measured 700×460 pixels and is in PNG format. These images feature have an 8-bit depth per channel and are composed of a 3-channel RGB color model. The collection took place from January 2014 to December 2014 in collaboration with the P&D Laboratory-Pathological Anatomy and Cytopathology, in Parana, Brazil. The dataset contains images of 2480 benign and 5429 malignant tumors. The benign tumors were categorized into four sub-classes: A, TA, PT, and F. Similarly, the malignant tumors were divided into four distinct sub-classes: DC, MC, LC, and PC Fig. 3. These samples were stained with H&E and were labeled by pathologists from the P&D Lab. All data in the dataset were anonymized and were split into two parts: 635 images for training and 550 images for testing. In the test set, there are 250 benign and 300 malignant images. The dataset is publicly accessible and frequently utilized for researching breast cancer classification issues.
- ***ICIAR 2018 dataset:*** The public ICIAR 2018 dataset (https://iciar2018-challenge.grand-challenge.org), part of the BACH (BreAst Cancer Histology images) challenge (Aresta et al. 2019), extended from the Bioimaging 2015 (http://www.bioimaging2015.ineb.up.pt/dataset.html). It's available under the CC BY-NC-ND license. This collection included high-resolution 2048×1536 pixels, H&E stained breast histology images of 24-bit RGB format. The images, extracted from WSIs biopsies, had pixel size of 0.42 µm × 0.42 µm, and pixel values ranging from 0 to 255. They were acquired at 200× magnification and stored in the TIFF Format. The BACH dataset was divided into two parts, A and B. Part A featured 400 annotated microscopy images of breast tissues wherein they were categorized into four classes: Normal tissue,

Table 1 Summary of histopathological image datasets for breast cancer classification

Dataset	Size	Classes	MF	Format	Access	Patient count	Patient gender
BreakHis (Spanhol et al. 2015)	7909 images	8	40×, 40×, 40×, 40×	PNG	Public	82	N/A
ICIAR 2018 (Aresta et al. 2019)	400 images	4	40×	TIFF	Public	N/A	N/A
KMC (Rashmi et al. 2023)	1516 images	2	40×	N/A	Private	N/A	N/A
BHI (Cruz-Roa et al. 2014; Janowczyk and Madabhushi 2016)	162 images	2	40×	N/A	Public	162	Female
BreCaHAD (Aksac et al. 2019)	23549 annotations	6	40×	PNG	Public	N/A	N/A
WDBC (Vig 2014)	569 images	2	N/A	N/A	Public	N/A	N/A

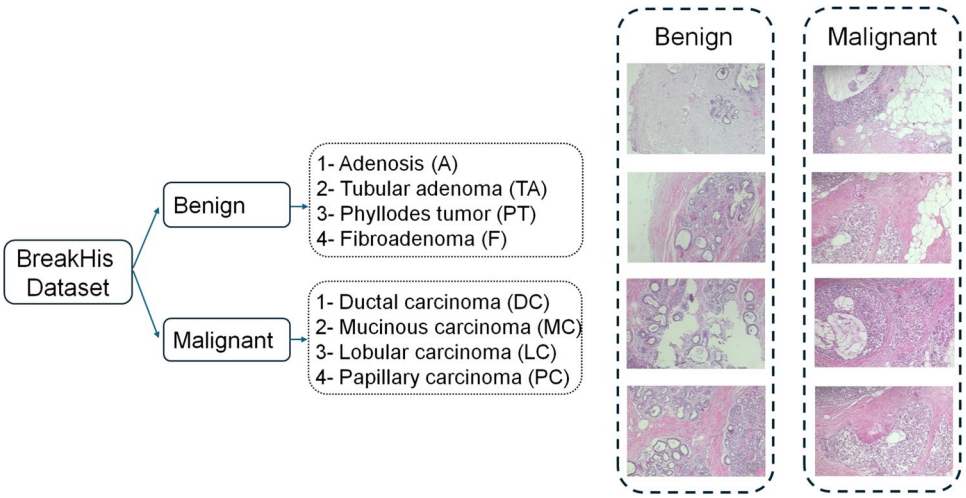

Fig. 3 Representative histopathological images from the BreakHis dataset categorizing benign and malignant breast tumors

Benign tissue, In-situ carcinoma, and Invasive carcinoma, with the ICIAR dataset containing 100 images per category. Part B included WSIs of the same classes, intended for evaluating pixel-wise labeling performance. Annotation was carried out by two medical experts and any images with disagreements between the experts were excluded from the dataset.

- ***KMC dataset:*** The KMC dataset was utilized in Rashmi et al. (2023) as a private dataset and was manually collected from the pathology department of Kasturba Medical College (KMC) in Manipal, India. It consisted of 1516 images, each magnified 400× and with a resolution of 1600×1200 pixels. The images, captured using an OLYMPUS CX31 microscope, primarily focused on F and DC subtypes. A domain expert annotated the dataset, which included 759 benign and 757 malignant cases. Subsequently, the dataset was divided into training and testing sets, comprising 916 and 600 images respectively. The test set, selected randomly, contained an equal number of benign and malignant images, totaling 300 each.
- ***BHI dataset:*** The BHI dataset is a public dataset (https://www.kaggle.com/datasets/paultimothymooney/breast-histopathology-images), consisted of 162 whole mount slide images of breast cancer from women who were diagnosed with IDC at the Hospital of the University of Pennsylvania and The Cancer Institute of New Jersey (Cruz-Roa et al. 2014; Janowczyk and Madabhushi 2016). These slides were digitized at a 40× magnification, providing a resolution of 0.25µm per pixel. This dataset was then divided into three distinct groups: 84 slides were used for training purposes (referred to as D1), 29 slides for validation (D2), and the remaining 49 slides were reserved for final testing (D3).
- ***BreCaHAD dataset:*** This dataset (Aksac et al. 2019), first published online on January 28, 2019, included a diverse array of files formats (xlsx, json, png, zip). The annotation details file (xlsx) presented the distribution of around 23549 annotations across six categories: mitosis, apoptosis, tumor nuclei, non-tumor nuclei, tubule, and non-tubule, all laid out in an excel spreadsheet. It included two .png files: 'original.png', which was the input image, and 'annotated.png', an example from the dataset presenting tumor nuclei marked in blue, non-tumor nuclei like blood cells and lymphocytes in pink, mitosis, and apoptosis in orange and green, tubules in light blue, and non-tubule regions such as fat and blood vessels in yellow. The 'data.json' file contained JSON-format annotations for the BreCaHAD dataset, providing examples with coordinates for annotated objects like mitosis and tumor nuclei. Finally, the 'BreCaHAD.zip' archive contained the dataset, organized into three folders: 'images' for original images, 'groundTruth' for JSON files, and 'groundTruth_display' for images with applied ground truth.
- ***WDBC dataset:*** The WDBC dataset is a resource for breast cancer research, focusing on the analysis of digitized images from fine needle aspirates (FNA) of breast masses. It encompasses a detailed set of features describing the characteristics of cell nuclei within these images. The dataset included data on 569 instances, with no missing attribute values, and was balanced with 357 benign and 212 malignant cases, labeled 'B' for benign and 'M' for malignant (Vig 2014). Available through the University of Wisconsin's FTP server and the UCI Machine Learning Repository (Wolberg William and Mangasarian Olvi 1995; Vig 2014).

Table 2 Summary of breast cancer genes expression datasets

Dataset	Size	Data details	Patient details
TCGA-BRCA (Brigham & Women's Hospital & Harvard Medical School Chin Lynda et al. 2012)	1098 cases	WSIs, genomic data, clinical data, 60,945 files, 6925 annotations	Patients with Luminal A, Luminal B, Triple-Negative, and HER2
TMA cohort (Mondol et al. 2023)	485 cases	WSIS, Luminal A and B subtypes, TIFF format, 40× magnification with 0.25µm/pixel resolution	n = 405 with ER+ invasive luminal tumors, subtyped into n = 309 luminal A, and n = 96 luminal B
CPTAC (Rudnick et al. 2016)	1500 patients	WSIs data, WGS, WXS, RNA sequences, proteomic data	N/A
ClinSeq-BC, SoS-BC-1, SCAN-B (Wang et al. 2022)	256 + 1358 + 1262 patients	WSIs, RNA-seq data, clinical data from the Swedish National Breast Cancer Registry	Female patients with an initial diagnosis of primary invasive breast cancer
STNet dataset (He et al. 2020)	23 patients	WSIs, ST data, 30612 spatially mapped gene expressions, 26949 distinct mRNA species	n = 12 with Triple-Negative, n = 12 with Luminal A, n = 15 with Luminal B, n = 14 with HER2-luminal, n = 15 with HER2-non-luminal
HER2+ (Andersson et al. 2021)	8 patients	WSIs, ST data	n = 36 with HER2+
BCMT (Liu et al. 2022)	1254 patients	WSIs with annotations for molecular subtypes: Luminal A, Luminal B, Her-2 overexpression, Basal-like	n = 313 with Luminal A, n = 382 with Luminal B, n = 316 with HER2, and n = 243 with Basal-like

3.2 Datasets for genes prediction from WSIs and gene expression data

In gene expression area, several platforms have made significant contributions, offering varied datasets. These include the Genomic Data Commons and 10x Genomics. The GDC, a national initiative, serves as a centralized portal for the collection, preservation of genomic, and clinical data. It offers a variety of cohorts such as TCGA-BRCA (Brigham & Women's Hospital & Harvard Medical School Chin Lynda et al. 2012, 2024) and CPTAC (Rudnick et al. 2016), including gene expression data related to multiple cancer types, in particular breast cancer with 4002 female and 32 male cases. For another, 10x Genomics provides a platform for spatial gene expression, enabling visualization of gene expression in the spatial tissues. As part of this study, we present datasets from several GDC and 10x Genomics sources and others, which are used in the latest studies mentioned in this article (Table 2).

- ***TCGA-BRCA:*** The TCGA-BRCA project (Brigham & Women's Hospital & Harvard Medical School Chin Lynda et al. 2012, 2024) was part of the larger TCGA initiative https://www.cancer.gov/ccg/research/genome-sequencing/tcga, which was established by the National Cancer Institute (NCI) and the National Human Genome Research Institute (NHGRI) in 2006, conducted detailed molecular profiling on more than 20000 primary cancer cases and their corresponding normal samples across 33 distinct cancer categories. The TCGA-BRCA project focused specifically on breast cancer and comprised 1098 cases, 60,945 files and 6925 annotations. The data contained 7 categories: sequencing reads, transcriptome profiling, simple nucleotide variation, copy number variation, DNA methylation, clinical, and biospecimen.
- ***TMA cohort:*** The TMA cohort, used Mondol et al. (2023), was derived from the St.George Breast Boost randomized radiotherapy clinical trial, conducted between 1996 and 2003, which included n = 485. This cohort consisted of 405 ER+ invasive luminal tumors, subtyped into 309 luminal A, and 96 luminal B categories, according to immunohistochemistry criteria (Goldhirsch et al. 2013). During the construction, of the TMAs tumours were examined on an H&E slide by a breast pathologist who marked appropriate regions for sampling. Then, these sections were scanned at 400× magnification with 0.25µm/pixel, saved as .TIF files and checked that every part was not blurry (Millar et al. 2020).
- ***CPTAC:*** The CPTAC dataset (Rudnick et al. 2016) was initiated by the National Cancer Institute (NCI) in 2006 and launched in 2011 to find and list proteins that come from changes in cancer genomes and how these

changes work https://gdc.cancer.gov/about-gdc/contributed-genomic-data-cancer-research. This initiative is commonly known as CPTAC2 (Rudnick et al. 2016) https://wiki.cancerimagingarchive.net/display/Public/CPTAC+Imaging+Proteomics. In 2016, the NCI made an expansion to the CPTAC program, which is known as CPTAC3 https://proteomics.cancer.gov/programs/completed. They also make this data available to everyone. CPTAC has submitted to the Genomic Data Commons genetic data from over 1500 patients suffering from various types of cancer, including specifically Breast (CPTAC-BRCA), brain, lung adenocarcinoma and squamous cell carcinoma, endometrial, pancreatic, ovarian, renal, head and neck, and colon cancers. DNA sequences from CPTAC, including WGS, WXS, and RNA sequences, were aligned with the GRCh38 reference genome through the use of GDC DNA-Seq analysis pipelines and mRNA analysis pipelines, respectively. This aligned genomic data from CPTAC is accessible on the GDC Data Portal. Proteomic data processed by the CPTAC Common Data Analysis Pipeline (CDAP) were made accessible on the CPTAC Data Portal. In addition, this proteomic data from CPTAC can be found in the Proteomic Data Commons (PDC). The dataset comprised data obtained from samples aimed at validating findings from CPTAC in TCGA samples. These samples, encompassing breast, ovarian, colon andlLung tumors, were collected following a protocol adapted for proteomics analysis.

