

# Deep Learning Based Methods for Breast Cancer Diagnosis: A Systematic Review and Future Direction

Maged Nasser  and Umi Kalsom Yusof \*

School of Computer Sciences, Universiti Sains Malaysia, Gelugor 11800, Penang, Malaysia

\* Correspondence: umiyusof@usm.my

**Abstract:** Breast cancer is one of the precarious conditions that affect women, and a substantive cure has not yet been discovered for it. With the advent of Artificial intelligence (AI), recently, deep learning techniques have been used effectively in breast cancer detection, facilitating early diagnosis and therefore increasing the chances of patients' survival. Compared to classical machine learning techniques, deep learning requires less human intervention for similar feature extraction. This study presents a systematic literature review on the deep learning-based methods for breast cancer detection that can guide practitioners and researchers in understanding the challenges and new trends in the field. Particularly, different deep learning-based methods for breast cancer detection are investigated, focusing on the genomics and histopathological imaging data. The study specifically adopts the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA), which offer a detailed analysis and synthesis of the published articles. Several studies were searched and gathered, and after the eligibility screening and quality evaluation, 98 articles were identified. The results of the review indicated that the Convolutional Neural Network (CNN) is the most accurate and extensively used model for breast cancer detection, and the accuracy metrics are the most popular method used for performance evaluation. Moreover, datasets utilized for breast cancer detection and the evaluation metrics are also studied. Finally, the challenges and future research direction in breast cancer detection based on deep learning models are also investigated to help researchers and practitioners acquire in-depth knowledge of and insight into the area.



**Citation:** Nasser, M.; Yusof, U.K. Deep Learning Based Methods for Breast Cancer Diagnosis: A Systematic Review and Future Direction. *Diagnostics* **2023**, *13*, 161. <https://doi.org/10.3390/diagnostics13010161>

Academic Editors: Kwangsoo Kim and Chang Ik Yoon

Received: 26 November 2022

Revised: 19 December 2022

Accepted: 19 December 2022

Published: 3 January 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** deep learning; artificial neural network; breast cancer; cancer detection; breast cancer diagnosis

## 1. Introduction

One of the common cancers identified globally among women is breast cancer, which has become the leading cause of death [1–3]. Based on the recent findings by the American cancer society, over 40,000 women and about 600 men died as a result of breast cancer disease [3]. There are four basic forms of breast cancer: benign, normal, in situ carcinoma and invasive carcinoma [1]. A benign tumor slightly alters the breast's anatomy, it is not toxic and does not fit the description of dangerous cancer [4]. In situ carcinoma, on the other hand, only affects the system of mammary duct lobules and does not spread to other organs [5]. This kind of cancer is not very harmful and is treatable if detected early. The most severe type of breast cancer is invasive carcinoma, which has the potential to spread to all other organs [6]. Over many years, breast cancer can be identified through several methods such as mammography, X-ray, ultrasound (US), Positron Emission Tomography (PET), Computed Tomography, temperature measurement and Magnetic Resonance Imaging (MRI) [2,7,8]. Usually, the golden standard approach for breast cancer diagnosis is a pathological process. In order to maximize visibility, the extracted tissue is stained in the lab before being subjected to imaging analysis. The staining procedure frequently employs Hematoxylin and Eosin (H&E) [9]. In most cases, Histopathological image analysis and genomics can both be utilized to identify breast cancer [4,10]. A histopathological image is a

microscopic picture of breast tissues and is very helpful in the early treatments of cancer [11]. The genomics field is primarily concerned with multi-scale connections between data on gene expression and medical imaging [4]. A more accurate diagnosis can be achieved with the use of radio-genomics [10]. In order to predict and identify cancer early, molecular analyses of tissues can be performed.

Recently, Computer-Aided Design (CAD) has been introduced [12,13] to simplify breast cancer identification. However, traditional computer-aided design systems generally depend on manually created features and therefore weaken the overall performance [13]. With the advent of machine learning and AI methods, deep learning-based techniques were recently studied for breast cancer detection [14,15]. Representation learning is the basis of deep learning techniques, which come from several layers. The representation is transformed from the lower to the higher levels at each end by combining the non-linear and simple modules, in which the lower-level features are more comprehensible and the higher-level features more abstract [14]. Compared to the ML methods [10,16], deep learning is more effective and requires fewer human interventions for the related pattern recognition schemes. This makes it capable of effectively solving complex problems in various areas such as image analysis [17], pattern recognition and natural language processing.

