DOI: http://dx.doi.org/10.26483/ijarcs.v15i5.7127

Volume 15, No. 5, September-October 2024

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

REVIEW ARTICLE

Available Online at www.ijarcs.info

Deep Learning Techniques for Subtype Classification and Prognosis in Breast Cancer Genomics: A Systematic Review and Meta-Analysis

Ayodeji G. Abiodun
Department of Computer Science
School of Computing, Babcock University
Nigeria

Ernest E. Onuiri
Department of Computer Science
School of Computing, Babcock University
Nigeria

Abstract: Breast cancer is a significant global health concern, characterized by its complex and heterogeneous nature, which presents challenges for accurate diagnosis and effective treatment. Traditional classification methods for breast cancer subtypes and prognosis prediction often lack precision. In contrast, recent advances in deep learning have shown great potential to enhance diagnostic accuracy and improve patient outcomes by leveraging complex genomic data. This systematic review and meta-analysis aim to evaluate the effectiveness of deep learning models in classifying breast cancer subtypes and predicting prognosis. By focusing on studies published between January 2013 and February 2024, sourced from Scopus and PubMed databases, this review analyzes the performance of models such as Convolutional Neural Networks (CNNs) and Graph Convolutional Networks (GCNs). The results from 221 studies highlight that deep learning models significantly outperform traditional methods, achieving an average AUC of 0.893 and accuracy rates between 65.92% and 93%. These models demonstrate their ability to detect subtle genomic patterns associated with disease progression and patient outcomes, marking a substantial advancement in personalized medicine and breast cancer diagnostics.

Keywords: Breast Cancer Subtypes, Deep Learning, Diagnostic Accuracy, Genomic Data, Machine Learning Models, Personalized Medicine, Predictive Analytics, Prognosis

I. INTRODUCTION

The global health landscape is facing a formidable challenge with cancer, the predicted number of cases to reach 28.4 million in 2040, a significant increase from 2020's 19.3 million diagnoses and 10 million fatalities. This underscores the need for improved prevention and cure strategies [1]. Scientific and medical communities are focusing on early diagnosis and on the creation of suitable therapies for diverse types of Cancer [2].

Among Cancers affecting women, Breast cancer remains the most prevalent cancer worldwide, where there were about 685000 deaths and 2.3million new cases reported in 2020 according to GLOBOCAN [1]. Because it is molecularly heterogeneous, it has different clinical behaviors, treatment outcomes, and survival rates. However, conventional approaches such as clinicopathological characteristics or gene expression profiles microarray methods have their own limitations like poor prognostic value, inconsistent interpretation, and low resolution [3].

Consequently, advanced genomic technologies have been applied to unravel the complexity of breast cancer through identification of numerous molecular subtypes, such as luminal A, luminal B, HER2-enriched, basal-like and normal-like, individually are characterized by specific gene expression patterns which affect prognosis. However, accurately classifying and forecasting these subtypes remains challenging due to the disease's intricate and variable nature [4]. Deep learning, a division of artificial intelligence, has proven beneficial in analyzing genomic datasets and making precise predictions about patient outcomes.

Recent interdisciplinary research highlights the transformative potential of computational methodologies across diverse domains. For instance, advancements in big data analytics have been showcased in legal decision-making, demonstrating the power of data-driven insights [5]. Similarly, applications of deep learning in mental health diagnostics through speech recognition have shown promising results in detecting conditions like depression and PTSD, achieving high accuracy rates of 94% and 95% respectively [6]. Moreover, recent advancements in knowledge representation, exemplified by ontology-driven approaches, offer structured frameworks crucial for holistic understanding in biomedical research [7]. These developments underscore the broader applicability of computational methods.

In cancer research, deep learning, a subset of artificial intelligence, has proven beneficial in analyzing intricate genomic data and making precise predictions about patient outcomes [8]. Deep learning involves creating of artificial neural networks that can learn hierarchical data representations on their own. Neural networks have the ability to analyze vast genomic datasets, reveal patterns that are not easily noticeable, and provide precise predictions about patient outcomes using molecular characteristics [9].

Multiple research efforts have shown how deep learning methods can enhance the precision of categorizing cancer subtypes and predicting outcomes. For instance, [10] created a deep learning system that can accurately classify skin cancer in dermoscopic images at a level comparable to dermatologists. In the same way, [11] employed deep learning to forecast mutations and categorize histopathology pictures associated with non-small cell lung tumor. These studies emphasize the wide range of applications and success of deep learning in examining various biomedical datasets, such as genomic information from individuals with cancer. Additionally, recent systematic reviews, such as study on machine learning prediction models for colorectal cancer

patient survival using clinical data and gene expression profiles [12], highlight the broader significance of machine learning in addressing challenges in cancer research.

Despite the positive findings from these research studies, there is still a requirement for a thorough examination of the current literature on utilizing deep learning techniques for subtype classification and prognosis in breast cancer genomics. This systematic review and meta-analysis seek to fill this gap by summarizing existing evidence, assessing deep learning models' effectiveness, recognizing major obstacles and constraints, and suggesting future areas for research.

A. Rationale:

This rationale for this research is rooted in the growing recognition of deep learning techniques for improving the ability to classify and predict breast cancer subtypes. Conventional methods require manual examination of genomic information, which is time-consuming and susceptible to biases. Therefore, this evaluation seeks to assess deep learning techniques and their effectiveness using criteria like precision, recall, true negative rate, AUC, accuracy, and survival rates. The objective is to identify the most effective tactics and their influence on making clinical decisions regarding breast cancer patients.