- ***ClinSeq-BC/SöS-BC-1/SCAN-B: The ClinSeq-BC*** dataset (Wang et al. 2022) derived from a study that included 256 patients treated at Stockholm South General Hospital in 2012 or at Karolinska University Hospital from 2001 to 2008. The ***SöS-BC-1*** cohort (Wang et al. 2022) had 1358 patients diagnosed at Stockholm South General Hospital from April 2012 to October 2014 and from October 2015 to May 2018. The ***SCAN-B-Lund*** subset (Wang et al. 2022) involved 1262 patients from Lund, diagnosed with breast cancer from 2010 to 2019. RNA-seq data was available for ***The ClinSeq-BC*** patients. clinical data for patients from ClinSeq-BC, ***SöS-BC-1***, and ***SCAN-B-Lund*** subset came from the Swedish National Breast Cancer Registry.
- ***The human breast cancer in situ capturing transcriptomics dataset:*** This dataset referred to STNet dataset (He et al. 2020), it was created from samples provided by 23 patients with the disease and contained 30612 gene spatially resolved gene expression data correspond to histopathology images. For each individual, three microscope images of stained tissue along with data detailing gene locations within the tissue were obtained. A total of 68 sections were divided as follows: n = 12 Triple-Negative, n = 12 Luminal A, n = 15 Luminal B, n = 14 HER2-luminal, and n = 15 HER2-non-luminal. Gene activity in small areas, each 100µm and spaced 200µm, was analyzed. Throughout the dataset, 26,949 distinct mRNA species were identified by researchers. As a result, each spot was represented by a 26,949-dimensional vector composed of non-negative integers.
- ***HER2+ dataset:*** The HER2+ dataset (Andersson et al. 2021) included samples from 8 patients. A total of 36 sections were collected from the tumors, with either three sections placed close together or six sections spaced evenly apart. The analysis was performed using ST technique. Initial data characterization included manual annotation by a pathologist, who identified areas such as carcinoma in situ, invasive carcinoma, adipose tissue, immune infiltration, and connective tissue. The spatial gene expression data were then normalized and visualized in 2D space using UMAP. The spatial acquisition locations (called spots) showed patient-specific clustering, indicating heterogeneity between patients. To accurately capture the molecular signature of each patient and ensure that weak signals were not missed, the data for each patient were analyzed individually.
- ***BCMT dataset:*** The BCMT dataset, given by Xiangya Hospital, included 1254 pathological WSIs from 1254 patients or cases, gathered between 2017 and 2019. It had slide-level annotations for breast cancer molecular subtypes offering 313 slides for Luminal A, 382 for Luminal B, 316 for Her-2 overexpression, and 243 for Basal-like subtype. The pathology WSIs used a pyramid storage structure to facilitate access and analysis of these high-resolution images (Liu et al. 2022).

4 Deep learning approaches for the classification of breast cancer histopathological images

The classification of breast cancer histopathological images (Fig. 4) can include simple binary classification as shown in Table 3, which separates benign (non-cancerous) from malignant (cancerous) conditions. It also involves multi-class classification as shown in Table 4, which identifies specific sub-types within benign and malignant categories, such as various types of tumors or carcinomas.

In this section, we highlight the methods employed to achieve binary and multi-class classification.

4.1 Approaches based on hybrid models

4.1.1 Binary classification

Al-Jabbar et al. (2023) proposed two approaches that were applied on BreakHis dataset with magnification of 40×

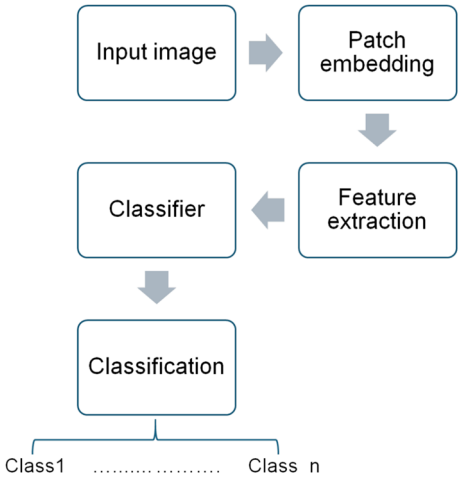

Fig. 4 Overview of the general image classification pipeline in most research

,100×,200×, and 400× to diagnose benign and malignant cells. The first was based on an hybrid technology CNN that used AlexNet and GoogleNet to extract features and save them in features vectors, then the classification was made by the SVM algorithm. As a result of using AlexNet+SVM with 40×, 100×, and 200× magnification dataset, the best accuracy was 98.8%. However, GoogLeNet+SVM technique outperformed AlexNet+SVM on the 400× magnification dataset with an accuracy of 96.7%. The second method combined CNN features (extracted by AlexNet and GoogLeNet) and handcrafted features (using a fusion features which are derived from fuzzy color histogram (FCH), local binary pattern (LBP) and gray level co-occurrence matrix (GLCM)). A PCA algorithm has been used to reduce the dimensions of the deep features by CNN and extract the most important features of each histological image. The merged features were classified afterwards with an ANN algorithm. The accuracy, precision, sensitivity, specificity, and AUC of the ANN algorithm based on hybrid features between AlexNet with handcrafted features were all 100% using images with a magnification of 400×. However, features extracted by GoogLeNet merging with handcrafted features gave 99.8% of accuracy with 200× magnification dataset.

In this work (Obayya et al. 2023), an arithmetic optimization algorithm with deep-learning based on histopathological breast cancer classification (AOADL-HBCC) technique has been developed by Obayya et al. (2023) to detect benign and malignant classes. It used a contrast enhancement process and a median Filter (MF) for noise removing in the image pre-processing. The features have been extracted with a SqueezeNet model and a binary classification was realized by deep belief network model (DBN) with an Adamax hyperparameter optimizer. A benchmark dataset was used with 100× dataset and 200× dataset magnification. With this proposed system, the highest accuracy achieved was 96.77% on the 200× images.

In their study (Clement et al. 2022) employed a machine learning approach for binary classification. This method utilized a bag of deep multi-resolution convolutional features (BoDMCF) extracted from histopathological images of the BreakHis dataset at four different resolution (40×, 100×, 200×, 400×) using pre-trained CNN models, including ResNet-50, EfficientNetb0, and Inception-v3. Then the BoDMCF were pooled using global average pooling and classified using the SVM classifier. An average accuracy of 99.92% was achieved.

Zou et al. (2022) presented a dual-pathway high-order neural network architecture termed DsHoNet. Specifically, a shallow convolutional neural network comprised six layers served as the backbone of this architecture. The first stream incorporated batch normalization to maintain original information with clearer representations, while the second stream integrated ghost modules to extract more detailed features through a variety of linear transformations. The feature sets derived from the two streams were combined and directed into a second-order pooling layer called MPN-COV (Li et al. 2017), facilitating the acquisition of information with high-order statistical characteristics. This process enabled the extraction of global features via the covariance pooling module. The classification of benign and malignant breast cancer on the BreakHis dataset was performed using three fully connected layers, demonstrating that DsHoNet achieved accuracies of 99.01% and 99.25% at the image and patient levels, respectively.

In their research, Atban et al. (2023) utilized a ResNet18 architecture, a pre-trained transfer learning network, to extract deep features. To tackle the issue of high dimensionality without losing information, they employed meta heuristic algorithms like PSO (Hu et al. 2022), ASO (Zhao et al. 2019a, b) and EO (Gao et al. 2020) to select a minimal subset of the most representative features. They employed traditional machine learning algorithms to evaluate the effectiveness of the optimized deep features in binary classification including KNN (Nayak et al. 2022), SVM (Cortes and Vapnik 1995) and DT (Kim and Kim 2022). The proposed method was assessed on the BreakHis dataset achieving an accuracy of 97.73% for 200× magnification using SVM with Gaussian and Radial-Based Functions (RBF) (Sharma and Kumar 2022).

In this research, Abbasniya et al. (2022), an approach called IRv2-CXL was proposed. It utilized feature vectors extracted from pre-trained CNNs, with a specific emphasis on the Inception-ResNet-v2 (Szegedy et al. 2017), a network that effectively combined the benefits of both residual and inception networks. Among sixteen diverse pre-trained networks evaluated in this study (Abbasniya et al. 2022),

Table 3 Summary for approaches used in binary classification of the breast cancer histopathological images: (Benign-Malignant classes (B-M), Fibroadenoma-Ductal Carcinoma subclasses (F-DC), IDC(+)-IDC(−) subclasses (IDC(−)-IDC(+)))

References	Methodology	Datasets	Acc %	Prec %	Rec %	Spec %	F1 %	AUC %	MF	Classes
Al-Jabbar et al. (2023)	AlexNet feature extractor + Handcrafted feature + PCA algorithm + ANN classifier	BreakHis	100	100	100	100	N/A	100	400×	B- M
Obayya et al. (2023)	SqueezeNet feature extractor + DBN classifier + Adamax optimizer	Benchmark	96.77	N/A	96.88	96.88	95.85	N/A	200×	B- M
Clement et al. (2022)	BoDMCF feature extractor (ResNet-50, EfficientNetb0, Inception-v3) + Global average pooling + SVM classifier	BreakHis	99.92	99.87	99.87	97.97	99.87	99.90	40×+ 100×+ 200×+ 400×	B- M
Hassan et al. (2023)	Combined feature extractors (DenseNet201 + ResNet152 + ResNet101 + ResNet50 + InceptionV3) + SVM classifier (L-SVM, G-SVM, S-SVM)	BreakHis	99.24	N/A	N/A	N/A	N/A	N/A	40×+ 100×+ 200×+ 400×	B- M
Karthik et al. (2022)	Ensemble of CSAResnet (ResNet-101 + channel/ spatial attention) and DAMCNN (DenseNet-201 + EfficientNet-B0)	BreakHis	99.55	99.00	99.00	N/A	99.00	N/A	40×+ 100×+ 200×+ 400×	B- M
Zou et al. (2022)	Dual-stream network: (stream 1: batch normalization and stream 2: ghost module) + Second-order pooling layer + 3 fully connected layers	BreakHis	99.25	99.77	99.08	99.40	99.42	99.00	200×	B- M
Zhang et al. (2022)	CNN of 10 layers feature extractor within the nucleus region + Three-channel feature fusion classification with softmax classifier	BreakHis	99.67	N/A	N/A	N/A	N/A	N/A	40×+ 100×+ 200×+ 400×	B- M
Majumdar et al. (2023)	Fuzzy rank-based ensemble (GoogleNet + VGG11 + MobileNetV3_Small) + Gamma function	BreakHis ICIAR-2018	99.16	99.40	98.67	N/A	99.02	98.67	40×	B- M
Jakhar et al. (2024)	Ensembling with stacked classifiers: base learners (Extra tree, Random forest, AdaBoost, Gradient boosting, KNN9) Final estimator: logistic regression	BreakHis WDBC	98.80	99.09	99.09	N/A	99.09	97.45	N/A	B- M

Table 3 (continued)

References	Methodology	Datasets	Acc %	Prec %	Rec %	Spec %	F1 %	AUC %	MF	Classes
Atban et al. (2023)	ResNet18 feature extractor + Feature selection meta heuristic algorithms (PSO, ASO, EO) + SVM classifier(gaussian and RBF)	BreakHis	97.73	97.75	97.75	N/A	97.75	N/A	200×	B- M
Abbasniya et al. (2022)	Inception-ResNet-v2 feature extractor + Ensemble classifier (CatBoost + XGBoost + LightGBM)	BreakHis	97.01	N/A	N/A	N/A	97.87	N/A	200×	B- M
Zerouaoui and Idri (2022)	DenseNet201 feature extractor + MLP classifier	BreakHis	93.93	93.10	94.97	N/A	93.98	N/A	200×	B- M
Nakach et al. (2022)	InceptionV3 feature extractor + XGB classifier with DT (200 trees) integrated	BreakHis	92.52	90.71	95.36	N/A	92.93	N/A	40×	B- M
Taheri and Golrizkhatami (2023)	MSB classification + DenseNet201	BreakHis	96.70	N/A	N/A	N/A	N/A	N/A	200×	B- M
Rashmi et al. (2023)	Combined CNN feature extractor with patch level and image level module + Colour channel attention module + Ensemble classifier (pool of 10 DT)	BreakHis KMC	96.00	95.50	95.50	N/A	95.50	N/A	200×	F-DC
Singh et al. (2023)	Deep QCNN (4 residual network blocks)	BHI	97.20	98.00	96.00	98.00	99.00	97.00	40×	IDC(-)- IDC(+)
Bagchi et al. (2022)	Patch feature extractors (VGG-19+ VGG-16+ Inception-ResNet v2) + Patch classification model (KNN+ SVM+ RF+ Adaboost+ XGBoost) + Image classification	ICIAR-2018	98.60	N/A	N/A	N/A	N/A	N/A	200×	Cancerous- Non-cancerous
Yu et al. (2023)	CDT based on SECS + Coordinate attention mechanism	BreakHis	100	100	100	N/A	100	N/A	40×+ 400×	B- M
Ahmed and Islam (2023)	Four-pathway CNN with EfficientNet-B0 as the backbone network	BreakHis	99.52	98.87	100	N/A	99.43	99.05	40×+ 100×+ 200×+ 400×	B- M
Bhowal et al. (2022)	Bottleneck feature extraction approach with DCNNs (VGG16, VGG19, Xception, InceptionV3, InceptionResnetV2)+ MLP classifier with choquet fuzzy integral	ICIAR-2018	96.00	96.00	96.00	N/A	N/A	N/A	200×	Carcinoma- Non-carcinoma
Ukwuoma et al. (2022)	Ensemble learning (DenseNet201+VGG16) + Multiple self-attention heads+ Multi-linear perceptron	BreakHis	99.80	N/A	N/A	N/A	N/A	N/A	N/A	B- M

the Inception-ResNet-v2 demonstrated the most deep feature extraction capability. For the classification of BreakHis images, this approach employed an ensemble of three fine-tuned classifiers: Categorical Boosting (CatBoost) (Abbasniya et al. 2022), Extreme Gradient Boosting (XGBoost) (Chelgani et al. 2021), and Light Gradient Boosting Machine (LightGBM) (Nasiri et al. 2022; Ezzoddin et al. 2022), with Soft-Voting applied (Gupta and Rani 2020). By implementing class activation mapping, IRv2-CXL achieved an accuracy rates of 96.83%, 95.84%, 97.01%, and 96.16% on the 40×, 100×, 200×, and 400× magnification factors, respectively.