Following its successful applications in breast cancer detection, the number of research works on deep learning-based approaches has increased exponentially recently. This begs for a systematic review and summary of the existing works to help successive researchers and practitioners gain better insight into and understanding of the field. In the past, several literature reviews on breast cancer detection have been published. For instance, Yassin et al. [2] investigated the existing ML-based methods for breast cancer diagnosis. The authors comprehensively assessed different image modalities as well as the different ML-based classifiers used for breast cancer detection. The authors in [18] reviewed various deep learning-based methods for classifying breast cancer based on image processing. However, this work focuses on shallow feed-forward networks, while other deep learning-based methods were not emphasized. A study in [19] summarized recent studies that used deep learning (DL) methods to detect breast cancer disease based on different imaging approaches. The authors specifically focused on the three breast cancer imaging approaches, namely, MRI, mammography and ultrasound.

The authors in [20] examined several DL- and traditional ML-based methods for breast cancer prediction by reviewing a total of 8 papers and 27 papers in DL and ML, respectively. The authors discovered that most of the reviewed literature employed the imaging process; however, only a few of the reviewed articles applied genetics. The authors in [21] reviewed several imaging methods based on mammography for breast cancer diagnosis. Gupta et al. [22] presented a brief survey of different systems and methods for the early detection of breast cancer. In this study, various imaging methods which comprise radar-based imaging and microwave tomography were examined. Oyelade et al. [23] examined different deep learning-based methods for breast cancer diagnosis from digital mammography. Husaini et al. [15] examined the application of ML techniques and thermography for detecting breast cancer problems. In this method, various ML methods were investigated to process the breast cancer thermographic images.

Considering the above-mentioned several reviews on the deep learning-based methods for breast cancer detection, it could be seen that most of the existing review works particularly focus on the image-based methods for breast cancer detection problems. It can be seen that most of the existing studies emphasize the traditional ML-based approaches, while those focused on the deep learning-based techniques particularly covered very limited studies, with no clear comprehensive and systematic analysis of the existing approaches.

Therefore, this SLR aims to provide a comprehensive SLR to systematically analyze the existing literature in the area of deep learning techniques for breast cancer detection and to have more in-depth knowledge that is required for the early detection of breast cancer and proper treatments. The SLR is mainly focused on methods based on histopathological images and genomics. To provide more recent developments in breast cancer diagnosis,

we consider studies conducted from 2010 to 2021. We also investigate the challenges and provide recommendations for future research to help researchers and practitioners in this area. The main contributions of this study are as follows:

- Summarizing the deep learning-based methods popularly applied for breast cancer detection.
- Identifying the deep learning-based methods with the best performances for breast cancer diagnosis.
- Investigating the datasets generally applied in the deep learning-based methods for breast cancer diagnosis.
- Summarizing the evaluation metrics used for breast cancer diagnosis using deep learning-based methods.
- Analyzing the research gaps and future direction for deep learning-based breast cancer detection.

The rest of the article is designed as follows: Section 2 describes the background of the deep learning-based methods. Sections 3–5 present the research methods, the research questions and the synthesis results. Finally, Sections 6 and 7 summarize the limitations of the study and the conclusion of the paper.

## 2. Background of Deep Learning Methods

Deep learning, in simple terms, is referred to as a machine learning method that employs learning representation to automatically determine feature representations from input data [11,24]. Unlike traditional learning (such as support vector machine (SVM), K-nearest neighbors (KNN), random forest (RF), etc.), deep learning does not need a human-engineered feature to optimally perform [10,12]. Several deep learning methods have been introduced in the past decades [25,26], which include the convolutional neural network (CNN), Restricted Boltzmann Machine (RBM), Recurrent Neural Network (RNN), Deep Autoencoder (AE), multi-layer perceptron and Generative Adversarial Network (GAN). These models have been applied and proved to be successful in several areas, including natural language processing [27], recommender systems [28], computer vision [29], medical imaging [30], etc. Brief explanations of these models are given in the following paragraphs:

**RBMs:** RBMs are generative deep learning-based methods that use blocks in a greedy, layer-by-layer form of network training and feature learning [31]. For the model to generate unbiased estimates of maximum likelihood learning, contrastive divergence (CD) is used during training. One of the popular examples of the RBM is the Deep Belief Network (DBN) [32]. By stacking many RBMs to data, the Deep Belief Network deep learning approach is learned in a greedy-wise, layer-by-layer style [33,34]. The top layer of the Deep Belief Network contains an undirected connection, which models the observed distribution between the hidden layer and vector space layer [13]. The lower layers of the Deep Belief Networks have direct connections. Similar to this, weight fine-tuning during training is carried out layer by layer via Contrastive Divergence (CD) [26]. Figure 1 illustrates the structure of the DBN technique.

**Autoencoder (AE):** The autoencoder technique replicates the copy of the input values as the output through the use of encoder and decoding units [35]. By reducing the dimensions of the data, autoencoder algorithms obtain the most discriminative features from unlabeled data [32]. In order to reduce error rates, the encoder converts the input data into hidden features that are further reconstituted using the decoder [36]. The method provided learning feature extraction techniques for avoiding handcrafted feature issues [35]. To generate a lower dimensional discriminative feature, the training of the auto-encoder is performed in such a way that the hidden layers are smaller than the inputs/outputs. [36]. A typical example of the autoencoder model is illustrated in Figure 2.

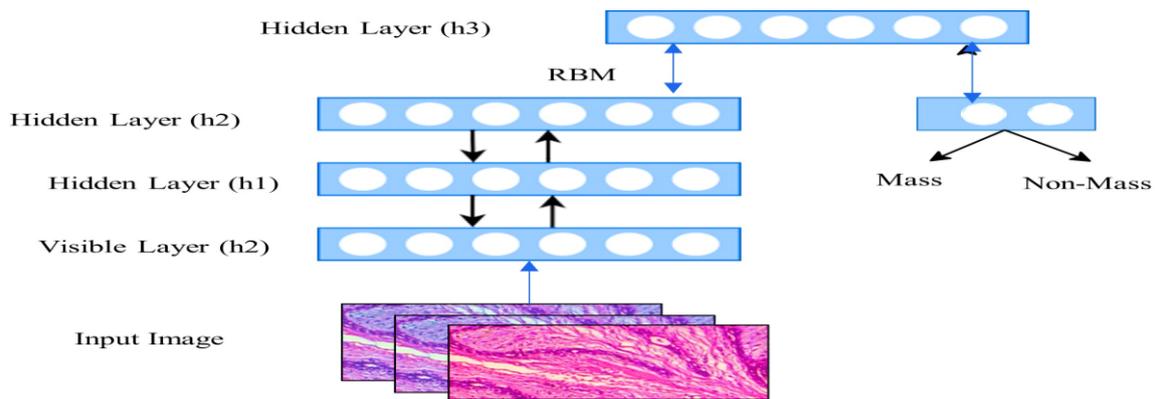


Figure 1. An illustration of the DBN for breast cancer detection.