B. Objectives:

The key intent of the meta-analysis and systematic review and is thus to summarize and give an overview on how deep learning is used for differentiating and predicting outcomes related to breast cancer. The review seeks to identify a number of significant issues with regards to the efficacy, accuracy and possible biases of deep learning models within this domain as outlines by PICOS framework. The review seeks to spotlight advanced methodologies and guide future research directions in breast cancer genomics by analyzing relevant research in detail. The PICOS framework directs the review to investigate key questions like:

1. Population: Which demographic and clinical characteristics are observed in breast cancer patients analyzed using deep learning methods for categorizing subtypes and forecasting outcomes?
2. Intervention: What deep learning models are specifically utilized in these studies, and what kinds of genomic data, such as gene expression profiles, are used as inputs for these models?
3. Comparison: How do different deep learning models excel in accurately identifying breast cancer subtypes and predicting patient outcomes using genomic data?
4. Result/outcome: Which deep learning models perform best in classifying subtypes and predicting the progression of breast cancer based on metrics such as the area under the curve (AUC), sensitivity, specificity, and accuracy?
5. Research/study Design: Evaluating the methodological soundness and potential bias in studies using deep learning models to identify and forecast breast cancer subtypes is essential. How do this research compare in regards to their design, population size, data management, model validation, and potential biases that could impact their findings?

II. METHODS

Following the PRISMA guidelines [13], a detailed search through of the literature was carried out using Scopus and PubMed databases. This inquiry specifically targeted publications written in English that were peer-reviewed and published from January 2013 to 27 February 2024. The search keywords employed were "deep learning," "AI," "subtype identification," "classification," "prognosis prediction," "breast cancer," and "genomics."

A. Scope of the Analysis:

This review specifically looks at research that have applied deep learning approach to categorize and forecast outcomes in breast cancer genomics.

B. Guidelines for Qualification:

Following the PICOS framework, the criteria for inclusion were determined as outlined below:

1. Population (P): Individuals who have received a diagnosis of breast cancer.
2. Intervention (I): Application of deep learning technologies in breast cancer genomics for subtype classification or prognosis prediction.
3. Comparison (C): Assessments that compare different deep learning techniques or juxtapose deep learning against traditional methods for determining breast cancer subtype or predicting prognosis.
4. Result/outcome(O): The results sought included measurements like precision, responsiveness, selectivity, area under the curve (AUC), survival percentages, and other relevant to the categorization or forecast of breast cancer genomics.
5. Research/study Design(S): The analysis included various types of studies, such as observational studies, clinical trials, experimental research, and validation studies which examined the use of deep learning for categorizing subtypes and predicting prognosis in breast cancer genomics.

C. Criteria for inclusion.

The systematic review's inclusion criteria were as listed below:

a) Research utilizing advanced deep learning methods to classify or predict the prognosis of breast cancer based on genomics.
b) Study with individuals who have been diagnosed with breast cancer.
c) Research that looks at the differences between various deep learning techniques or assesses how well deep learning models perform in comparison to traditional methods for categorizing breast cancer subtypes or predicting prognosis.
d) Research investigating breast cancer genomics in terms of subtype classification or prognosis prediction, analyzing accuracy, sensitivity, specificity, AUC, survival rates, and other relevant metrics.
e) Publications in the English language.
f) Articles that were published between January 2013 and February 2024.

D. Criteria for exclusion.

The guidelines for omitting studies from this review were determined in the following manner:

a) Research that did not utilize deep learning techniques or focused on interventions different from deep learning.
b) Research that did not yield results on classifying breast cancer subtypes or predicting prognosis, or explored outcomes unrelated to breast cancer.
c) Research involving animals, in vitro studies, or research not directly linked to breast cancer.

d) Non-English publications.
e) Articles for which the complete text was unavailable for assessment.
f) Varieties of papers like conferences and abstracts, letters to the editor and editorial pieces, case studies and systematic evaluations, and also meta-analytical reviews.

E. Data Source:

The comprehensive search was carried out in Scopus and PubMed databases to locate relevant studies, including the medical subject headings (MeSH) and terms connected to deep learning, AI, subtype identification, prognosis assessment, breast cancer, and genomics. The search method was adjusted for the specific requirements of each database to guarantee a thorough investigation [14]

F. Search Strategy:

The approach to searching was methodically formulated to include comprehensive range of literature representing the updated research, following specified criteria for inclusion:

a) Study using deep learning for classification and prediction of outcomes in breast cancer genomics.
b) Study involving Patients with confirmed diagnosis of breast cancer.
c) Studies focused on categorizing or forecasting in breast cancer genomics, highlighting metrics like accuracy, sensitivity, specificity, AUC, and survival rates, among others.
d) Articles that were published in English from January 2013 to February 2024.

The systematic review titled Deep Learning technique for Subtype Classification and Prognosis in Breast Cancer Genomics utilized thorough search protocols on both Scopus and PubMed databases.

Approach for utilizing the Scopus Database:

The approach for finding relevant literature in the Scopus database was organized as follows:

TITLE-ABS-KEY (("deep learning" OR "neural network" OR "machine learning" OR "artificial intelligence" OR "predictive modeling") AND ((subtype OR subset OR branch* OR subpopulat*) AND (classif* OR sort* OR group*)) AND (prognos* OR predict* OR forecast* OR projection) AND ("breast cancer" OR "breast neoplasm" OR "mammary carcinoma") AND (genom* OR genet*)) AND PUBYEAR > 2013 AND PUBYEAR < 2025 AND (LIMIT-TO (SRCTYPE , "j")) AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (SUBJAREA , "BIOC") OR LIMIT-TO (SUBJAREA , "MEDI") OR LIMIT-TO (SUBJAREA , "COMP") OR LIMIT-TO (SUBJAREA , "PHAR") OR LIMIT-TO (SUBJAREA , "IMMU") OR LIMIT-TO (SUBJAREA , "NEUR") OR LIMIT-TO (SUBJAREA , "HEAL")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (LANGUAGE , "English")). **Which resulted in 148 documents.**

PubMed Database Strategy:

The following query string was designed for PubMed database to include:

(("deep learning"[All Fields] OR "neural network"[All Fields] OR "machine learning"[All Fields] OR "artificial intelligence"[All Fields] OR "predictive modeling"[All Fields]) AND (("subtype"[All Fields] OR "subtyped"[All Fields] OR "subtypes"[All Fields] OR "subtyping"[All Fields] OR "subtypings"[All Fields] OR ("subset"[All Fields] OR "subsets"[All Fields]) OR "branch*"[All Fields] OR "subpopulat*"[All Fields]) AND ("classif*"[All Fields] OR "sort*"[All Fields] OR "group*"[All Fields])) AND ("prognos*"[All Fields] OR "predict*"[All Fields] OR "forecast*"[All Fields] OR ("forecasting"[MeSH Terms] OR "forecasting"[All Fields] OR "projected"[All Fields] OR "projecting"[All Fields] OR "projection"[MeSH Terms] OR "projection"[All Fields] OR "projections"[All Fields] OR "projectional"[All Fields])) AND ("breast cancer"[All Fields] OR "breast neoplasm"[All Fields] OR "mammary carcinoma"[All Fields]) AND ("genom*"[All Fields] OR "genet*"[All Fields])) AND ((fha[Filter]) AND (fft[Filter]) AND (humans[Filter]) AND (female[Filter]) AND (english[Filter]) AND (2014:2024[pdat])). **This strategy yielded 73 results.**

G. Data Management:

The results of the searches conducted on Scopus and PubMed were saved as CSV files and imported into Rayyan [15] to remove any duplicate entries. Titles, abstracts, and full texts were carefully examined to effectively narrow down 221 articles meeting the criteria, with the search concluding on February 27, 2023.

H. Selection of studies:

The selection process was skillfully carried out with the use of Rayyan [16], a tool for managing systematic reviews, in order to prevent duplication. This process involved carefully examining the titles and abstracts to ensure they met the inclusion criteria and checking their content to determine their relevance. Research that met the criteria was kept for in-depth analysis, whereas studies that didn't meet the criteria were excluded.

I. Data Extraction:

The study extensive data collection concentrated on gathering crucial information, such as research characteristics, participant demographics, total number of participants, clinical findings, trends in gene expression, deep learning methods used, prediction models, and resulting outcomes. All the specifics were carefully and extensively recorded. In cases of disagreement, conversations were started, or an outside opinion was sought to reach agreement. The primary aim of this com prehensive collection was to gather all essential information needed to successfully meet the study's goals. This involved describing the characteristics of breast cancer patients, determining the deep learning models utilized, documenting the evaluation criteria, and examining the research methodology and potential biases in the analyzed studies. The complete procedure, starting from study selection to data compilation, was meticulously recorded following PRISMA guidelines to guarantee the transparency and reproducibility of the research, as depicted in Figure 1 [17]

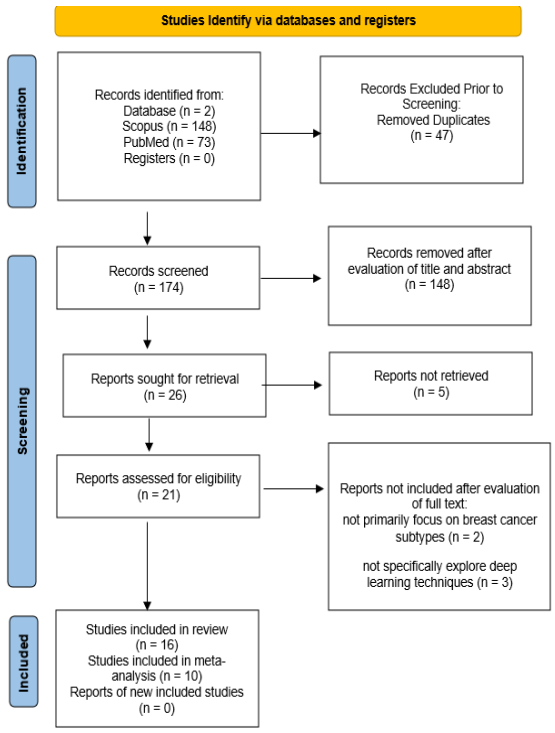


Figure 1. The screened studies documented using PRISMA flowchart

J. Bias assessment

The bias assessment was conducted carefully, following a rigorous screening protocol created by the researchers. This process required strict adherence to established criteria to exclude studies that did not conform to the specified search parameters or were irrelevant to the main focus of the review. By utilizing this systematic selection process, the study aimed to minimize bias and ensure that only relevant and high-quality articles were included in the review.

III. OUTCOME

Researchers assessed the effectiveness of various deep learning models in predicting the outcomes of breast cancer genomics in patient through a review and meta-analysis on deep learning methods for subtype classification and prognosis. The research emphasized the value of utilizing high-throughput genomic information for precise categorization and prediction. The review highlighted the potential of artificial intelligence in enhancing accuracy, sensitivity, specificity, and survival rates in breast cancer research by contrasting deep learning methods with traditional approaches like manual interpretation of genomic data. This thorough study explores using deep learning models to refine clinical decision-making and enhance patient care in breast cancer genomics.

A. Investigating Possible Progressions

Recent studies in breast cancer genomics highlights the effectiveness of deep learning methods in accurately identifying subtypes and predicting prognoses. By analyzing and integrating intricate data patterns in large genomic datasets, advanced deep learning technologies can greatly improve the precision and reliability of results for patients [2]. Researchers are committed to overcoming common challenges and limitations in traditional methods by utilizing these advancements, aiming to develop more personalized and effective techniques in breast cancer research. This analysis of how deep learning can be used in genomics underlines the crucial impact of artificial intelligence in transforming our knowledge and approaches to dealing with breast cancer.