The work of Zerouaoui and Idri (2022) introduced and assessed 28 hybrid frameworks that combined deep learning techniques for extracting features with traditional machine learning classifiers for binary classification. These architectures incorporated 7 distinct deep learning models (DenseNet201, InceptionV3, InceptionResNetV2, MobileNetV2, ResNet50, VGG16, VGG19) along with 4 classifiers (MLP, SVM, DT, KNN). The Scott Knott (Worsley 1977) statistical test was employed to group architectures with similar performance, while the Borda (Emerson 2013) count method ranked them based on the four criteria (accuracy, precision, recall, and F1-score). The findings indicated that architectures utilizing DenseNet201 for feature extraction and MLP for classification achieved the highest accuracy, with results of 92.61%, 92%, 93.93%, and 91.73% on BreakHis (Spanhol et al. 2015) images at 40×, 100×, 200×, and 400× magnification levels, respectively. The hybrid architecture that featured DenseNet201 and MLP (MDEN) secured the top rank overall, surpassing the other structures. The second-best was SVM with DenseNet201 (SDEN). KNN with DenseNet 201 (KDEN) secured the third position on BreakHis. For another, the DT classifier generated the least effective architectures.

Nakach et al. (2022) examined the utilization of hybrid models for the classification of the images sourced from the BreakHis dataset. This integration involved the fusion of deep learning and boosting techniques. The employed deep learning strategies encompassed the application of three distinct methods: DenseNet201, MobileNetV2, and InceptionV3, each serving to extract features from the images. Simultaneously, boosting was harnessed for image classification, incorporating four boosting methodologies, namely AdaBoost (ADB), Gradient Boosting Machine (GBM), LightGBM (LGBM), and XGBoost (XGB). Within the framework of boosting, DTs were integrated. The models were subjected to testing with varying numbers of trees (50, 100, 150, 200) and a tendency toward improved performance with a higher number of trees was observed. To select the best models and establish their rankings, the Scott Knott, five-fold cross-validation, and Borda methods were employed. In terms of results, the most effective boosting ensemble attained an accuracy rate of 92.52%. This accomplishment was realized by employing XGBoost with 200 trees, combined with InceptionV3 as the feature extractor.

The hybrid method presented by Rashmi et al. (2023), called CWA-NET, combined CNN-based feature extraction with an ensemble classification approach based on a dynamic classifier selection (DCS) (Woods et al. 1997) for the discrimination of F and DC subtypes. Two pools of classifiers were used, the first utilized 10 DT while the second utilized SVM, DT, RF, and Gaussian Naive Bayes. The evaluation of this method was conducted on two datasets: the KMC dataset and the BreakHis dataset, wherein the focus is on the 400× images. To provide effective training of these high resolution images, the authors applied the Reinhard method (Reinhard et al. 2001) to normalize the color distribution of the input images. To preserve the information contained in the original images without loss, a two-module feature extractor was designed, comprising patch-level and image-level feature extraction. These features were then concatenated to obtain the final feature vector representing the input image. In both the patch-level and image-level feature extractors, a color channel attention module was employed, inspired by the methodology proposed in Fu et al. (2019). This module took multi-color channels as input and generates attention masks to refine the features and regions of interest, such as nuclei, which is necessary in malignancy determination due to their distinctive characteristics and structure. The performance of the proposed model was evaluated on the KMC and BreakHis datasets, resulting in accuracy rates of 95% and 96%, respectively. In addition, the model achieved Kappa scores of 90% and 91% for the respective datasets (Table 3).

4.1.2 Multi-class classification

Bagchi et al. (2022) developed a deep learning-based method. Firstly, the images were divided into 35 patches, each with a dimension of 512 × 512 × 3 and assigned the same label as the original image they belong to. These patches underwent pre-processing steps including stain normalization, regularization, and augmentation techniques. Secondly, a feature extraction model based on VGG-19, VGG-16, and Inception-ResNet v2 architectures was utilized. A patch classification model was developed to classify the image patches into four classes: normal, benign, in situ carcinoma, and invasive carcinoma, using for this purpose classification algorithms such as KNN, SVM, Random Forest, Adaboost, and XGBoost. Thirdly, the deep learning-based model utilized the output of the patch classification model to classify images into two classes: cancerous and non-cancerous. Then, the 2-class classification probabilities were used to further classify the

images into the four additional classes: normal, benign, in situ carcinoma, and invasive carcinoma. Bagchi et al. (2022) adopted the ICIAR-2018 dataset and the results showed an accuracy of 97.50% on image-level and an accuracy of 82.50% on patch-level for the multi-class classification. For the binary classification task, the images accuracy achieved was 98.60%.

Shankar et al. (2022) introduced the CSSADTL-BCC model. They employed a gaussian filtering (GF) technique to eliminate noise. A set of feature vectors was generated by a MixNet-based feature extraction model which is a sequence of lightweight convolutional operators and depthwise separable convolutional layer. Then, a stacked gated recurrent unit (SGRU) classification approach with adjusted hyperparameters, composed of various GRU units, was used to assign class labels. To evaluate the model's effectiveness, the BreakHis dataset was employed and after training for 1500 epochs, the CSSADTL-BCC methodology yielded average accuracy reaching approximately 98.61% for the 8-class classification.

Deer-Canid based deep CNN, incorporating the combined characteristics of DHO (Tian et al. 2020) and GWO's (Nadimi-Shahraki et al. 2021) encircling behavior, was proposed by (Bhausaheb and Kashyap 2023). Initially, localization was performed, wherein the YOLO architecture accurately determined the position of cell nuclei using probability charts, without the need for a nuclei class. The segmentation process was employed through the V-Net architecture to extract the relevant regions. The classification of these regions was carried out utilizing the Deer Canid optimization-based deep CNN, resulting in the categorization of the output as normal, benign, or malignant. The input data was sourced from the BreCaHAD and BHI datasets. On the BreCaHAD dataset, the achieved accuracy was 92.96%. Similarly, on the BHI dataset, the corresponding accuracy was 93.05% for the 40× magnification.

Yu et al. (2023) incorporated a coordinated attention mechanism (Hou et al. 2021) to improve the DenseNet architecture,. This mechanism, known as a spatial domain attention module (Chaudhari et al. 2021; Niu et al. 2021; Usama et al. 2019), was employed to refine the network's feature extraction capabilities for challenging image textures across different magnifications. They developed the CA-BreastNet model, which served as the fundamental classifier by combining a CNN network with seven types of convolutional decision trees (CDT) using a specialized enhanced classifying strategy (SECS). This integration raised the network's efficiency by mitigating the limitations imposed by dataset structures on its accuracy. The CDT achieved on the BreakHis dataset accuracy of 100% for binary classification at the 40× and 200× magnifications, 96.43% for eight-class classification at 400× magnification. By utilizing focal loss (Lin et al. 2020), the network focused on challenging hard samples particularly LC and DC, resulting in improved the stability.

In their investigation, Ahmed and Islam (2023) adopted the Multiple Instance Learning (MIL) (Quellec et al. 2017) concept to address bags consisting of four images at varying magnifications. Instead of the conventional MIL assumption, they employed the feature-vector to infer the bag's label. To achieve this, a DCNN model with four input pathways was deployed wherein the EfficientNet-B0 was as backbone network. The BreakHis images were categorized into binary, 4-class, and 8-class classes. This model achieved an accuracy scores of 99.52% for binary classification, 99.65% for 4-class classification, and 99.04% for 8-class classification.

The study of Bhowal et al. (2022) proposed an approach that classified breast cancer into 2 classes (carcinoma vs non-carcinoma) or 4 classes (normal, benign, in situ carcinoma, invasive carcinoma) using an ensemble of five pre-trained DCNN models: VGG16 (Simonyan and Zisserman 2014), VGG19 (Simonyan and Zisserman 2014), Xception (Chollet 2017), InceptionV3 (Szegedy et al. 2016), and InceptionResnetV2 (Szegedy et al. 2017). A bottleneck feature extraction approach was employed from the fine-tuned DCNN models as done in previous work (Kassani et al. 2019; Rakhlin et al. 2018). The extracted features were input into a Multi-layer MLP classifier, which were then integrated using a choquet fuzzy integral (Li et al. 2013). This integration involved unique fuzzy measures calculated through coalition game theory and information theory and was achieved through a novel fuzzy integral method that utilized shapley values to maintain monotonicity constraints. The ICIAR-2018 dataset underwent data pre-processing such as the stain normalization method introduced by Macenko et al. (2009). As a result, the proposed fuzzy ensemble method achieved accuracies of 96% and 95% for 2-class and 4-class categorizations, respectively.

In their study, Joseph et al. (2022) introduced a method that applied both handcrafted feature extraction and a DNN classifier for multi-class classification. Handcrafted techniques such as color histogram (Nahid et al. 2018), Hu moment (Hu 1962), and Haralick texture (Haralick et al. 1973) were employed to extract color, shape and texture features from BreakHis dataset images. These features were then fed into a DNN classifier with four dense layers and softmax activation for the purpose of classification. The results showed that the proposed method of handcrafted feature extraction combined with the DNN classifier achieved best accuracy scores of 97.89% for the 40× magnification.

The study conducted by Liu et al. (2022) investigated the classification of breast cancer images through a novel approach known as MSMV-PFENet (multi-scale, multi-view progressive feature encoding network). The methodology implicated the selection of image patches at various scales

based on the density of cell nuclei using the key-region-extraction module (KREM) (Liu et al. 2022), which were then input into the PFENet network for feature extraction. These extracted features were encoded into vectors. An essential step in the process involved the analysis of these vectors using a bidirectional long short-term memory (BiLSTM) network to obtain category scores. The feature discrimination network (FDNet) (Liu et al. 2022) examined the scores and determined the cancer type through the process of majority voting. The efficacy of the method was assessed using the ICIAR-2018 dataset, specifically for the four categories of normal, benign, in situ carcinoma, and invasive carcinoma. The achieved accuracy on patches was 93% and on full images reached 94.80%.

Ukwuoma et al. (2022) worked on multi-class classification and binary classification at different magnifications using DEEP_Patchi architecture (Fig. 5) based on DenseNet201 and VGG16 as ensemble learning, multiple self-attention heads and multi-linear perceptron. For more generalized features, the extraction was made with the ensemble learning. Afterwards, the identification and classification was realized after extracting the regions of interest (spatial information). This technique achieved a highest accuracy of 100% for 400× magnification on multi-class classification on the BreakHis dataset and 99.8% for binary classification, whereas on the the ICIAR-2018 dataset the accuracy was 99.9% (Table 4).

4.2 Approaches based on CNN models

4.2.1 Binary classification

In their study, Karthik et al. (2022) introduced an innovative ensemble-based deep learning strategy for the categorization of histopathological breast cancer images as benign or malignant. To achieve this, they developed two distinct customized CNN architectures: CSAResnet (a fusion of ResNet-101 with channel/spatial attention) and DAMCNN (a combination of DenseNet-201 and EfficientNet-B0). While CSAResnet focused on capturing features across both channels and spatial dimensions through the incorporation of attention modules, DAMCNN extracted multi-scale features in parallel pathways. The integration of a weighted ensemble technique was employed to improve overall performance by aggregating predictions from the two networks. Therefore, the proposed model achieved an accuracy of 99.55% on the BreakHis dataset.

A rank-based ensemble approach has been proposed by Majumdar et al. (2023) to perform binary classification on the BreakHis and ICIAR-2018 datasets. Using three CNN models, in particular GoogleNet (Szegedy et al. 2015), VGG16 (Simonyan and Zisserman 2014), and MobileNetV3_Small (Howard et al. 2019), binary classification was achieved on the BreakHis and ICIAR-2018 datasets. The utilization of the Gamma function in this study (Majumdar et al. 2023), along with a fuzzy rank-based ensemble, allowed for the combination of confidence scores from the three CNNs to make the final decision. This approach achieved an accuracy of 99.16%, 98.24%, 98.67%, and 96.16% for magnifications of 40×, 100×, 200×, and 400×, respectively on the BreakHis images. On the ICIAR-2018 dataset, the achieved accuracy was 96.95%.

Jakhar et al. (2024) created a self-based ensemble learning framework (SELF) wherein the top five classifiers, namely Extra tree (Abbas et al. 2021), Random Forest (Jackins et al. 2021), AdaBoost (Assegie et al. 2021), Gradient Boosting (Deif et al. 2021), and KNN9 (Khorshid and Abdulazeez 2021), were selected as base-learners from several others based to their accuracy measure after training. Finally, the logistic regression model was considered as the final estimator. For evaluation purposes, the BreakHis dataset (Spanhol et al. 2015) and the WDBC were utilized. The obtained results demonstrated an accuracy of 94.35% on the BreakHis dataset. In addition, an accuracy of approximately 99% was achieved on the WDBC.

Using the BreakHis dataset, two systems were assessed in the work of Taheri and Golrizkhatami (2023). The first system was based on magnification-specific binary classification (MSB) (Benhammou et al. 2020), employed a fine-tuned DenseNet201 CNN architecture. The second system consisted of magnification-independent binary classification (MIB) (Benhammou et al. 2020), utilized score-level (Taheri and Golrizkhatami 2023) fusion of four MSB subsystems, each equipped with a specific DenseNet201 model to deal with the diverse magnification factors. The evaluation metrics employed in this approach (Taheri and Golrizkhatami 2023) included image-level accuracy (ILA) (Benhammou et al. 2020) and patient-level accuracy (PLA) (Benhammou et al. 2020). The MSB classification achieved ILA results of 93.76% and PLA of 96.7% for 40× and 200× magnifications, respectively. For another, the MIB classification obtained ILA scores of 92.54% and 93.67% for 40× and 100× magnifications, respectively.