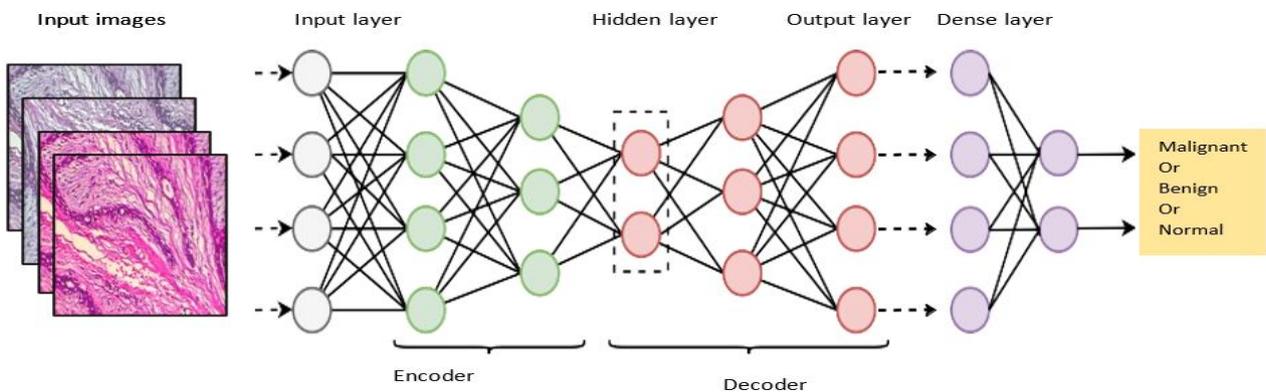


Figure 2. An illustration of the deep autoencoder.

CNN: CNN is a neural network that has an interconnected structure. A CNN method is one of the popular deep learning methods that form convolutional operations on raw data [37]. It has been applied in various applications such as speech recognition, sentence modeling, image classification and, recently, medical imaging, including a breast cancer diagnosis. Basically, three layers make up the CNN: a convolutional layer, a pooling layer and a fully connected layer. These layers are stacked to create a deep architecture for automatically extracting the features [37]. Recently, several of the CNN models have been introduced by different researchers: VGG [38], AlexNet [39] and GoogleNet [40]. Figure 3 illustrates the structure of the CNN technique.

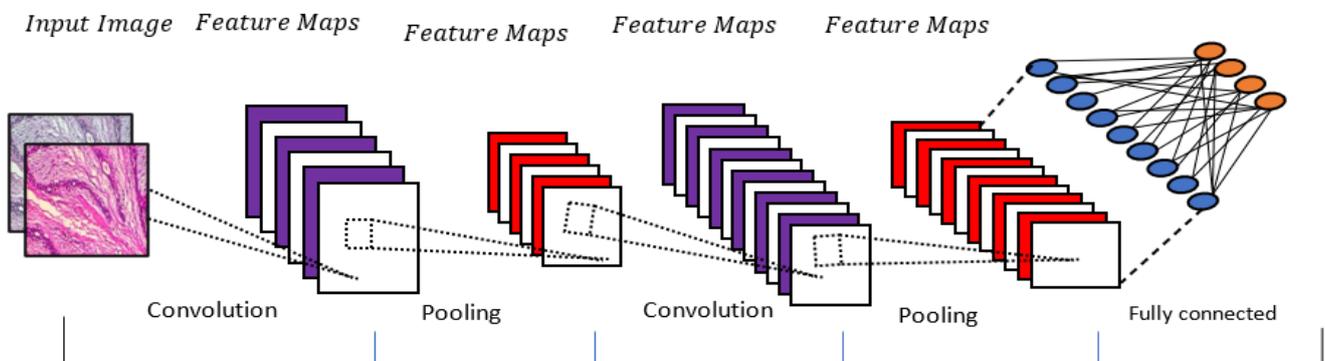


Figure 3. An illustration of the CNN model.

RNN: RNN is a supervised deep learning-based technique used for sequential data [41]. The hidden unit of the recurrent cell is used by RNN to learn complex changes by integrating a temporal layer for capturing sequential information [5]. Depending on the

information accessible to the network, which is automatically updated to represent the network’s current state, the hidden unit cells may change [41]. However, the model suffers from vanishing or exploding gradients and is difficult to train, which limits its application for modeling temporal dependencies and longer sequences in a dataset [42]. To mitigate the issue of the vanishing and exploding gradients of RNN, new models such as Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) were introduced [36]. As a way to manage the flow of information into the network, LSTM included memory cells to store relevant data [34]. The LSTM can represent temporal dependencies and effectively capture the global features in sequential data to improve the speed [42]. However, one of the problems of the LSTM is the issue of the numerous parameters that are required to be updated in the course of the model training [34]. To reduce the parameter update, the GRU with fewer parameters, which makes it faster and less complicated, is introduced. The next hidden state’s updating process and contents exposure technique are different for LSTM and GRU [25]. While GRU updates the subsequent hidden states based on the correlation, subject to the amount of time required to maintain such information in the memory, the LSTM updates the hidden states using the summing operation. The basic structure of the RNN model is shown in Figure 4.