Table I. Summary of the selected studies

S/N	*Study title*	*Result and Findings*	*Deep learning model*	*AUC*	*Accuracy*
1	A deep learning image-based intrinsic molecular subtype classifier of breast tumors reveals tumor heterogeneity that may affect survival [18].	The image-based classifier utilizing deep learning accurately categorized most of the Samples from a distinct group of tumors. Substantial diversity was noted in assigned subtypes Within an individual whole slide image., indicating that sophisticated deep machine learning techniques using solely whole slide images that are regularly collected could mimic tests molecular due to RNA sequencing, like PAM50. Additionally, patients who had tumors categorized as heterogeneous experienced survival rates that fell between those of Luminal A and Basal patients.	RBF kernel with Multiclass one-vs-rest SVM	0.8259	Overall: 65.92% for unselected test patients, 56.73% for low-confidence test patients
2	Attention-based GCN Integrates Multi-omics Data for Breast Cancer Subtype Classification and Patient-specific Gene Marker Identification [19].	During 5-fold cross-validation, the cAGCN model performed better than other techniques, reaching AUC score at 0.9816, accuracy at 0.8743, and Matthew's correlation coefficient at 0.8151. LRP found individualized biomarkers linked to the growth and advancement of breast tumor, highlighting the effectiveness of Graph Convolutional Networks (GCN) and attention mechanisms in evaluating diverse omics data. data for classifying subtypes and offering relevant biological insights into model interpretations.	Column-wise Attention-based GCN	0.9816	0.8793
3	Breast Cancer Type Classification Using Machine Learning [20].	In the four machine learning algorithms examined, the SVM algorithm turned out the highest accuracy, recall and specificity values in differentiating breast cancer as TNBC and non-TNBC. SVM had lower misclassification error rate than other algorithms. Results indicate that ML algorithms are able to distinguish various types of breast cancer and thus may serve as a starting point for further investigations	SVM, K-nearest neighbor, Naïve Bayes, and Decision Trees.	N/A	SVM: 90%, kNN: 87%, NB and DT: Not provided separately

		into subtype classification and use of genomic data in personalized treatments strategies.			
4	Prognostic power assessment of clinical parameters to predict neoadjuvant response therapy in HER2-positive breast cancer patients: A machine learning approach [21].	The model, developed on a subset of key characteristics, obtained an AUC of 0.732 and accuracy of 71.67%, along with sensitivity and specificity rates of 72.58% and 72.22%, in that order. The model found that the status of ER, Pgr, and HER2 score were significantly linked to pCR to NAC. The findings indicate that relying only on clinical features may not be enough to create a clinical support system for predicting neoadjuvant therapy response. This highlights the necessity for more research in larger validation studies and considering the use of radiomics analysis of biomedical images.	Random Forest	73.27%	71.67%
5	Classifying Breast Cancer Subtypes Using Deep Neural Networks Based on Multi-Omics Data [22].	The DeepMO model, utilizing multi-omics data, outperformed single omics data and methods such as MKL in binary classification tasks. It demonstrated superior predictive accuracy in multi-classification when compared to alternative data integration techniques. Feature selection was discovered to enhance the model's effectiveness. Examination of important genes showed an abundance of gene ontology terms and biological pathways connected to cell cycle and morphogenesis, highlighting their importance in differentiating breast cancer subtypes.	DeepMO	0.908	0.782
6	Classifying Breast Cancer Subtypes Using Multiple Kernel Learning Based on Omics Data [23].	Utilizing the SMO-MKL algorithm to merge multi-omics data resulted in a notable enhancement in accurately breast cancer subtypes classification and increasing AUC, as opposed to relying on single omics data methods. The model showed improved ability in differentiating various types of breast tumor, such as luminal A, luminal B, HER2-positive, and (TNBC). Feature selection pinpointed important genes and pathways which are close to the onset alongside progression of breast cancer, providing understanding of the underlying mechanisms that distinguish different subtypes of the disease. The research showed how merging various Omics data improve precision to identify breast tumor subtypes, emphasizing the need for a holistic approach in cancer biology and clinical decision-making.	SMO-MKL	0.916	0.798
7	Clinicopathological Features of Triple-Negative Breast Cancer Epigenetic Subtypes [24].	Four distinct TNBC epitypes were discovered, known as Epi-CL-A, Epi-CL-B, Epi-CL-C, and Epi-CL-D. Patients diagnosed with Epi-CL-B cancer experienced notably reduced disease-free and overall survival rates. Gene expression and mutation variations in TNBC epitypes indicate potential alternate pathway activation, which may be targeted for supplementary therapy. The recently identified TNBC transcriptomic subtypes were enhanced by the addition of epigenetic subtypes, showing notable clinical and molecular distinctions among the epitypes.	Machine Learning-based Epigenetic Classifiers	N/A	N/A
8	Deep-learning approach to identifying cancer subtypes using high-dimensional genomic data [25].	DeepType drastically surpassed current techniques in detecting more resilient cancer subtypes using fewer genes. It showed the ability to efficiently manage large datasets with very high dimensionality and was resilient to label noise. The research introduced a fresh approach for identifying precise and resilient molecular cancer subcategories through intricate genomic information.	DeepType	N/A	N/A
9	Predicting Breast Cancer Gene Expression Signature by Applying Deep Convolutional Neural Networks from Unannotated Pathological Images [26].	Deep learning models trained on multiple types of omics data outperformed single omics data and traditional methods such as MKL in binary classification tasks. They also demonstrated better predictive accuracy in multi-classification in comparison to alternative data integration methods. Feature selection was discovered to enhance the performance of the model. Examination of important genes showed an abundance in terms of gene ontology and biological pathways linked to cell cycle and morphogenesis, highlighting their importance in differentiating breast cancer subtypes.	VGG16, ResNet50, ResNet101, Xception, ResNet101_imgnet	0.88-0.94	0.68-0.78