The objective of the study of Zhang et al. (2022) was to analyze the separation and reconstruction of the R, G, and B channels in the BreakHis dataset to obtain pure three-channel color images from all magnification images. To achieve this, the authors proposed the ColorDeep CNN model, which consists of a 10-layer architecture designed to extract reconstructed slice features within the nucleus region. This approach does not require nucleus segmentation and is suitable for low-complexity models. Following the feature extraction, a three-channel feature fusion classification was performed using a softmax classifier. The results highlighted the importance of the B

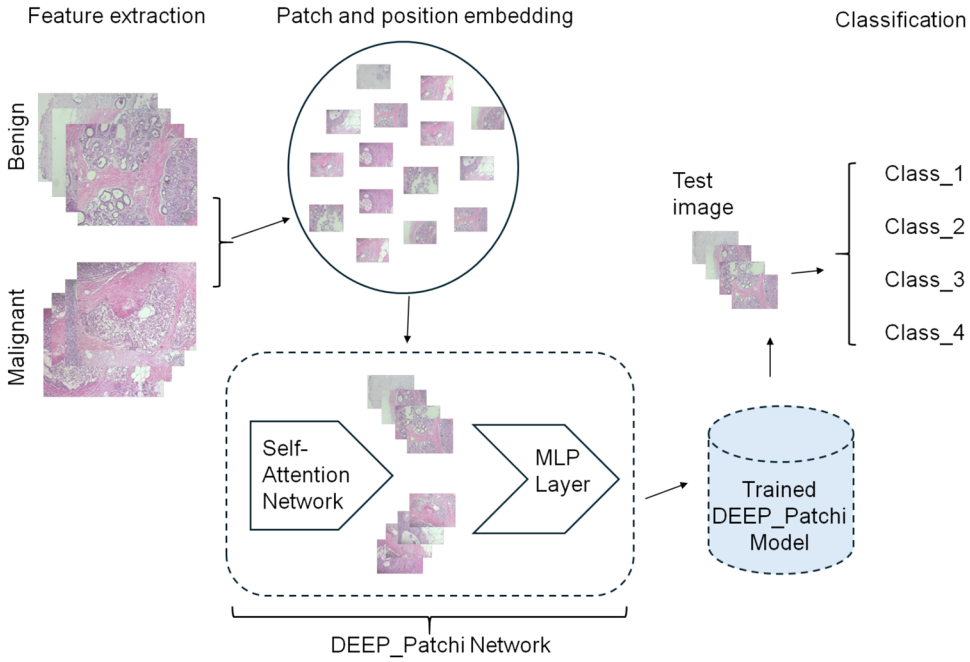

Fig. 5 Reproduction of the DEEP_Patchi proposed method in Ukwuoma et al. (2022)

channel which among the three channels, it had the most significant impact on the identification process across all four magnifications. The model 1, which utilized a fully connected layer, achieved an accuracy of 99.67%, while the model 2, employing global average pooling, achieved an accuracy of 96.89%. It is worth noting that the only difference between these two models lies in the structure of the classification layer.

Singh et al. (2023) studied the utilization of an QCNN (Zhu et al. 2018) to construct a residual network comprising four residual blocks, enabling the gradual construction of a deeper residual network (Trabelsi et al. 2017). This network consisted of multiple quaternion convolution layers. The breast histopathological images were transformed into the quaternion domains and to capture the relationship among the multiple channels of the input image, the authors employed the quaternion number systems, which provide powerful properties as generalizations of complex numbers. Hence, the images were converted into quaternion matrices to be served as inputs to the model. In this regard, the dataset used in this study (Singh et al. 2023) comprised images captured at the Hospital of the University of Pennsylvania and the Cancer Institute of New Jersey. The goal was to classify the images into two categories: IDC-negative and IDC-positive. The Deep Quaternion CNN yielded the following outcomes: an accuracy rate of 97.20%, precision of 98%, recall of 96%, specificity of 98%, F1-score of 99% and AUC of 97% (Table 3).

4.2.2 Multi-class classification

Nneji et al. (2023) introduced an approach called the Lightweight Separable Convolution Network (LWSC) to address the issue of low image quality in histopathological images. The LWSC model utilized a contrast enhancement that used contrast algorithm on small regions called tiles rather than entire image and edge detection techniques which applied gaussian smoothing for noise detection to extract visually informative trainable features from the images. It incorporated parallel stacks of separable convolution layers with varying filter sizes and it encompassed the process of factorizing convolution layers and employing bottleneck convolution layers to decrease the dimensionality of the model, therefore the model allowed for a reduction in the amount of trainable parameters as well as the computational cost. On top of that, this work (Nneji et al. 2023) served to classify both binary and multi-class categories on the BreakHis database with its different magnifying factors. The best results were obtained on the 40× magnifying factor with an accuracy of 93.12% on binary category and 97.23% for the multi-class category (Table 4).

5 Deep learning approaches for predicting gene expression based on WSIs and gene expression data

Recently, the deep learning have changed the way of the prediction of genetic expression in breast cancer, offering new perspectives. At the core of this innovation, lies the

Table 4 Summary for approaches used in multi-class classification of the breast cancer histopathological images: (Benign subtypes in BreakHis dataset: A, TA, PT, and F. Malignant subtypes in BreakHis dataset: DC, MC, LC, and PC)

References	Methodology	Datasets	Acc %	Prec %	Rec %	Spec %	F1 %	AUC %	MF	Classes
Nneji et al. (2023)	Parallel stacks of separable convolution layer + Contrast algorithm + Edge detection techniques	BreakHis	97.23	98.11	97.71	97.93	97.98	98.02	40×	A, F, TA, PT, DC, LC, MC, PC
Bagchi et al. (2022)	Patch feature extractors (VGG-19+ VGG-16+ Inception-ResNet v2) + Patch classification model (KNN+ SVM+ RF+ Adaboost+ XGBoost) + Image classification	ICIAR-2018	97.50	98.25	97.75	N/A	97.75	N/A	200×	Normal, Benign, In situ carcinoma, Invasive carcinoma
Shankar et al. (2022)	MixNet-based feature extractor (sequence of lightweight convolutional operators and depthwise separable convolutional layer) + Stacked Gated Recurrent Unit (SGRU) classification approach	BreakHis	98.61	92.80	91.48	N/A	92.10	N/A	40×+ 100×+ 200×+ 400×	A, F, TA, PT, DC, LC, MC, PC
Bhausaheb and Kashyap (2023)	Localization of cell nuclei with YOLO + Extraction of region with V-Net + Classification with Deer Canid optimization-based deep CNN	BreCaHAD BHI	93.05	94.42	93.54	N/A	92.98	N/A	40×	Normal, Benign, Malignant
Yu et al. (2023)	CDT based on SECS + coordinate attention mechanism	BreakHis	96.43	97.28	95.98	N/A	96.63	N/A	400×	A, F, TA, PT, DC, LC, MC, PC
Ahmed and Islam (2023)	Four-pathway CNN with EfficientNet-B0 as the backbone network	BreakHis	99.65	99.50	99.56	N/A	99.34	N/A	40×+ 100×+ 200×+ 400×	4-class separately for: -(A, F, TA, PT) -(DC, LC, MC, PC)
Bhowal et al. (2022)	Bottleneck feature extraction approach with DCNNs (VGG16, VGG19, Xception, InceptionV3, and InceptionResnetV2)+ MLP classifier with choquet fuzzy integral	ICIAR-2018	95.00	95.00	95.00	N/A	N/A	N/A	200×	Normal, Benign, In situ carcinoma, Invasive carcinoma
Joseph et al. (2022)	Handcrafted feature extraction + DNN classifier	BreakHis	97.89	N/A	N/A	N/A	N/A	N/A	40×	A, F, TA, PT, DC, LC, MC, PC
Liu et al. (2022)	MSMV-PFENet + KREM + BiLSTM + FDNet	ICIAR-2018	94.80	94.80	94.80	N/A	94.40	N/A	200×	Normal, Benign, In situ carcinoma, Invasive carcinoma

Table 4 (continued)

References	Methodology	Datasets	Acc %	Prec %	Rec %	Spec %	F1 %	AUC %	MF	Classes
Ukwuoma et al. (2022)	Ensemble learning (DenseNet201+VGG16) + Multiple self-attention heads+ Multi-linear perceptron	BreakHis	100	98.50	100	100	99.50	99.50	400×	A, F, TA, PT, DC, LC, MC, PC

development of algorithms capable of analysing histopathological images, including WSI, with genetic data, clinical data etc, to determine active genes linked to breast cancer and other tumors. Table 5 presents several methods used for this purpose, along with the various results.

5.1 Approaches based on CNN

One such noteworthy work is the ST-Net model, developed by He et al. (2020) using DenseNet-121, which identified over 100 genes, including 37 well-known cancer markers. The algorithm operated at a resolution of 100 µm and demonstrated the ability to generalize its predictions to several datasets without the need for re-training. The ST-Net model, underwent training using an ST data matched to WSIs, consisting of 30612 spots from 68 breast tissue sections, derived from 23 breast cancer patients and demonstrated a predictive capabilities for 250 genes. In the held-out patient, 5 genes were highly predicted wherein 4 genes were known as cancer biomarker, GNAS, ACTG1, FASN, DDX5, and XBP1 having respectively a median correlation across patient 0.34, 0.33, 0.31, 0.30, and 0.29. On external validation, ST-Net was able to predict 207 out of 234 genes in 10x Genomics Spatial Gene Expression dataset and 177 out of 249 genes from TCGA dataset with a positive correlation for both.

The hist2RNA model was trained on TCGA images with n = 335 when in the transcriptome-wide RNA sequencing data was obtained and have represented mRNA expression levels for 20438 genes in the study of (Mondol et al. 2023). It employed five pre-trained networks (EfficientNet, RegNet, DenseNet, Inception, ResNet) for extracting image features and then aggregated these features across patches to create a slide-level representation. Three convolutional blocks (C1, C2, C3) are utilized, each featuring a 1D convolutional layer with the ReLU activation function. C1 uses 256 kernels with a size of 5x5, while C2 and C3 use 512 kernels with a size of 1×1. The output layer comprised 138 neurons with a linear activation function, followed by a global average pooling layer. On the held-out test set the model using EfficientNet had 0.82 of correlation across patients and 0.29 across genes. In addition, when in using TMA cohort as an external dataset with n = 498, out of 32 genes were predicted, 17 of them had a strong relationship with PAM50 and among the top 20 genes, 13 were from the PAM50 group. Also, the model was able to predict the expression of common biomarkers such as ESR1, PGR, and ERBB2 for cases that are positive and negative of ER, PR and HER2, within both the TCGA and TMA datasets and similarly for the prediction of MKI67 gene and average Ki67% protein expression.

Qu et al. (2021) provided models that involved two steps. First, the feature extractor utilized the convolutional layers of ResNet-101 to derive feature vectors representing

the input tiles. Second, the MLP predictor, incorporating three fully connected layers and a self-attention layer. The study involved a WSIs dataset of n = 659 obtained from the Genomic Data Commons database. The models were trained for two distinct tasks. In the first task, the model was trained on the point mutations data and predicted 6 genes out of 18 genes, with notable emphasis on TP53, a gene frequently mutated in breast cancer and known for its prognostic significance, achieving an AUC of 0.729. In addition, the models predicted mutations in genes like RB1 (AUC 0.852), CDH1 (AUC 0.776), NF1 (AUC 0.768), and NOTCH2 (AUC 0.740). In the second task, the models were trained to predict CNAs in six genes. The classifier displayed an important predictive capability (AUC > 0.65) for the CNA data in breast cancer, encompassing 6 genes out of 35 genes: FGFR1, EIF4EBP1, KAT6A, HEY1, ZNF217, and RAB25, with corresponding AUC values of 0.794, 0.742, 0.732, 0.715, 0.693, and 0.686.

In their study, Rahaman et al. (2023) proposed BrST-Net framework. They trained 10 models on the STNet dataset without pre-trained weights: ResNet101, Inception-v3, EfficientNet-b0, EfficientNet-b1, EfficientNet-b2, EfficientNet-b3, EfficientNetb4, EfficientNet-b5, ViT-B-16, and ViT-B-32 with an without the introduced auxiliary network (AuxNet) into the framework. By integrating EfficientNet-b0 with AuxNet, 237 out of 250 genes were identified as having a positive correlation, including 24 genes with a PCC value above 0.50. The EfficientNet architecture, particularly the EfficientNet-b0 followed by EfficientNet-b4, was predominant. The genes B2M, ACTG1, ACTB, TMSB10, GNAS, PTMA, PTPRF, ERBB2, PRDX1, and TMSB4X emerged as the top predictions, with PCC values of 0.6325, 0.6233, 0.6204, 0.6197, 0.6139, 0.5902, 0.5926, 0.5897, 0.5702, and 0.5665, respectively.

Phan et al. (2021) proposed an approach employing deep learning models for the prediction of the PAM50 gene expression signature. This process involved the training of four pre-trained CNN models: VGG16, ResNet50, ResNet101, and Xception, utilizing a two-step transfer learning strategy. In the first step, these models underwent training using a dataset of pathological images from patients with recurrent breast cancer on the TCGA-BRCA dataset, with the primary objective of distinguishing between those at high and low risk of recurrence. In the second step, the models received further training to predict one of four PAM50 subtypes: basal, HER2-enriched, luminal A, or luminal B. The best average score of accuracy, precision, recall, and F1-score on patch-wise was 78% achieved by ResNet101. In addition, on the slide-wise, the ResNet101 achieved an accuracy of 89%, 93.75%, 96.90%, 85.70% on basal, HER2-enriched, luminal A, and luminal B subtypes respectively.