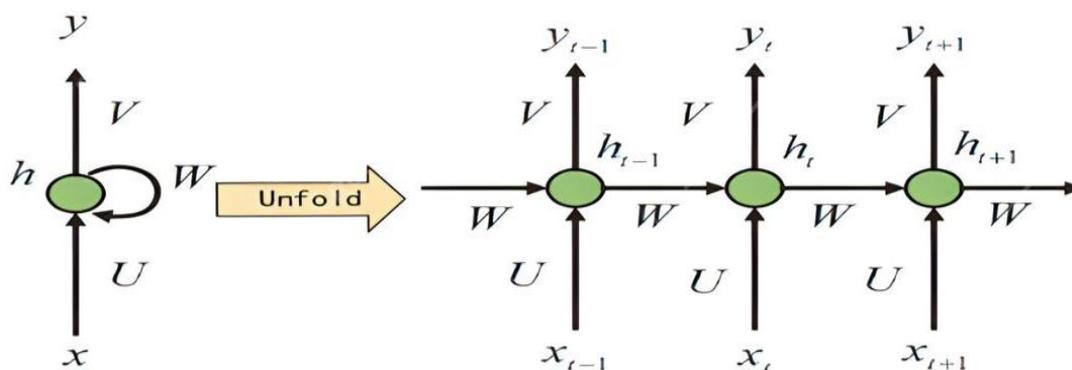


Figure 4. An illustration of the RNN model.

Multilayer Perceptron (MLP): The MLP technique for feed-forward neural networks employs additional layers and nonlinear activation functions [43]. The simplest deep learning architecture is considered to be the multilayer perceptron [25]. It consists of a minimum of one hidden layer that is coupled in a feed-forward manner. For the most part, deep learning architectures are basically built using this framework [43]. The model’s linear approach can be changed into nonlinear models for neural performance using MLP. Consequently, they have been applied in different applications including natural language processing, a recommendation system and medical imaging. Figure 5 illustrates the structure of the MLP.

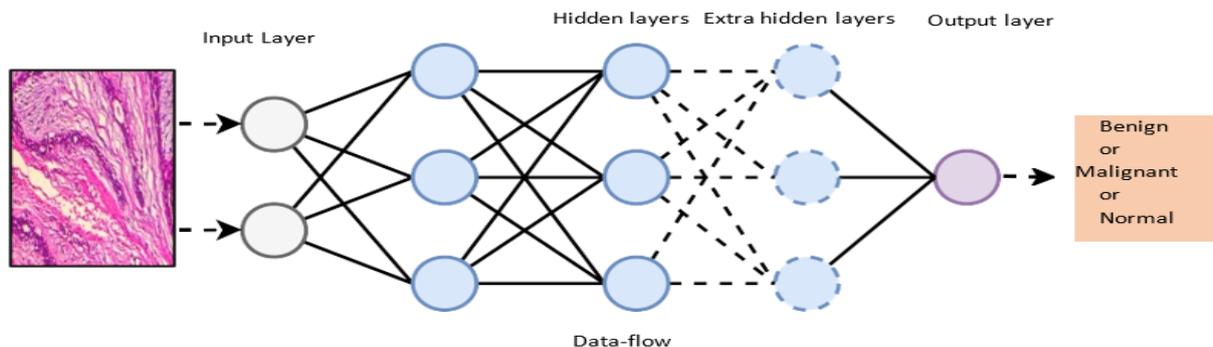


Figure 5. An illustration of the MLP.

Generative Adversarial Network (GAN): One recently introduced deep learning technique is the GAN, which was first discussed in [44]. For the maximum likelihood estimation technique, GAN offers an alternate approach. In a zero-sum game involving two neural networks, it uses both supervised and unsupervised learning techniques. It particularly aims to train a generative method that seeks to infer the distribution of the target data from the training data. Additionally, it makes use of the discriminative model, which provides an approximation of the likelihood that a sample of data is drawn from actual training data as opposed to the output. The learning rates and other parameters, such as the structure of the model, have a significant impact on how the GAN is trained. Numerous ad hoc methods are frequently needed to achieve effective convergence so as to increase the fidelity of the data generated. In addition, several extensions of the GAN methods were also devised to ease the complexity and improve the training process convergence. The Wasserstein Generative Adversarial Network and Loss Sensitive Generative Adversarial Network (LSGAN) are examples of this [45]. Research on the GAN is still in its early stages, though. Studies published recently suggested that GAN might be used for supervised learning applications. In the cases of recommender systems and information retrieval applications, using GAN’s unsupervised learning capabilities seems exciting [46]. An illustration of the GAN model is given in Figure 6.