10	Moanna: Multi-Omics Autoencoder-Based Neural Network Algorithm for Predicting Breast Cancer Subtypes [27].	Moanna excelled in accurately predicting breast cancer subtypes and ER status, surpassing traditional approaches and showcasing successful incorporation of multi-omics information. It could identify important biological patterns and had the potential to improve breast cancer subtype classification, as it strongly correlated with patient survival rates more than current PAM50 subtypes. The findings show Moanna's skill in utilizing advanced genetic information for better categorization of cancer subtypes and its possible use in improving clinical results by providing more accurate molecular characterization.	Moanna	0.88-0.94	ER: 96%, Basal-like: 98%, PAM50: 85%
11	moBRCA-net: a breast cancer subtype classification framework based on multi-omics attention neural networks [28].	Experimental findings validated moBRCA-net's superior performance compared to alternative methods, showcasing the efficacy of multi-omics integration and focus on omics-level analysis. The method helped in grasping the significance of features in various omics layers for predicting breast cancer subtypes more accurately using deep learning methods, showing potential for improvement. This research emphasized the importance of combining multiple omics data and the effectiveness of attention mechanisms in pinpointing essential attributes for categorizing diseases.	moBRCA-net	N/A	0.891
12	omicsGAT: Graph Attention Network for Cancer Subtype Analyses [29].	omicsGAT showed better results in combining sample neighborhood information and creating an embedding vector to enhance the prediction of disease phenotypes, stratify cancer patients, and cluster cells. The attention matrix, formed by the multi-head attention coefficients, produced valuable insights than the correlation-based adjacency matrix, resulting in improved classification and clustering results. The model's effectiveness in cancer subtype analysis and potential use in precision medicine were demonstrated by its capacity to recognize significant neighbors, incorporate various types of omics data, and be easily interpretable through the attention mechanism.	omicsGAT	ER: 0.9636, PR: 0.9065, TN: 0.9611	N/A
13	Reducing variability of breast cancer subtype predictors by grounding deep learning models in prior knowledge [30].	Integrating previous information to the loss function of deep learning models greatly decreased predictor inconsistency and improved reliability while maintaining model performance with the accuracy. The consistency of pathway enrichment analysis improved, leading to enhanced generalization capabilities of the model. The research showed that incorporating biological information into machine learning systems can enhance the precision of predictions and make deep learning models in breast cancer subtype classification more understandable.	Custom Neural Network	N/A	N/A
14	RNA-Seq-Based Breast Cancer Subtypes Classification Using Machine Learning Approaches [31].	Utilizing weighted differentially expressed genes in binary classification for each subtype could accurately predict outcomes for new samples, demonstrating the efficiency of the suggested methods. The innovative refined Gene Ontology terms applied through GOEGCN for both the control and the experimental groups of each subtype illustrated variations in biological functions to an extent, based on the bilateral division of co-expression network structures. The research emphasized how crucial it is to incorporate regulatory data when choosing DEGs, and the ability of GOEGCN to uncover alterations in particular biological functions across breast cancer subtypes.	Naive Bayes, Random Forest, SVM	Basal-like: 0.9847, Her2: 0.9562, Luminal A: 0.9134, Luminal B: 0.9075, Normal-like: 0.9125	Basal-like: 0.9607, Her2: 0.8868, Luminal A: 0.8637, Luminal B: 0.8383, Normal-like: 0.8502
15	DeepTRIAGE: interpretable and individualised biomarker scores using attention mechanism for the classification of breast cancer sub-types [32]	DeepTRIAGE accurately categorizes different types of cancer and also gives understandable, personalized scores for biomarkers, revealing hidden variations within each type. The utilization of the model in luminal A and B breast tumor subtypes identifies important genes and gene sets, even those not included in the PAM50 signature. This indicates the model's ability to improve clinical decision-making and propose novel research ideas regarding breast cancer diversity.	DeepTRIAGE	N/A	N/A
16	Deep learning generates custom-made logistic regression models for explaining	The results indicated (PWL) model used genes associated to pathways related to the cell cycle for analyzing breast cancer subtypes, indicating that deep learning can predict cancer subtypes accurately and	PWL Model	RNA-seq: Up to 0.985, Copy	N/A

	how breast cancer subtypes are classified [33].	might reveal new and unconventional knowledge about gene expressions. This analysis approach helped uncover the underlying processes of breast cancer, and demonstrated its ability to enhance patient outcomes by offering a deeper understanding of the molecular profiles of cancer subtypes using advanced learning methods.		number: Up to 0.862	

B. Findings.

This systematic review examined 16 studies and meta-analysis review the model in 10 studies which demonstrated the flexibility and efficiency of deep learning in categorizing and predicting outcomes for different types of breast cancer. The studies revealed a wide range of methods including CNNs, GCNs, MKL, autoencoder-based networks, attention mechanisms, and GATs. Utilizing multiple omics data sources, including gene expression, DNA methylation, copy number variations, and pathological images, these researches showed how deep learning can handle the complexity of breast cancer genomics. One important finding from these studies is the variation in AUC values, showing a strong ability to differentiate between different types of breast cancer. This confirms capability of deep learning to enhance the accuracy of diagnosing and predicting outcomes in breast cancer research. This thorough examination highlights how deep learning technologies have the potential in advancing our understanding and treatment of breast cancer, leading to more customized and accurate medical treatments.