Liu et al. (2022) devised an other sophisticated deep learning framework to predict molecular subtypes, including Luminal A, Luminal B, HER2-enriched, and Basal-like. This framework was structured around a five-stage procedure and was formulated as a multi-instance learning challenge with the presence of noisy or weak labels at the patch level. To begin, they segmented WSIs into patches, a ResNet, trained using the co-teaching method, extracted features for each of these patches. In the following phase, they employed the Local Outlier Factor (LOF) to select distinctive patches based on their characteristics, wherein they fine-tuned the ResNet model in next stage, by incorporating both the entire slide and patch loss. Finally, they applied the optimized model to predict the ultimate molecular subtypes for individual patches as well as complete WSIs. their proposed framework was applied to the breast cancer H&E pathological image dataset from Xiangya Hospital (BCMT), achieving an accuracy of 64% and an F1 score of 0.685.

5.2 Approaches based on vision transformer

HisToGene was an attention-based model designed by Pang et al. (2021) to predict gene expression from WSIs using ST data. The architecture utilized a modified vision transformer, adept at handling image recognition tasks and modeling spatial dependencies. The model extracted patches from WSIs based on the spatial coordinates of ST spots, embedded these patches using learnable linear layers, and employed multi-head attention layers to generate latent embedding for gene expression prediction. The HER2+ dataset (Andersson et al. 2021) was used for analysis. The top four genes predicted by HisToGene are GNAS, MYL12B, FASN, and CLDN4, with mean R values of 0.32, 0.27, 0.27, and 0.26, respectively.

A deep learning model called Hist2ST was created by Zeng et al. (2022) using WSIs and ST data. The model integrated three primary modules: convmixer, transformer, and GNN. The convmixer module used convolution procedures to capture 2D vision features inside an image patch, allowing it to extract specific, localized features from WSIs. The transformer module employed a self-attention mechanism to capture global spatial dependencies between image patches, improving the understanding of spatial linkages by modeling global relationships. The GNN module utilized the GraphSAGE framework to learn local spatial dependencies from neighboring spots. The HER2+ (Andersson et al. 2021) dataset was used in this study. The performance of the Hist2ST model in predicting gene expression was evaluated using several metrics, including the average – log10 p-values, the average R, and the average p-values for each gene. The top five predicted genes based on these metrics were FN1, GNAS, SCD, MYL12B, FASN, and STMN1. FN1 had the highest average– log10 p-value of 16.68, an

Table 5 Summary of gene expression and cancer prognosis studies

References	Methodology	Dataset details	Results	Additional details
He et al. (2020)	ST-Net: DenseNet-121	WSIs and gene expression data, STNet dataset: 30612 spots from 68 sections and 23 patients, ST data	**Median correlation across patient:** (GNAS) = 0.34 (ACTG1) = 0.33 (FASN) = 0.31 (DDX5) = 0.30 (XBP1) = 0.29	External test: 207/234 genes in 10x Genomics Spatial Gene Expression dataset 177/249 on TCGA dataset
Qu et al. (2021)	ResNet-101 + MLP for predicting mutations and CNAs	WSIs and gene expression data, Genomic Data Commons: n = 659	**Mutation AUC:** (RB1) = 0.852 (CDH1) = 0.776 (NF1) = 0.768 (NOTCH2) = 0.740 (TP53) = 0.729 **CNA AUC:** (FGFR1) = 0.794 (EIF4EBP1) = 0.742 (KAT6A) = 0.732 (HEY1) = 0.715 (ZNF217) = 0.693 (RAB25) = 0.686	6/18 genes predicted in mutations 6/35 genes predicted in CNAs
Gao et al. (2023)	IGI-DL: CNN + GNN	WSIs and gene expression data, HER2+, ST data	**Mean correlation:** (ERBB2) = 0.374 (ACTG1) = 0.374 (ACTB) = 0.367 (GNAS) = 0.363 (TMSB10) = 0.355	External test: **Mean correlation:** (B2M) = 0.703 (FN1) = 0.633 (TMSB4X) = 0.633 (TPT1) = 0.625 (HLA-B) = 0.609
Rahaman et al. (2023)	BrST-Net: EfficientNet-b0 + AuxNet	WSIs and gene expression data, STNet dataset, ST data	**PCC:** (B2M) = 0.632 (ACTG1) = 0.623 (ACTB) = 0.620 (TMSB10) = 0.619 (GNAS) = 0.613 (PTMA) = 0.590 (PTPRF) = 0.592 (ERBB2) = 0.589 (PRDX1) = 0.570 (TMSB4X) = 0.566	237/250 genes predicted
Wang et al. (2021)	GPDBN: Inter-BFEM + Intra-BFEM + FC	WSIs and gene expression data, TCGA: 20436 gene expressions	**AUC =** 0.817 **C-index =** 0.723	N/A
Mondol et al. (2023)	Hist2RNA: Convolutional blocks + EfficientNet + ReLU + Global average pooling	WSIs and gene expression data, TCGA: n = 335 and mRNA expression analysis of 20438 genes	**Correlation:** Across patients = 0.82 Across genes = 0.29	External test: 17 genes from PAM50 on TMA dataset

Table 5 (continued)

References	Methodology	Dataset details	Results	Additional details
Mejia et al. (2023)	Based GNN model	WSIs and gene expression data, 10x Genomics dataset, STNet dataset	**MAE** = 0.630 **MSE** = 0.708 **PCC-Gene** = 0.383 **R2-Gene** = 0.106 **PCC-Patch** = 0.928 **R2-Patch** = 0.853	N/A
Wang et al. (2021)	GPDBN: Inter-BFEM + Intra-BFEM + FC	WSIs and gene expression data, TCGA: 20436 gene expressions	**AUC** = 0.817 **C-index** = 0.723	N/A
Liu et al. (2022)	ResNet: trained via Co-teaching	WSIs, BCMT dataset	**Accuracy** = 64% **F1-score** = 0.685	Prediction of 4 molecular subtypes: Luminal A, Luminal B, HER2-enriched, and Basal-like
Yang et al. (2022)	MCB model	WSIs and clinical information, 123 HER2+ from TCGA, 127 HER2+ from Chinese hospital	**AUC** = 0.76	External test: **AUC** = 0.72
Phan et al. (2021)	ResNet101: transfer learning	WSIs and gene expression data, TCGA-BRCA	**Accuracy:** (Patch-wise) = 78% (Slide-wise) = 96.90%	Prediction of 4 molecular subtypes: Luminal A, Luminal B, HER2-enriched, and Basal-like
Wang et al. (2023)	Deep-ITH: Elastic-net regularised Cox PH model	WSIs and gene expression data,TCGA-BC: n = 675, Clinseq-BC: n = 256	**HR** = 2.11 **CI** = 1.22-3.60 **p** = 0.007	External test: **HR** = 1.84 on SöS-BC-1 **HR** = 1.84 on SCAN-B-Lund
Pang et al. (2021)	HisToGene: vision transformer	WSIs and gene expression data, HER2+: 32 samples with 9612 spots and 785 genes, ST data	**Mean R:** (GNAS) = 0.32 (MYL12B) = 0.27 (FASN) = 0.27 (CLDN4) = 0.26	N/A
Zeng et al. (2022)	Hist2ST: convmixer, transformer, GNN	WSIs and gene expression data, HER2+: 32 samples of 7 patients with at least 180 spots/section, ST data	**Average R:** (FN1) = 0.39 (GNAS) = 0.37 (SCD) = 0.35 (MYL12B) = 0.33 (FASN) = 0.34 (STMN1) = 0.31	N/A
Yang et al. (2023)	EGN: a vision transformer backbone	WSIs and gene expression data, STNet dataset, 10xProteomic: 32032 slide windows and gene expressions pairs from 5 slides, ST data	**Results on STNet dataset:** $MSE_{\times 10^2}$= 4.10 $MAE_{\times 10^1}$= 1.61 $PCC@F_{\times 10^1}$= 1.51 $PCC@S_{\times 10^1}$= 2.25 $PCC@M_{\times 10^1}$= 2.02	Results on 10xProteomic dataset: $MSE_{\times 10^2}$= 5.49 $MAE_{\times 10^1}$= 1.55 $PCC@F_{\times 10^1}$= 6.78 $PCC@S_{\times 10^1}$= 7.21 $PCC@M_{\times 10^1}$= 7.07

Table 5 (continued)

References	Methodology	Dataset details	Results	Additional details
Dawood et al. (2021)	NSL	WSIs and gene expression data, HER2+: 378 spots per slide and 13620 total spots, ST data	**Median Pearson correlation** (GNAS) = 0.54 (FASN) = 0.52 (ACTG1) = 0.58 (ACTB) = 0.56 (ERBB2) = 0.54 (PFN1) = 0.50	215/250 genes Predicted

average R of 0.39194732, and an average p-value of 0.008. GNAS followed with an average– log10 p-value of 12.49, an average R of 0.37, and an average p-value of 0.0000031. SCD had an average– log10 p-value of 11.92, an average R of 0.35, and an average p-value of 0.017. MYL12B had an average– log10 p-value of 11.02, an average R of 0.33, and an average p-value of 0.002. FASN had an average– log10 p-value of 10.97, an average R of 0.34, and an average p-value of 0.004. Finally, STMN1 had an average– log10 p-value of 10.82, an average R of 0.34, and an average p-value of 0.001.

The Exemplar Guided Network (EGN) proposed by Yang et al. (2023) consisted of three main components: a representation extractor for unsupervised exemplar retrievals for the slide images, a vision transformer backbone for extracting input representations, and an exemplar bridging block to refine these representations using the nearest exemplars. The EGN model was trained and evaluated on two benchmark datasets, the STNet dataset and the 10xProteomic dataset that include WSIs and ST data. On the STNet dataset, the results were as follows: $MSE_{\times 10^2}$= 4.10, $MAE_{\times 10^1}$= 1.61, $PCC@F_{\times 10^1}$= 1.51, $PCC@S_{\times 10^1}$= 2.25, and $PCC@M_{\times 10^1}$= 2.02. On the 10xProteomic dataset, the results were: $MSE_{\times 10^2}$= 5.49, $MAE_{\times 10^1}$= 1.55, $PCC@F_{\times 10^1}$= 6.78, $PCC@S_{\times 10^1}$= 7.21, and $PCC@M_{\times 10^1}$= 7.07.

5.3 Approaches based on GNN

To maximize the benefits of both pixel intensity and structural characteristics, Gao et al. (2023) combined CNNs and GNNs to interpret TME heterogeneity using public dataset used in the study of(Andersson et al. 2021). The model called IGI-DL had two input streams and a fused output stream, which were designed to extract features from both the Nuclei-Graph and image patches with 100 µm, enabling it to predict spatial gene expression independently. In addition, it demonstrated good effectiveness in studying breast cancer and various other types of cancer. Across eight held-out patients, the top 5 genes predicted on leave-one-out validation set were ERBB2, ACTG1, ACTB, GNAS, and TMSB10 with a mean correlation of 0.374, 0.374, 0.367, 0.363, and 0.355 respectively. On the external test set, B2M, FN1, TMSB4X, TPT1, and HLA-B were predicted with a mean correlation of 0.703 0.633 0.633, 0.625, and 0.609 respectively. In addition, it provided novel risk scores derived from a super-graph, significantly improving the accuracy of cancer prognosis.

Mejia et al. (2023) introduced an approach that employed GNN, wherein each graph encapsulated information from multiple patches to represent spatial data. The model consisted of two distinct phases: local learning, which was the traditional method of fine-tuning an image encoder and spatial learning wherein the input patch along with its

neighboring patches are modeled as a graph. In this graph, the central node is the image for which gene expression prediction is targeted. This graph-based approach enabled the model to incorporate visual features from both the immediate location and its adjacent areas. In this study the 10x Genomics breast cancer and STNet dataset were used wherein the best results showed on Visium datasets with MAE of 0.630, MSE of 0.708, PCC-Gene of 0.383, R2-Gene of 0.106, PCC-Patch of 0.928, and R2-Patch of 0.853.

5.4 Other approaches

Yang et al. (2022) introduced a multi-modal compact bilinear model, MCB, for predicting the risk of recurrence and metastasis in HER2-positive breast cancer patients. It utilized both histopathology images and clinical data, such as age, tumor stage, and lymph node status. The model's objective was to classify patients into those at risk of recurrence and those not at risk. To develop the model, data of 127 HER2+ breast cancer patients from a Chinese hospital were used. The model achieved an AUC of 0.76 during 2-fold cross-validation. For independent validation, 123 HER2+ patients from TCGA were employed, the model achieved an AUC of 0.72 in predicting recurrence risk. Feature importance analysis was conducted using random forest on the clinical dataset. Then, logistic regression was used for classification. Histopathology images were processed by dividing them into patches and passing them through a ResNet50 network for feature extraction. These image features, along with the clinical data, were combined using the MCB fusion method, followed by batch normalization and a linear layer.

The authors (Wang et al. 2023) discussed a study on ITH in breast cancer and trained deep learning model (Deep-ITH) to predict spatial gene expression and quantify ITH in histopathology WSIs, focusing on the PAM50 gene panel. For each predicted transcript, 12 measures of heterogeneity were extracted. The prognostic model, based on ITH, was developed using an elastic-net regularised Cox proportional hazards model to divide patients into Deep-ITH low-risk and high-risk groups. The Deep-ITH model's development involved a training and an internal validation cohort comprising 931 breast cancer patients (TCGA-BC with N = 675 and Clinseq-BC with n = 256). The model was then evaluated in two separate and independent cohorts, one comprising 1358 patients called SöS-BC-1 and the other comprising 1262 patients called SCAN-B-Lund. The study found an increased risk of recurrence in the high-risk group, with a hazard ratio (HR) of 2.11 (95% CI 1.22–3.60; p = 0.007) in the cross validation, affirming the prognostic value of ITH features. In the SöS-BC-1 external test set, a significant difference in recurrence-free survival (RFS) rates between low and high Deep-ITH risk groups was observed (HR = 1.84; 95% CI 1.03–3.3; p = 0.04) and similar prognostic value was found in ER+ and HER2-negative subgroups. The SCAN-B-Lund cohort showed significant prognostic stratification in patients with grade 2 tumours and in patients with small tumours (less than 20 mm), although the overall adjusted HR in this cohort did not reach statistical significance (HR = 1.19; 95% CI 0.81–1.7; p = 0.375).