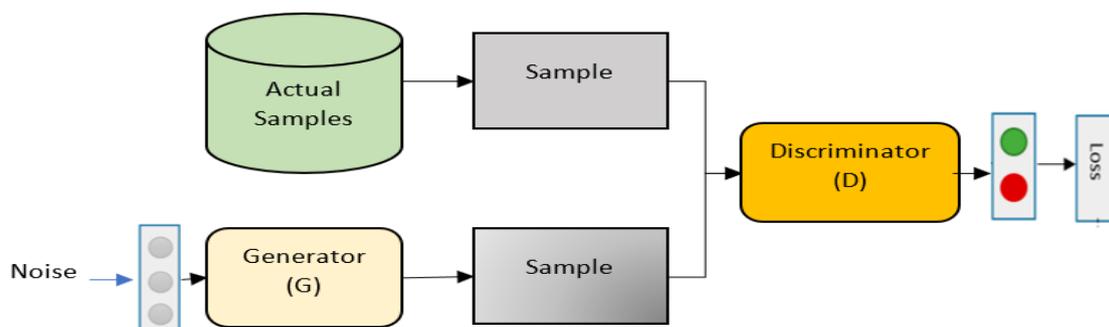


Figure 6. An illustration of the GAN.

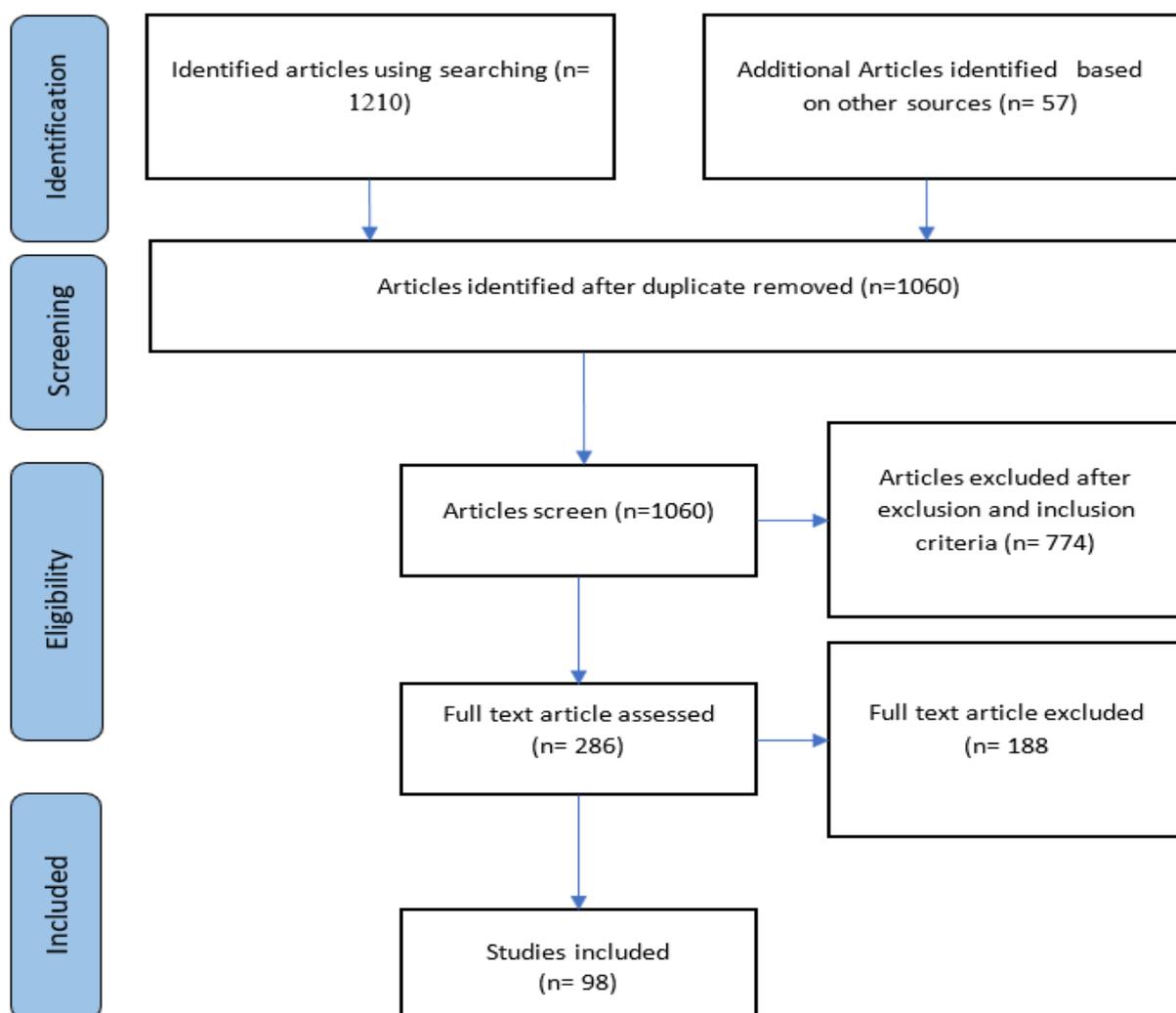
The various deep learning techniques discussed above have made state-of-the-art methods in breast cancer detection. The main benefit of the deep learning-based method is its ability to automatically learn from unlabeled raw data. Such methods offer diverse capabilities for the detection of breast cancer. Table 1 summarizes the descriptions, weaknesses and strengths of each deep learning-based technique.