Table II. Data Extracted for Meta-Analysis

Year	***Deep Learning Model***	***AUC***	***Accuracy***
2020	Multiclass one-vs-rest SVM with RBF kernel	0.826	65.92%
2022	Column-wise Attention-based GCN	0.982	87.93%
2021	Support Vector Machines (SVM)	-	90%
2023	Random Forest	0.733	71.67%
2020	DeepMO	0.908	78.2%
2019	SMO-MKL	0.916	79.8%
2023	Machine Learning-based Epigenetic Classifiers	-	-
2020	DeepType	-	-
2021	VGG16, ResNet50, ResNet101, Xception, ResNet101_imgnet	0.91	73%
2023	Moanna	0.91	93%
2023	PWL Model	0.985	N/A
2023	moBRCA-net	-	89.1%
2021	Custom Neural Network	-	-
2022	omicsGAT	0.964	-
2020	DeepTRIAGE	-	-
2020	Naive Bayes, Random Forest, SVM	0.935	88%

Table III. Deep learning model and AUC of selected studies for meta-analysis

Year	***Model***	***AUC***
2020	Multiclass SVM	0.826
2022	Column-wise GCN	0.982
2021	Random Forest	0.733
2023	PWL	0.985
2020	DeepMO	0.908
2019	SMO-MKL	0.916
2021	VGG16, ResNet50, ResNet101, Xception	0.910
2023	Moanna	0.910
2023	omicsGAT	0.964
2020	Naive Bayes, RF, SVM	0.935

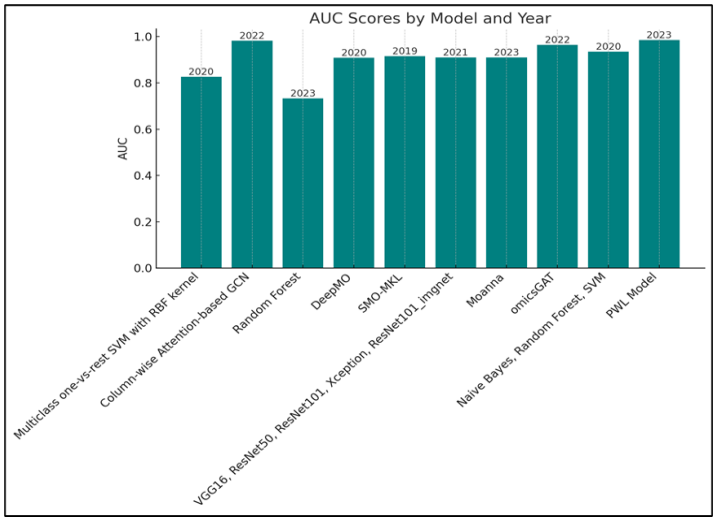
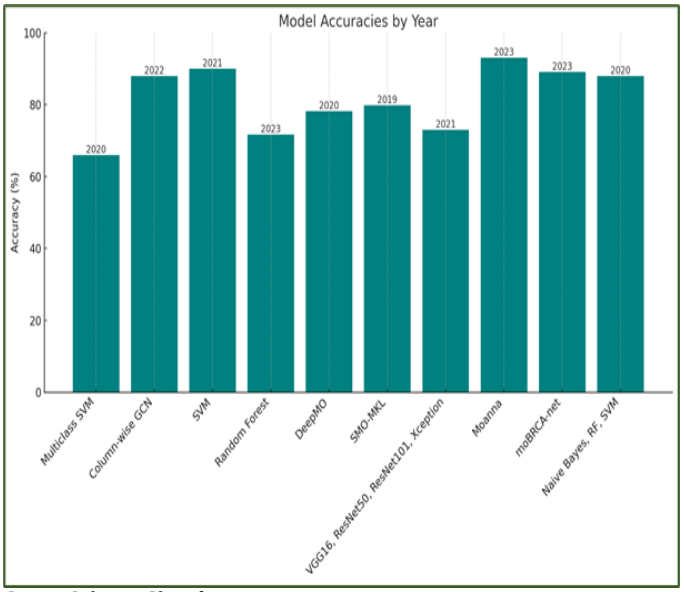


Figure 2. Annual AUC Performance of Deep Learning Models in Breast Cancer Subtype Classification

Table IV. Deep learning model and Accuracy of selected studies for meta-analysis

Year	***Model***	***Accuracy***
2020	Multiclass SVM	65.92%
2022	Column-wise GCN	87.93%

2021	SVM	90%
2023	Random Forest	71.67%
2020	DeepMO	78.2%
2019	SMO-MKL	79.8%
2021	VGG16, ResNet50, ResNet101, Xception	73%
2023	Moanna	93%
2023	moBRCA-net	89.1%
2020	Naive Bayes, RF, SVM	88%



Cancer Subtype Classification

Figure 3. Yearly Trend Analysis of Model Accuracy in Breast Cancer Genomic Prediction.

C. Meta-Analysis Report on Selected Studies

This meta-analysis investigates the most effective deep learning models at categorizing breast cancer subtypes, analyzing both AUC scores and accuracy percentages. Examining ten deep learning models between 2019 and 2023, this report seeks to provide a comprehensive view of model performance, highlighting advancements and obstacles in the important medical sector.

D. Analysis of AUC Scores

The average AUC value from the models studied is around 0.893, indicating an excellent performance in classifying breast cancer subtypes. The models displayed variety of AUC scores ranging from 0.733 to 0.985, demonstrating notable differences in their ability to discriminate. Exceptional accuracy in differentiating cancer subtypes is demonstrated by high-performing models such as the "PWL Model" (AUC: 0.985), "Column-wise Attention-based GCN" (2022, AUC: 0.982), and "OmicsGAT" (AUC: 0.964). On the contrary, the Random Forest model (AUC: 0.733) achieved the lowest score, indicating areas that can be enhanced in future versions.

The AUC scores' standard deviation, estimated at around 0.076, indicates a fairly narrow spread around the average, even though some outliers display greatly superior or slightly inferior performances. This difference highlights the fast progress in model skills and the various methods for addressing breast cancer subtype classification.

E. Analysis of the percentage of accuracy.

In terms of model accuracy, the meta-analysis found that the selected deep learning models varied in effectiveness, with accuracy rates ranging from 65.92% to 93%. The mean accuracy of the models was around 81.4%, showcasing potential for enhancing performance to achieve more stable high-level results. The success of the latest models like the "Moanna" (2023, Accuracy: 93%) and the "SVM" (2021, Accuracy: 90%) highlights their ability to offer extremely precise classifications.