Wang et al. (2021) introduced the GPDBN framework (genomic and pathological deep bilinear network), to ameliorate the precision of predicting whether a breast cancer patient would have a poor or good prognosis by combining genomic data (gene expression profiles) and pathological images. The model had three essential components: An inter-modality bilinear feature encoding module (Inter-BFEM) designed to model complicated connections between genomic and image features, an intra-modality bilinear feature encoding modules (Intra-BFEMs) to capture relationships within each modality (genomic and images) and a multi-layer neural network that merged both inter and intra-modality features for the final prognosis prediction. WSIs and gene expression profiles, consisting of 20436 genes, of breast cancer patient samples were utilized from TCGA dataset. The framework achieved an AUC of 0.817 and a C-index of 0.723.

Dawood et al. (2021) created the Neural Stain Learning (NSL) model, which used related ST data to predict spatial gene expression profiles from WSIs. NSL learned a stain deconvolution matrix and used it to predict the relationship between stain absorption and patterns of gene expression. Using spatial coordinate data, image patches representing each ST spot were generated from each WSI for the gene expression prediction task. To learn the problem-specific stain matrix, the spot-level gene expression profile was utilized as a target label. Then, deconvolved pixel values from the image patches were combined to create a single patch-level feature for gene expression prediction. The model was tested on the HER2+ dataset. Results showed that NSL predicted the expression of 215 out of 250 genes and achieved a median Pearson correlation greater than 0.5 for 12, including GNAS (0.54), FASN (0.52), ACTG1 (0.58), ACTB (0.56), ERBB2 (0.54), and PFN1 (0.50).

6 Discussion

In this section we will discuss two parts, data collecting and pre-processing and model results, while focusing on classification and gene expression prediction.

6.1 Data collecting and pre-processing

In breast cancer research reviewed in this work, collecting and pre-processing datasets include handling

histopathological images, WSIs and gene expression sequences. This section summarizes the data collecting, pre-processing techniques used in the various studies. These techniques are designed to improve dataset quality, image resolution, and model training strength.

For classification purposes, several public datasets of histopathological images have been utilized in different studies. The BreakHis dataset (Spanhol et al. 2015), which is the most frequently used, includes 7909 images, featuring four magnification factors (40×, 100×, 200×, 400×) and eight subtypes (four benign and four malignant). The ICIAR 2018 dataset (Aresta et al. 2019) comprises 400 images with a magnification of 200× across four classes, while the BHI dataset (Cruz-Roa et al. 2014; Janowczyk and Madabhushi 2016) offers 162 images in two classes at a 40× magnification. In addition, the BreCaHAD dataset (Aksac et al. 2019) contains 23549 images in six classes, also at 40× magnification. The WDBC dataset provides 569 images in two classes and the private KMC dataset utilized in Rashmi et al. (2023) includes 1516 images with 400× magnification.

In gene expression from WSIs studies, different datasets have been employed for training, internal validation or external validation. The TCGA-BRCA dataset (Brigham & Women's Hospital & Harvard Medical School Chin Lynda et al. 2012, 2024) offers data for 1098 cases, including RNA sequences and clinical data. The CPTAC dataset (Rudnick et al. 2016) covers 1500 patients, while the TMA cohort (Mondol et al. 2023) includes data from 485 cases. Other datasets, such as the STNet dataset (He et al. 2020), encompasses 23 patients, BCMT dataset (Liu et al. 2022) contains 1254 patients and the ClinSeq-BC, SoS-BC-1, and SCAN-B datasets (Wang et al. 2022) include respectively 256, 1358 and 1262 cases.

Data pre-processing is useful for the consistence of deep learning models by introducing necessary variability and increasing the dataset size. Here's how these techniques are implemented across different references in this review: **For histopathological images analysis:**

1. Rotations: Images are rotated at random angles to simulate various viewing angles. This technique is used to add variability in image presentation and is referenced in several studies (Yu et al. 2023; Ahmed and Islam 2023; Ukwuoma et al. 2022; Hassan et al. 2023; Abbasniya et al. 2022; Zerouaoui and Idri 2022; Nakach et al. 2022; Taheri and Golrizkhatami 2023).
2. Flipping: Both horizontal and vertical flips are used to make sure that the original and rotated images represent different orientations, offering an important view of potential image perspectives. This method is documented in multiple research papers (Yu et al. 2023; Ahmed and Islam 2023; Joseph et al. 2022; Ukwuoma et al. 2022; Al-Jabbar et al. 2023; Hassan et al. 2023; Zerouaoui and Idri 2022; Nakach et al. 2022; Taheri and Golrizkhatami 2023).
3. Zooming: This involves using various factors and dynamic zoom ratios to adjust the scale of images, simulating different viewing distances. This technique is applied in several studies to improve the detail visibility (Ahmed and Islam 2023; Ukwuoma et al. 2022; Zerouaoui and Idri 2022; Nakach et al. 2022).
4. Contrast Limited Adaptive Histogram Equalization (CLAHE): Employed on a localized basis (tile-by-tile), this method enhances image contrast, which is crucial for identifying subtle differences in tissue samples (Zerouaoui and Idri 2022; Nakach et al. 2022).
5. Edge detection: techniques like Gaussian smoothing followed by hysteresis thresholding are utilized to reduce noise and clarify edge (Nneji et al. 2023).
6. Image resizing: standardizing image sizes to the dimensions required by various CNN models guarantees compatibility in model input, facilitating more consistent analysis.
7. Stain normalization: addressing one of the main challenges in histopathological image, stain normalization algorithms correct color inconsistencies that arise from variations in scanners and staining chemicals. A histogram specification-based algorithm is employed to normalize these variations (Taheri and Golrizkhatami 2023).

For gene expression analysis:

1. Gene selection: selecting genes related to cellular functions and breast cancer upgrade the relevance of the training data. Techniques including breast cancer-specific gene panels such as Oncotype DX, Mammaprint, Prosigna (PAM50), EndoPredict, Breast Cancer Index (BCI), and Mammostrat were applied (Phan et al. 2021; Mondol et al. 2023; Wang et al. 2023).
2. Standardization and normalization: RNA-seq protocols are standardized and transformations like log2(1+x) are applied to normalize gene expression levels. This provide data suitability for advanced analyses and consistency across studies (Gao et al. 2023; Mondol et al. 2023).

6.2 Models results discussion

In multi-class classification, the DEEP_Patchi model, combining DenseNet201 and VGG16, achieved an impressive accuracy of 100% on the BreakHis dataset at 40× magnification. This performance is attributed to the model utilizing multiple self-attention heads and a multi-linear perceptron,

Table 6 Summary of the top 3 on binary and multi-class classification accuracies in different studies

References	Model	Task	Dataset	MF	Accuracy
Ukwuoma et al. (2022)	DEEP_Patchi	Multi-class	BreakHis	40×	100%
Ahmed and Islam (2023)	Four-pathway CNN	Multi-class	BreakHis	Various	99.65%
Shankar et al. (2022)	CSSADTL-BCC	Multi-class	BreakHis	Various	98.61%
Al-Jabbar et al. (2023)	Hybrid CNN+ANN	Binary	BreakHis	40×	100%
Yu et al. (2023)	CA-BreastNet	Binary	BreakHis	40×, 400×	100%
Clement et al. (2022)	BoDMCF+ SVM	Binary	BreakHis	Various	99.92%
Ukwuoma et al. (2022)	DEEP_Patchi	Binary	BreakHis	Various	99.80%

demonstrating the efficacy of integrating deep learning approaches for detailed classification of histopathological images (Ukwuoma et al. 2022). Following closely, the four-pathway CNN model using EfficientNet-B0 as the backbone network also showed excellent results with an accuracy of 99.65% on the BreakHis dataset for four-class classification. This approach highlighted the use of multiple instance learning to handle various image magnifications, achieving high performance by Ahmed and Islam (2023). The CSSADTL-BCC model employed a combination of feature extraction by MixNet and classification through stacked gated recurrent units (SGRU) to achieve an accuracy of 98.61% on the BreakHis dataset for eight-class classification (Shankar et al. 2022), showing the impact of combining lightweight convolutional operators with deep recurrent structures for complicated classification tasks.

In binary classification, Al-Jabbar et al. (2023) developed a binary classification method using a combination of AlexNet and handcrafted features with an ANN algorithm, achieving perfect scores of 100% in accuracy at 40× magnification on the BreakHis dataset. Also Yu et al. (2023) achieved an accuracy of 100% with the CA-BreastNet model on the BreakHis dataset at both 40× and 40× magnifications. Clement et al. (2022) achieved an average accuracy of 99.92% using a machine learning approach with pre-trained CNN models including ResNet-50, EfficientNetb0, and Inception-v3 on the BreakHis dataset across magnifications of 40×, 40×, 40×, and 40×. Ukwuoma et al. (2022) reached an accuracy of 99.80% for binary classification using the DEEP Patchi architecture. The Table 6 shows all details about the top 3 accuracies of each task.

In gene expression prediction, the ST-Net model He et al. (2020) using the DenseNet-121, identified over 237 genes on the STNet dataset. The correlation of the top three genes were 0.34 for GNAS, 0.33 for ACTG1, and 0.31 for FASN. The hist2RNA model (Mondol et al. 2023) using convolutional blocks, EfficientNet, RELU, and global average pooling, achieved correlations as high as 0.82 across patients on the TCGA dataset. Similarly, the model of Qu et al. (2021) integrated ResNet-101 and MLP, predicted mutations, and CNAs with notable accuracy on the genomic data commons dataset, highlighted by mutation AUC reaching up to 0.852 on RB1 and CNA AUC reaching up to 0.794 on FGFR1. Gao et al. (2023) integrated CNNs and GNNs to analyze tumor microenvironment heterogeneity, showing correlations of 0.374, 0.374, 0.367 on ERBB2, ACTG1, and ACTB respectively on the HER2+ dataset and showed over 0.70 in external tests. The graph-based and network-integrated approaches of Mejia et al. (2023) and (Rahaman et al. 2023) emphasize the strength of combining spatial and histological data for more precise gene prediction. The SEPAL model (Mejia et al. 2023) achieved PCC-Gene of 0.383 and PCC-Patch of 0.928 on the Visium dataset. The top three gene predicted with the BrST-Net model (Rahaman et al. 2023) were B2M, ACTG1, and ACTB with PCC of 0.632, 0.623, and 0.620 respectively on ST dataset. On the prognosis task, the frameworks of Wang et al. (2021) achieved AUC of 0.817 on the TCGA dataset and Liu et al. (2022) achieved an accuracy of 64% on the BCMT dataset. The two models have effectively combined genomic and pathological data to predict cancer outcomes with high prognostic accuracy. Using vision transformer based models, HisToGene (Pang

Table 7 Summary of the top 3 predicted genes in different studies

References	Model	Dataset	Top1	Top2	Top3
He et al. (2020)	ST-Net	STNet	GNAS	ACTG1	FASN
Gao et al. (2023)	IGI-DL	HER2+	ERBB2	ACTG1	ACTB
Rahaman et al. (2023)	BrST-Net	STNet	B2M	ACTG1	ACTB
Qu et al. (2021)	ResNet-101+ MLP	GDC	RB1	FGFR1	CDH1
Pang et al. (2021)	HisToGene	HER2+	GNAS	MYL12B	FASN
Zeng et al. (2022)	Hist2ST	HER2+	FN1	GNAS	SCD
Dawood et al. (2021)	NSL	HER2+	GNAS	FASN	ACTG1

et al. 2021), was capable to predict top four genes, GNAS, MYL12B, FASN, and CLDN4, with mean R of 0.32, 0.27, 0.27, and 0.26 respectively on the HER2+ dataset, with the same dataset, the top six genes predicted by Hist2ST (Zeng et al. 2022) were FN1, GNAS, SCD, MYL12B, FASN, and STMN1 with average R of 0.39, 0.37, 0.35, 0.33, 0.34, and 0.31.

These studies collectively highlight the transformative impact of deep learning in treating breast cancer through refined gene expression across varied datasets. The Table 7 shows the top 3 predicted gene in different studies.

Overall, we can see that the best-performing models used architectures based on CNN for both binary and multi-class classification tasks on histopathological images especially for highest magnification factors. As these images have high resolution, the models are able to extract most information during training. For another, in the area of gene prediction from WSIs and gene expression data based on the provided studies, the best performance can be evaluated based on various metrics and goals of each study, rather than a direct comparison, as each study addresses different aspects of gene prediction with varying methodologies and datasets. As research progressed, the limitations of CNNs in capturing inter-relations or dependencies between different regions of an image or between different data types became apparent. GNNs were introduced to address these limitations, especially useful in medical research for modeling complicated relationships and interactions within data. GNNs can interpret not only the features within individual cells or tissue structures but also the interactions between them, which are essential in studying cancer progression and gene expression dynamics. In addition, the use of the vision transformer was shown in some studies, demonstrating its capabilities in this task.

7 Conclusion

This paper has highlighted the important deep learning techniques used to improve breast cancer diagnostics. From enhancing the accuracy of histopathological image classification to predicting gene expression patterns, to support breast cancer early detection. We also talked about the use of hybrid models, ensemble learning, CNN, vision transformer, and GNN proposed in several works. These models have been capable to differentiate between benign and malignant tumors with remarkable results. Similarly, deep learning's capacity to extract gene expression patterns from histopathological images has emerged as a significant insights.