Table 1. Comparison of the deep learning techniques.

Techniques	Description	Weakness	Strength
DBN	Comprises two top-layer undirected connections and a directed connection at the lower layer.	The lengthy parameter initialization method makes it hard to train.	Able to extract hidden features from multiple data with great power.
DBM	Deep neural network with undirected connections between layers.	Due to the extensive parameters adjustment, the computation took a long time.	Enhances feature extraction using unsupervised training.
MLP	Models the data which have a simple correlation.	Delayed convergence and high complexity.	Nonlinear transformation.
AE	A neural network that has been trained to reconstruct the inputs under certain restrictions.	Lower ability to scale to high-dimensional data. Relies on high parameter tuning and numerical optimization.	Using an unsupervised learning strategy, it can learn more complicated feature representation.
CNN	Interconnected structural design motivated by the biological visual cortex.	Requires high parameterization tuning.	Highly effective for feature extraction with contextual data.
RBM	Undirected, bipartite graph with visible and hidden layers.	Although not tractable, the contrastive divergence can be utilized to learn the parameters.	Suitable for simple representation learning.
GAN	Deep neural network with a generator and discriminator.	Difficulty in convergence. Unstable learning process.	Appropriate for both supervised and unsupervised learning.

### 3. Methods

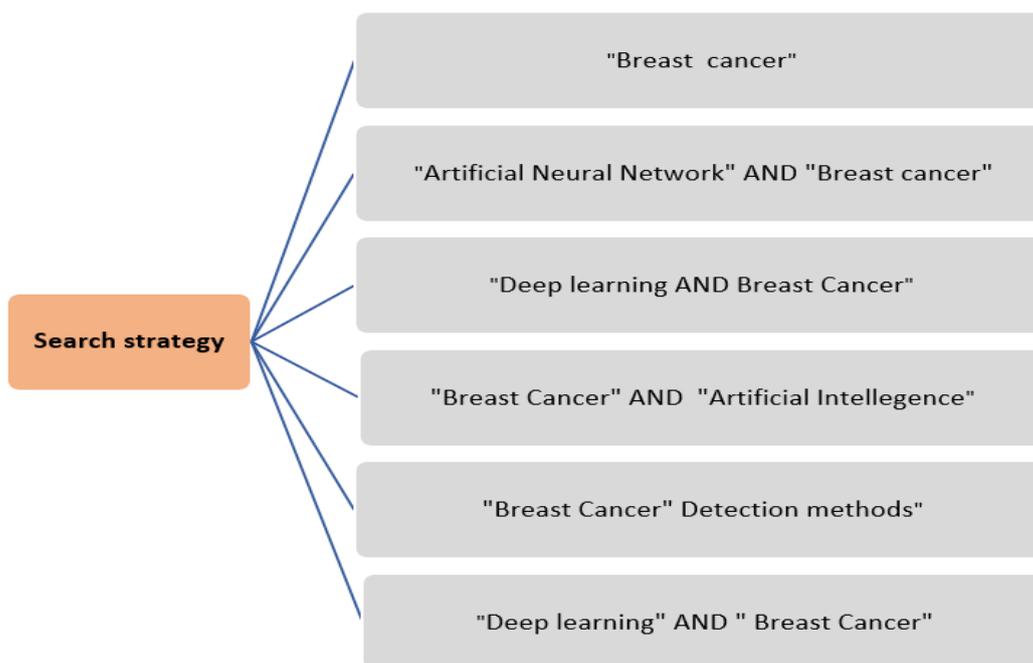
This study employs the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guiding principles for conducting systematic reviews [47]. PRISMA provides a replicable and consistent method for identifying, selecting and critically examining the existing studies. It also offers direction for choosing, recognizing and evaluating the studies. Figure 7 displays the PRISMA process of the SLR. Details of the review process are given in the following subsection.



**Figure 7.** PRISMA method of the SLR.

#### 3.1. Data Sources and Search Strategy

To gather more relevant studies, eight different bibliographic databases were considered for conducting the search process. The searched digital libraries include Scopus, Google Scholar, IEEE Xplore Library, Web of Science, SpringerLink, ScienceDirect, ACM Digital Library and PubMed. To obtain the latest and most comprehensive reviews, 2010 to 2021 was considered the timespan for the review. The search string comprised “breast cancer” or “cancer” or “((Artificial Neural Network)” or “((deep learning) AND (breast cancer))” AND (breast cancer))” or “((Artificial Intelligence) AND (breast cancer) AND (detection techniques))” or “breast cancer diagnosis”). Figure 8 demonstrates the search strings used in the searching process.