The fluctuation in precision, similar to the AUC scores, demonstrates the changing environment of deep learning uses in breast cancer diagnosis. The advancement in recent years indicates a favorable shift toward models that have improved accuracy in predicting breast cancer subtypes, essential for customized treatment plans.

F. Analysis comparing and potential future ways forward.

Comparing AUC scores and accuracy rates provides a more detailed understanding of model performance, showing that certain models may excel in one aspect but not necessarily in the other. This dual nature highlights the significance of taking into account numerous performance indicators when assessing models for clinical use. Future research needs to concentrate on closing the distance between high AUC scores and accuracy rates, making sure that models can effectively differentiate between cancer subtypes and do so reliably. Moreover, delving into ensemble methods and advanced machine learning techniques may also improve the model's performance. Giving importance to understanding how models work and incorporating different types of data sources could offer a better understanding of the intricate characteristics of breast cancer, resulting in more precise and customized diagnostic instruments.

IV. DISCUSSION

This comprehensive meta-analysis assesses the effectiveness of different deep learning models used in categorizing breast cancer subtypes and predicting prognosis, with a specific emphasis on AUC scores and accuracy rates from research conducted from 2019 to 2023. The study highlighted the strong effectiveness and significant differences between models, with AUC scores ranging from 0.733 to 0.985 and accuracy rates from 65.92% to 93%. The diverse performance levels of models like the "PWL Model," "Column-wise Attention-based GCN," and "OmicsGAT" showcase the continuous improvement of deep learning which enables the medical researchers to arrive at precise and rapid diagnoses. However, the Random Forest model's decreased performance indicates opportunities for improvement, highlighting the complex nature of deep learning in cancer genomics.

The diversity in computational approaches used in these models and the complexity of breast cancer itself are both reflected in the observed variability. This highlights an important concept: combining multi-omics data and advanced computational techniques can improve greatly the accuracy and credibility of subtypes of breast cancer classification, which is

essential for personalized treatment plans. Nonetheless, the obstacles mentioned, such as the requirement for explicable models and the incorporation of extensive data sources, lay out a direction for future research efforts.

V. CONCLUSION

The results of this systematic review and meta-analysis demonstrate an intriguing story of how deep learning is playing a growing part in the fight against breast cancer. Deep learning is leading the way in transforming breast cancer subtype classification and prognosis prediction with its advancements in technology and model performance. The study emphasizes how these models can improve diagnostic accuracy and bring about a new era of personalized medicine that caters to the genetic complexities of each patient. Future studies, ready to fill the identified gaps, have the potential to improve clinical results by combining advanced machine learning methods with multi-dimensional genomic data. Progressing ahead will lead to a greater emphasis on improving the understandability and dependability of models, as well as integrating genomic knowledge fully, making sure that deep learning remains a crucial tool in fighting breast cancer for oncologists.

VI. REFERENCES

[1] H. Sung et al., "Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries," CA Cancer J Clin, vol. 71, no. 3, pp. 209–249, May 2021, doi: 10.3322/CAAC.21660.

[2] J. Wardle, K. Robb, S. Vernon, and J. Waller, "Screening for prevention and early diagnosis of cancer," Am Psychol, vol. 70, no. 2, pp. 119–133, Feb. 2015, doi: 10.1037/A0037357.

[3] C. M. Perou et al., "Molecular portraits of human breast tumours," Nature 2000 406:6797, vol. 406, no. 6797, pp. 747–752, Aug. 2000, doi: 10.1038/35021093.

[4] C. Curtis et al., "The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups," Nature, vol. 486, no. 7403, pp. 346–352, Jun. 2012, doi: 10.1038/NATURE10983.

[5] O. C. Ngige, F. Y. Ayankoya, J. A. Balogun, E. Onuiri, C. Agbonkhese, and F. A. Sanusi, "A dataset for predicting Supreme Court judgments in Nigeria," Data Brief, vol. 50, p. 109483, Oct. 2023, doi: 10.1016/J.DIB.2023.109483.

[6] S. A. Bindas and E. E. Onuiri, "A DEEP LEARNING APPROACH TO SPEECH RECOGNITION FOR DETECTION OF MENTAL DISORDERS," CTICTR, vol. 2, no. 1, pp. 28–46, 2023.

[7] O. Ernest, O. Komolafe, S. O., and A. Oludele, "Ontology: A Case for Disease and Drug Knowledge Discovery," Communications on Applied Electronics, vol. 5, no. 9, pp. 6–13, Sep. 2016, doi: 10.5120/CAE2016652362.

[8] J. Zou, M. Huss, A. Abid, P. Mohammadi, A. Torkamani, and A. Telenti, "A primer on deep learning in genomics," Nat Genet, vol. 51, no. 1, pp. 12–18, Jan. 2019, doi: 10.1038/S41588-018-0295-5.

[9] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," Nature 2015 521:7553, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.

[10] A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," Nature 2017 542:7639, vol. 542, no. 7639, pp. 115–118, Jan. 2017, doi: 10.1038/nature21056.

[11] N. Coudray et al., "Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning," Nat Med, vol. 24, no. 10, pp. 1559–1567, Oct. 2018, doi: 10.1038/S41591-018-0177-5.

[12] E. E. Onuiri, O. Akande, O. B. Kalesanwo, T. Adigun, K. Rosanwo, and K. C. Umeaka, "A Systematic Review of Machine Learning Prediction Models for Colorectal Cancer Patient Survival Using Clinical Data and Gene Expression Profiles," Revue d'Intelligence Artificielle, vol. 37, no. 5, pp. 1273–1280, 2023, doi: 10.18280/RIA.370520.

[13] D. Moher et al., "Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement," PLoS Med, vol. 6, no. 7, Jul. 2009, doi: 10.1371/JOURNAL.PMED.1000097.