Despite these developments, there are still issues, namely with regard to heterogeneity of data, the requirement for substantial annotated datasets and the incorporation of multi-modal data sources. Continuous improvements in model architecture, training approaches, and data handling strategies are needed to meet these challenges. Future research may look also into more sophisticated transformer-based models, which are starting to make significant progress in image bioinformatics and have demonstrated amazing results in natural language processing. Transformers may be able to handle the sequential nature of gene expression data and simulate difficult connections in images data even better.

These advances can improve breast cancer diagnosis and therapy, leading to more appropriate medical interventions.

Funding This research was enabled in part by support provided by the Natural Sciences and Engineering Research Council of Canada (NSERC), funding reference number RGPIN-2024-05287.

References

Abbas S, Jalil Z, Javed AR, Batool I, Khan MZ, Noorwali A, Gadekallu TR, Akbar A (2021) Bcd-wert: a novel approach for breast cancer detection using whale optimization based efficient features and extremely randomized tree algorithm. PeerJ Comput Sci 7:390. https://doi.org/10.7717/peerj-cs.390

Abbasniya MR, Sheikholeslamzadeh SA, Nasiri H, Emami S (2022) Classification of breast tumors based on histopathology images using deep features and ensemble of gradient boosting methods. Comput Electr Eng 103:108382. https://doi.org/10.1016/j.compeleceng.2022.108382

Abo-El-Rejal A, Ayman S, Aymen F (2024) Advances in breast cancer segmentation: a comprehensive review. Acadlore Trans AI Mach Learn 3(2):70–83. https://doi.org/10.56578/ataiml030201

Ahmed M, Islam MR (2023) A combined feature-vector based multiple instance learning convolutional neural network in breast cancer classification from histopathological images. Biomed Signal Process Control 84:104775. https://doi.org/10.1016/j.bspc.2023.104775

Aksac A, Demetrick DJ, Ozyer T, Alhajj R (2019) Brecahad: a dataset for breast cancer histopathological annotation and diagnosis. BMC Res Notes 12(82):1–3. https://doi.org/10.1186/s13104-019-4121-7

Al-Jabbar M, Alshahrani M, Senan EM, Ahmed IA (2023) Multi-method diagnosis of histopathological images for early detection of breast cancer based on hybrid and deep learning. Mathematics 11(6):1429. https://doi.org/10.3390/math11061429

Andersson A, Larsson L, Stenbeck L, Salmén F, Ehinger A, Wu SZ, Al-Eryani G, Roden D, Swarbrick A, Borg Å et al (2021) Spatial deconvolution of her2-positive breast cancer delineates tumor-associated cell type interactions. Nat Commun 12(1):6012. https://doi.org/10.1038/s41467-021-26271-2

Aresta G, Araújo T, Kwok S, Chennamsetty SS, Safwan M, Alex V, Marami B, Prastawa M, Chan M, Donovan M et al (2019) Bach: grand challenge on breast cancer histology images. Med Image Anal 56:122–139. https://doi.org/10.1016/j.media.2019.05.010

Assegie TA, Tulasi RL, Kumar NK (2021) Breast cancer prediction model with decision tree and adaptive boosting. IAES Int J Artif Intell 10(1):184. https://doi.org/10.11591/ijai.v10.i1.pp184-190

Atban F, Ekinci E, Garip Z (2023) Traditional machine learning algorithms for breast cancer image classification with optimized deep features. Biomed Signal Process Control 81:104534. https://doi.org/10.1016/j.bspc.2022.104534

Bagchi A, Pramanik P, Sarkar R (2022) A multi-stage approach to breast cancer classification using histopathology images. Diagnostics 13(1):126. https://doi.org/10.3390/diagnostics13010126

Benhammou Y, Achchab B, Herrera F, Tabik S (2020) Breakhis based breast cancer automatic diagnosis using deep learning: Taxonomy, survey and insights. Neurocomputing 375:9–24. https://doi.org/10.1016/j.neucom.2019.09.044

Bhausaheb DP, Kashyap KL (2023) Detection and classification of breast cancer availing deep canid optimization based deep cnn. Multimed Tools Appl 82(12):18019–18037. https://doi.org/10.1007/s11042-022-14268-y

Bhowal P, Sen S, Velasquez JD, Sarkar R (2022) Fuzzy ensemble of deep learning models using choquet fuzzy integral, coalition game and information theory for breast cancer histology classification. Expert Syst Appl 190:116167. https://doi.org/10.1016/j.eswa.2021.116167

Brigham & Women's Hospital & Harvard Medical School Chin Lynda, PPJKR, Medicine Creighton Chad J, Donehower Lawrence AG, Systems Biology Reynolds Sheila 31 Kreisberg Richard B, Bernard Brady Bressler Ryan Erkkila Timo Lin Jake Thorsson Vesteinn Zhang Wei Shmulevich Ilya, I, et al (2012) Comprehensive molecular portraits of human breast tumours. Nature 490(7418): 61–70. https://doi.org/10.1038/nature11412. Accessed 2024.

Brigham & Women's Hospital & Harvard Medical School Chin Lynda, PPJKR, Medicine Creighton Chad J, Donehower Lawrence AG, Systems Biology Reynolds Sheila 31 Kreisberg Richard B, Bernard Brady Bressler Ryan Erkkila Timo Lin Jake Thorsson Vesteinn Zhang Wei Shmulevich Ilya, I, et al (2024) Comprehensive molecular portraits of human breast tumours. https://portal.gdc.cancer.gov/projects/TCGA-BRCA

Chaudhari S, Mithal V, Polatkan G, Ramanath R (2021) An attentive survey of attention models. ACM Trans Intell Syst Technol (TIST) 12(5):1–32. https://doi.org/10.1145/3465055

Chelgani SC, Nasiri H, Alidokht M (2021) Interpretable modeling of metallurgical responses for an industrial coal column flotation circuit by xgboost and shap-a conscious-lab development. Int J Min Sci Technol 31(6):1135–1144. https://doi.org/10.1016/j.ijmst.2021.10.006

Chollet F (2017) Xception: Deep learning with depthwise separable convolutions. In: Proceedings of the IEEE Conference on computer vision and pattern recognition, pp 1251–1258

Clement D, Agu E, Obayemi J, Adeshina S, Soboyejo W (2022) Breast cancer tumor classification using a bag of deep multi-resolution convolutional features. Informatics. https://doi.org/10.3390/informatics9040091

Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20:273–297. https://doi.org/10.1007/BF00994018

Cruz-Roa A, Basavanhally A, González F, Gilmore H, Feldman M, Ganesan S, Shih N, Tomaszewski J, Madabhushi A (2014) Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In: Medical Imaging 2014: Digital Pathology, p. 904103. https://doi.org/10.1117/12.2043872

Dawood M, Branson K, Rajpoot NM, Minhas FuAA (2021) All you need is color: image based spatial gene expression prediction using neural stain learning. In: Joint European Conference on machine learning and knowledge discovery in databases, pp. 437–450. https://doi.org/10.1007/978-3-030-93733-1_32

Deif M, Hammam R, Solyman A (2021) Gradient boosting machine based on pso for prediction of leukemia after a breast cancer diagnosis. Int J Adv Sci Eng Inf Technol 11(2):508–515. https://doi.org/10.18517/ijaseit.11.2.12955

Emerson P (2013) The original Borda count and partial voting. Soc Choice Welf 40(2):353–358. https://doi.org/10.1007/s00355-011-0603-9

Ezzoddin M, Nasiri H, Dorrigiv M (2022) Diagnosis of covid-19 cases from chest x-ray images using deep neural network and lightgbm. In: 2022 International Conference on machine vision and image processing (MVIP), pp. 1–7. https://doi.org/10.1109/MVIP53647.2022.9738760

Fu J, Liu J, Tian H, Li Y, Bao Y, Fang Z, Lu H (2019) Dual attention network for scene segmentation. In: 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3141–3149. https://doi.org/10.1109/CVPR.2019.00326

Gao Y, Zhou Y, Luo Q (2020) An efficient binary equilibrium optimizer algorithm for feature selection. IEEE Access 8:140936–140963. https://doi.org/10.1109/ACCESS.2020.3013617

Gao R, Yuan X, Ma Y, Wei T, Johnston L, Shao Y, Lv W, Zhu T, Zhang Y, Zheng J, et al (2023) Predicting gene spatial expression and cancer prognosis: An integrated graph and image deep learning approach based on he slides. bioRxiv, 2023–07, https://doi.org/10.1101/2023.07.20.549824

Goldhirsch A, Winer EP, Coates A, Gelber R, Piccart-Gebhart M, Thürlimann B, Senn H-J, Albain KS, André F, Bergh J et al (2013) Personalizing the treatment of women with early breast cancer: highlights of the St Gallen international expert consensus on the primary therapy of early breast cancer 2013. Ann Oncol 24(9):2206–2223. https://doi.org/10.1093/annonc/mdt303

Gupta D, Rani R (2020) Improving malware detection using big data and ensemble learning. Comput Electr Eng 86:106729. https://doi.org/10.1016/j.compeleceng.2020.106729

Haralick RM, Shanmugam K, Dinstein IH (1973) Textural features for image classification. IEEE Trans Syst Man Cybern 6:610–621. https://doi.org/10.1109/TSMC.1973.4309314

Hassan AM, Yahya A, Aboshosha A (2023) A framework for classifying breast cancer based on deep features integration and selection. Neural Comput Appl 35(16):12089–12097. https://doi.org/10.1007/s00521-023-08341-2

He B, Bergenstråhle L, Stenbeck L, Abid A, Andersson A, Borg Å, Maaskola J, Lundeberg J, Zou J (2020) Integrating spatial gene expression and breast tumour morphology via deep learning. Nat Biomed Eng 4(8):827–834. https://doi.org/10.1038/s41551-020-0578-x

Hou Q, Zhou D, Feng J (2021) Coordinate attention for efficient mobile network design. In: Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition, pp. 13713–13722

Howard A, Sandler M, Chu G, Chen L-C, Chen B, Tan M, Wang W, Zhu Y, Pang R, Vasudevan V, et al (2019) Searching for mobilenetv3. In: Proceedings of the IEEE/CVF International Conference on computer vision, pp 1314–1324

Hu M-K (1962) Visual pattern recognition by moment invariants. IRE Trans Inf Theory 8(2):179–187. https://doi.org/10.1109/TIT.1962.1057692

Hu P, Pan J-S, Chu S-C, Sun C (2022) Multi-surrogate assisted binary particle swarm optimization algorithm and its application for feature selection. Appl Soft Comput 121:108736. https://doi.org/10.1016/j.asoc.2022.108736

Jackins V, Vimal S, Kaliappan M, Lee MY (2021) Ai-based smart prediction of clinical disease using random forest classifier and naive bayes. J Supercomput 77(5):5198–5219. https://doi.org/10.1007/s11227-020-03481-x

Jakhar AK, Gupta A, Singh M (2024) Self: a stacked-based ensemble learning framework for breast cancer classification. Evol Intel 17(3):1341–1356. https://doi.org/10.1007/s12065-023-00824-4

Janowczyk A, Madabhushi A (2016) Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. J Pathol Inf 7(1):29. https://doi.org/10.4103/2153-3539.186902

Joseph AA, Abdullahi M, Junaidu SB, Ibrahim HH, Chiroma H (2022) Improved multi-classification of breast cancer histopathological images using handcrafted features and deep neural network (dense layer). Intell Syst Appl 14:200066. https://doi.org/10.1016/j.iswa.2022.200066

Karthik R, Menaka R, Siddharth M (2022) Classification of breast cancer from histopathology images using an ensemble of deep multiscale networks. Biocybern Biomed Eng 42(3):963–976. https://doi.org/10.1016/j.bbe.2022.07.006

Kassani SH, Kassani PH, Wesolowski MJ, Schneider KA, Deters R (2019) Breast cancer diagnosis with transfer learning and global pooling. In: 2019 International Conference on information and communication technology convergence (ICTC), pp 519–524. https://doi.org/10.1109/ICTC46691.2019.8939878

Khorshid SF, Abdulazeez AM (2021) Breast cancer diagnosis based on k-nearest neighbors: a review. PalArch's J Archaeol Egypt/Egyptol 18(4):1927–1951

Kim A, Kim H (2022) A new classification tree method with interaction detection capability. Comput Stat Data Anal 165:107324. https://doi.org/10.1016/j.csda.2021.107324

Li C, Zeng-tai G, Gang D (2013) Genetic algorithm optimization for determining fuzzy measures from fuzzy data. J Appl Math 2013(1):542153. https://doi.org/10.1155/2013/542153

Lin T-Y, Goyal P, Girshick R, He K, Dollár P (2020) Focal loss for dense object detection. IEEE Trans Pattern Anal Mach Intell 42(2):318–327. https://doi.org/10.1109/TPAMI.2018.2858826

Liu L, Feng W, Chen C, Liu M, Qu Y, Yang J (2022) Classification of breast cancer histology images using msmv-pfenet. Sci Rep 12(1):17447. https://doi.org/10.1038/s41598-022-22358-y

Liu H, Xu W-D, Shang Z-H, Wang X-D, Zhou H-Y, Ma K-W, Zhou H, Qi J-L, Jiang J-R, Tan L-L et al (2022) Breast cancer molecular subtype prediction on pathological images with discriminative patch selection and multi-instance learning. Front Oncol 12:858453. https://doi.org/10.3389/fonc.2022.858453

Li P, Xie J, Wang Q, Zuo, W (2017) Is second-order information helpful for large-scale visual recognition? In: 2017 IEEE International Conference on Computer Vision (ICCV), pp 2089–2097. https://doi.org/10.1109/ICCV.2017.228

Macenko M, Niethammer M, Marron JS, Borland D, Woosley JT, Guan X, Schmitt C, Thomas NE (2009) A method for normalizing histology slides for quantitative analysis. In: 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp 1107–1110. https://doi.org/10.1109/ISBI.2009.5193250

Majumdar S, Pramanik P, Sarkar R (2023) Gamma function based ensemble of cnn models for breast cancer detection in histopathology images. Expert Syst Appl 213:119022. https://doi.org/10.1016/j.eswa.2022.119022

Mejia G, Cárdenas P, Ruiz D, Castillo A, Arbeláez P (2023) Sepal: spatial gene expression prediction from local graphs. In: Proceedings of the IEEE/CVF International Conference on computer vision, pp. 2294–2303

Millar EK, Browne LH, Beretov J, Lee K, Lynch J, Swarbrick A, Graham PH (2020) Tumour stroma ratio assessment using digital image analysis predicts survival in triple negative and luminal breast cancer. Cancers 12(12):3749. https://doi.org/10.3390/cancers12123749

Mondol RK, Millar EK, Graham PH, Browne L, Sowmya A, Meijering E (2023) hist2rna: an efficient deep learning architecture to predict gene expression from breast cancer histopathology images. Cancers 15(9):2569. https://doi.org/10.3390/cancers15092569

Nadimi-Shahraki MH, Taghian S, Mirjalili S (2021) An improved grey wolf optimizer for solving engineering problems. Expert Syst Appl 166:113917. https://doi.org/10.1016/j.eswa.2020.113917

Nahid A-A, Mehrabi MA, Kong Y (2018) Histopathological breast cancer image classification by deep neural network techniques guided by local clustering. Biomed Res Int 2018(1):2362108. https://doi.org/10.1155/2018/2362108

Nakach F-Z, Zerouaoui H, Idri A (2022) Hybrid deep boosting ensembles for histopathological breast cancer classification. Heal Technol 12(6):1043–1060. https://doi.org/10.1007/s12553-022-00709-z

Nasiri H, Kheyroddin G, Dorrigiv M, Esmaeili M, Nafchi AR, Ghorbani MH, Zarkesh-Ha P (2022) Classification of covid-19 in chest x-ray images using fusion of deep features and lightgbm. In: 2022 IEEE World AI IoT Congress (AIIoT), pp. 201–206. https://doi.org/10.1109/AIIoT54504.2022.9817375

Nayak S, Bhat M, Reddy NS, Rao BA (2022) Study of distance metrics on k-nearest neighbor algorithm for star categorization. J Phys Conf Ser. https://doi.org/10.1088/1742-6596/2161/1/012004

Niu Z, Zhong G, Yu H (2021) A review on the attention mechanism of deep learning. Neurocomputing 452:48–62. https://doi.org/10.1016/j.neucom.2021.03.091

Nneji GU, Monday HN, Mgbejime GT, Pathapati VSR, Nahar S, Ukwuoma CC (2023) Lightweight separable convolution network for breast cancer histopathological identification. Diagnostics 13(2):299. https://doi.org/10.3390/diagnostics13020299

Obayya M, Maashi MS, Nemri N, Mohsen H, Motwakel A, Osman AE, Alneil AA, Alsaid MI (2023) Hyperparameter optimizer with deep learning-based decision-support systems for histopathological breast cancer diagnosis. Cancers 15(3):885. https://doi.org/10.3390/cancers15030885

Pang M, Su K, Li M (2021) Leveraging information in spatial transcriptomics to predict super-resolution gene expression from histology images in tumors. BioRxiv, 2021–11, https://doi.org/10.1101/2021.11.28.470212

Phan NN, Huang C-C, Tseng L-M, Chuang EY (2021) Predicting breast cancer gene expression signature by applying deep convolutional neural networks from unannotated pathological images. Front Oncol 11:769447. https://doi.org/10.3389/fonc.2021.769447

Qu H, Zhou M, Yan Z, Wang H, Rustgi VK, Zhang S, Gevaert O, Metaxas DN (2021) Genetic mutation and biological pathway prediction based on whole slide images in breast carcinoma using deep learning. NPJ Precis Oncol 5(1):87. https://doi.org/10.1038/s41698-021-00225-9

Quellec G, Cazuguel G, Cochener B, Lamard M (2017) Multiple-instance learning for medical image and video analysis. IEEE Rev Biomed Eng 10:213–234. https://doi.org/10.1109/RBME.2017.2651164

Rahaman MM, Millar EK, Meijering E (2023) Breast cancer histopathology image-based gene expression prediction using spatial transcriptomics data and deep learning. Sci Rep 13(1):13604. https://doi.org/10.1038/s41598-023-40219-0

Rakhlin A, Shvets A, Iglovikov V, Kalinin AA (2018) Deep convolutional neural networks for breast cancer histology image analysis. In: Image Analysis and Recognition: 15th International Conference, ICIAR 2018, Póvoa de Varzim, Portugal, June 27–29, 2018, Proceedings 15, pp. 737–744. https://doi.org/10.1007/978-3-319-93000-8_83

Rashmi R, Prasad K, Udupa CBK (2022) Breast histopathological image analysis using image processing techniques for diagnostic purposes: a methodological review. J Med Syst 46(1):7. https://doi.org/10.1007/s10916-021-01786-9

Rashmi R, Prasad K, Udupa CBK (2023) Region-based feature enhancement using channel-wise attention for classification of breast histopathological images. Neural Comput Appl 35(8):5839–5854. https://doi.org/10.1007/s00521-022-07966-z

Reinhard E, Adhikhmin M, Gooch B, Shirley P (2001) Color transfer between images. IEEE Comput Graphics Appl 21(5):34–41. https://doi.org/10.1109/38.946629

Rudnick PA, Markey SP, Roth J, Mirokhin Y, Yan X, Tchekhovskoi DV, Edwards NJ, Thangudu RR, Ketchum KA, Kinsinger CR et al (2016) A description of the clinical proteomic tumor analysis consortium (cptac) common data analysis pipeline. J Proteome Res 15(3):1023–1032. https://doi.org/10.1021/acs.jproteome.5b01091

Shankar K, Dutta AK, Kumar S, Joshi GP, Doo IC (2022) Chaotic sparrow search algorithm with deep transfer learning enabled breast cancer classification on histopathological images. Cancers 14(11):2770. https://doi.org/10.3390/cancers14112770

Sharma S, Kumar S (2022) The xception model: a potential feature extractor in breast cancer histology images classification. ICT Express 8(1):101–108. https://doi.org/10.1016/j.icte.2021.11.010

Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, https://doi.org/10.48550/arXiv.1409.1556

Singh S, Tripathi B, Rawat SS (2023) Deep quaternion convolutional neural networks for breast cancer classification. Multimed Tools Appl 82(20):31285–31308. https://doi.org/10.1007/s11042-023-14688-4

Soliman A, Li Z, Parwani AV (2024) Artificial intelligence's impact on breast cancer pathology: a literature review. Diagn Pathol 19(1):1–18. https://doi.org/10.1186/s13000-024-01453-w

Spanhol FA, Oliveira LS, Petitjean C, Heutte L (2015) A dataset for breast cancer histopathological image classification. IEEE Trans Biomed Eng 63(7):1455–1462. https://doi.org/10.1109/TBME.2015.2496264

Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: Proceedings of the IEEE Conference on computer vision and pattern recognition, pp 1–9

Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on computer vision and pattern recognition, pp 2818–2826

Szegedy C, Ioffe S, Vanhoucke V, Alemi A (2017) Inception-v4, inception-resnet and the impact of residual connections on learning. In: Proceedings of the AAAI Conference on artificial intelligence. https://doi.org/10.1609/aaai.v31i1.11231

Tafavvoghi M, Bongo LA, Shvetsov N, Busund L-TR, Møllersen K (2024) Publicly available datasets of breast histopathology h &e whole-slide images: a scoping review. J Pathol Inform 15:100363. https://doi.org/10.1016/j.jpi.2024.100363

Taheri S, Golrizkhatami Z (2023) Magnification-specific and magnification-independent classification of breast cancer histopathological image using deep learning approaches. SIViP 17(2):583–591. https://doi.org/10.1007/s11760-022-02263-7

Thakur N, Kumar P, Kumar A (2024) A systematic review of machine and deep learning techniques for the identification and classification of breast cancer through medical image modalities. Multimed Tools Appl 83(12):35849–35942. https://doi.org/10.1007/s11042-023-16634-w

Tian M-W, Yan S-R, Han S-Z, Nojavan S, Jermsittiparsert K, Razmjooy N (2020) New optimal design for a hybrid solar chimney, solid oxide electrolysis and fuel cell based on improved deer hunting optimization algorithm. J Clean Prod 249:119414. https://doi.org/10.1016/j.jclepro.2019.119414

Trabelsi C, Bilaniuk O, Zhang Y, Serdyuk D, Subramanian S, Santos JF, Mehri S, Rostamzadeh N, Bengio Y, Pal CJ (2017) Deep complex networks. arXiv preprint arXiv:1705.09792, https://doi.org/10.48550/arXiv.1705.09792

Ukwuoma CC, Hossain MA, Jackson JK, Nneji GU, Monday HN, Qin Z (2022) Multi-classification of breast cancer lesions in histopathological images using deep_pachi: Multiple self-attention head. Diagnostics 12(5):1152. https://doi.org/10.3390/diagnostics12051152

Unger M, Kather JN (2024) Deep learning in cancer genomics and histopathology. Genome Med 16(1):44. https://doi.org/10.1186/s13073-024-01315-6

Usama M, Ahmad B, Yang J, Qamar S, Ahmad P, Zhang Y, Lv J, Guna J (2019) Removed: Equipping recurrent neural network with cnn-style attention mechanisms for sentiment analysis of network reviews. Comput Commun 148:98. https://doi.org/10.1016/j.comcom.2019.08.002

Vig L (2014) Comparative analysis of different classifiers for the Wisconsin breast cancer dataset. Open Access Libr J 1(06):1. https://doi.org/10.4236/oalib.1100660

Wang Z, Li R, Wang M, Li A (2021) Gpdbn: deep bilinear network integrating both genomic data and pathological images for breast cancer prognosis prediction. Bioinformatics 37(18):2963–2970. https://doi.org/10.1093/bioinformatics/btab185

Wang Y, Acs B, Robertson S, Liu B, Solorzano L, Wählby C, Hartman J, Rantalainen M (2022) Improved breast cancer histological grading using deep learning. Ann Oncol 33(1):89–98. https://doi.org/10.1016/j.annonc.2021.09.007

Wang Y, Ali MA, Vallon-Christersson J, Humphreys K, Hartman J, Rantalainen M (2023) Transcriptional intra-tumour heterogeneity predicted by deep learning in routine breast histopathology slides provides independent prognostic information. Eur J Cancer 191:112953. https://doi.org/10.1016/j.ejca.2023.112953

Wolberg William SN, Mangasarian Olvi WS (1995) Breast Cancer Wisconsin (Diagnostic). UCI Machine Learning Repository. https://doi.org/10.24432/C5DW2B

Woods K, Kegelmeyer WP, Bowyer K (1997) Combination of multiple classifiers using local accuracy estimates. IEEE Trans Pattern Anal Mach Intell 19(4):405–410. https://doi.org/10.1109/34.588027

Worsley K (1977) A non-parametric extension of a cluster analysis method by Scott and Knott. Biometrics. https://doi.org/10.2307/2529369

Yang J, Ju J, Guo L, Ji B, Shi S, Yang Z, Gao S, Yuan X, Tian G, Liang Y et al (2022) Prediction of her2-positive breast cancer recurrence and metastasis risk from histopathological images and clinical information via multimodal deep learning. Comput Struct Biotechnol J 20:333–342. https://doi.org/10.1016/j.csbj.2021.12.028

Yang Y, Hossain MZ, Stone EA, Rahman S (2023) Exemplar guided deep neural network for spatial transcriptomics analysis of gene expression prediction. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 5039–5048, https://doi.org/10.48550/arXiv.2210.16721

Yu D, Lin J, Cao T, Chen Y, Li M, Zhang X (2023) Secs: an effective cnn joint construction strategy for breast cancer histopathological image classification. J King Saud Univ-Comput Inform Sci 35(2):810–820. https://doi.org/10.1016/j.jksuci.2023.01.017

Zeng Y, Wei Z, Yu W, Yin R, Yuan Y, Li B, Tang Z, Lu Y, Yang Y (2022) Spatial transcriptomics prediction from histology jointly through transformer and graph neural networks. Brief Bioinform 23(5):297. https://doi.org/10.1093/bib/bbac297

Zerouaoui H, Idri A (2022) Deep hybrid architectures for binary classification of medical breast cancer images. Biomed Signal Process Control 71:103226. https://doi.org/10.1016/j.bspc.2021.103226

Zhang C, Bai Y, Yang C, Cheng R, Tan X, Zhang W, Zhang G (2022) Histopathological image recognition of breast cancer based on three-channel reconstructed color slice feature fusion. Biochem Biophys Res Commun 619:159–165. https://doi.org/10.1016/j.bbrc.2022.06.004

Zhao W, Wang L, Zhang Z (2019a) Atom search optimization and its application to solve a hydrogeologic parameter estimation problem. Knowl-Based Syst 163:283–304. https://doi.org/10.1016/j.knosys.2018.08.030

Zhao W, Wang L, Zhang Z (2019b) A novel atom search optimization for dispersion coefficient estimation in groundwater. Futur Gener Comput Syst 91:601–610. https://doi.org/10.1016/j.future.2018.05.037

Zhu X, Xu Y, Xu H, Chen C (2018) Quaternion convolutional neural networks. In: Proceedings of the European Conference on computer vision (ECCV), pp 631–647

Zou Y, Chen S, Che C, Zhang J, Zhang Q (2022) Breast cancer histopathology image classification based on dual-stream high-order network. Biomed Signal Process Control 78:104007. https://doi.org/10.1016/j.bspc.2022.104007