[14] M. M. Voets, J. Veltman, C. H. Slump, S. Siesling, and H. Koffijberg, "Systematic Review of Health Economic Evaluations Focused on Artificial Intelligence in Healthcare: The Tortoise and the Cheetah," Value Health, vol. 25, no. 3, pp. 340–349, Mar. 2022, doi: 10.1016/J.JVAL.2021.11.1362.

[15] "Rayyan: a systematic reviews web app for exploring and filtering searches for eligible studies for Cochrane Reviews | Colloquium Abstracts." Accessed: Mar. 04, 2024. [Online]. Available: https://abstracts.cochrane.org/2014-hyderabad/rayyan-systematic-reviews-web-app-exploring-and-filtering-searches-eligible-studies

[16] M. Ouzzani, H. Hammady, Z. Fedorowicz, and A. Elmagarmid, "Rayyan-a web and mobile app for systematic reviews," Syst Rev, vol. 5, no. 1, pp. 1–10, Dec. 2016, doi: 10.1186/S13643-016-0384-4/FIGURES/6.

[17] J. A. C. Sterne et al., "RoB 2: a revised tool for assessing risk of bias in randomised trials," BMJ, vol. 366, 2019, doi: 10.1136/BMJ.L4898.

[18] M. I. Jaber et al., "A deep learning image-based intrinsic molecular subtype classifier of breast tumors reveals tumor heterogeneity that may affect survival," Breast Cancer Research, vol. 22, no. 1, Jan. 2020, doi: 10.1186/s13058-020-1248-3.

[19] H. Guo, X. Lv, Y. Li, and M. Li, "Attention-based GCN integrates multi-omics data for breast cancer subtype classification and patient-specific gene marker identification," Brief Funct Genomics, vol. 22, no. 5, pp. 463–474, 2023, doi: 10.1093/bfgp/elad013.

[20] J. Wu and C. Hicks, "Breast cancer type classification using machine learning," J Pers Med, vol. 11, no. 2, pp. 1–12, 2021, doi: 10.3390/jpm11020061.

[21] A. Fanizzi et al., "Prognostic power assessment of clinical parameters to predict neoadjuvant response therapy in HER2-positive breast cancer patients: A machine learning approach," Cancer Med, vol. 12, no. 22, pp. 20663–20669, Nov. 2023, doi: 10.1002/cam4.6512.

[22] Y. Lin, W. Zhang, H. Cao, G. Li, and W. Du, "Classifying breast cancer subtypes using deep neural networks based on multi-omics data," Genes (Basel), vol. 11, no. 8, pp. 1–18, Aug. 2020, doi: 10.3390/genes11080888.

[23] M. Tao et al., "Classifying Breast Cancer Subtypes Using Multiple Kernel Learning Based on Omics Data.," Genes (Basel), vol. 10, no. 3, Mar. 2019, doi: 10.3390/genes10030200.

[24] M. L. DiNome et al., "Clinicopathological Features of Triple-Negative Breast Cancer Epigenetic Subtypes," Ann Surg Oncol, vol. 26, no. 10, pp. 3344–3353, Oct. 2019, doi: 10.1245/s10434-019-07565-8.

[25] R. Chen, L. Yang, S. Goodison, and Y. Sun, "Deep-learning approach to identifying cancer subtypes using high-dimensional genomic data," Bioinformatics, vol. 36, no. 5, pp. 1476–1483, 2020, doi: 10.1093/bioinformatics/btz769.

[26] N. N. Phan, C.-C. Huang, L.-M. Tseng, and E. Y. Chuang, "Predicting Breast Cancer Gene Expression Signature by Applying Deep Convolutional Neural Networks From Unannotated Pathological Images," Front Oncol, vol. 11, 2021, doi: 10.3389/fonc.2021.769447.

[27] R. Lupat, R. Perera, S. Loi, and J. Li, "Moanna: Multi-Omics Autoencoder-Based Neural Network Algorithm for Predicting Breast Cancer Subtypes," IEEE Access, vol. 11, pp. 10912–10924, 2023, doi: 10.1109/ACCESS.2023.3240515.

[28] J. M. Choi and H. Chae, "moBRCA-net: a breast cancer subtype classification framework based on multi-omics attention neural networks," BMC Bioinformatics, vol. 24, no. 1, Dec. 2023, doi: 10.1186/s12859-023-05273-5.

[29] S. Baul, K. T. Ahmed, J. Filipek, and W. Zhang, "omicsGAT: Graph Attention Network for Cancer Subtype Analyses," Int J Mol Sci, vol. 23, no. 18, 2022, doi: 10.3390/ijms231810220.

[30] P. Anderson, R. Gadgil, W. A. Johnson, E. Schwab, and J. M. Davidson, "Reducing variability of breast cancer subtype predictors by grounding deep learning models in prior knowledge," Comput Biol Med, vol. 138, Nov. 2021, doi: 10.1016/j.compbiomed.2021.104850.

[31] Z. Yu, Z. Wang, X. Yu, and Z. Zhang, "RNA-Seq-Based Breast Cancer Subtypes Classification Using Machine Learning Approaches," Comput Intell Neurosci, vol. 2020, 2020, doi: 10.1155/2020/4737969.

[32] A. Beykikhoshk, T. P. Quinn, S. C. Lee, T. Tran, and S. Venkatesh, "DeepTRIAGE: Interpretable and individualised biomarker scores using attention mechanism for the classification of breast cancer sub-Types," BMC Med Genomics, vol. 13, Feb. 2020, doi: 10.1186/s12920-020-0658-5.

[33] T. Shibahara et al., "Deep learning generates custom-made logistic regression models for explaining how breast cancer subtypes are classified," PLoS One, vol. 18, no. 5 May, May 2023, doi: 10.1371/journal.pone.0286072.