**Figure 8.** The search string used in the search strategy.

3.2. Selection Criteria

In addition to the search string applied in the automatic search, we also conducted a manual search to comprehensively identify the relevant articles. In the identification stage, a total of 1267 studies were obtained from both automatic and manual searches. In the screening stage, 1060 published papers were chosen after filtering duplicate, unsuitable and irrelevant papers. A total of 774 articles were discarded after the removal criteria were applied to the remaining ones, and 286 articles were obtained. Finally, 98 research articles were designated for review after the quality assessment of the selected articles. Exclusively, papers utilizing genetic expression and imaging were included, and we narrowed down the number of publications by concentrating on journal and conference papers only. In addition to several kinds of gene expression and gene sequencing, the imaging modalities we took into consideration included ultrasound, magnetic resonance imaging (MRI), mammography and radiography. In this study, we concentrated on publications that employ deep learning-based approaches to implement the detection of breast cancer, as well as the publications that focused on breast cancer detection using both image and gene data. The exclusion and inclusion principles for the systematic reviews are stated in Table 2.

**Table 2.** Inclusion and exclusion.

Inclusion	Exclusion
Studies involving experimental results only	Studies without experimental results were not considered
Studies published from 2010 to 2021	Studies published before 2010 were excluded
Studies involving breast cancer detection only	Studies involving other cancer detections
Papers focusing on deep learning-based breast cancer detection	Papers focusing on other techniques used for breast cancer diagnosis
Studies written in English only	Studies written in other languages
Only journals and conferences are used	Other sources such as books, theses and magazines were excluded

3.3. Quality Assessment

Analyzing the data contained in an SLR and evaluating the quality of the evidence it contains is equally significant. A bias resulting from a methodology may affect the results

of a poorly conducted research work and therefore requires careful interpretation. Such articles must be explicitly excluded or at least identified as such in the systematic review. It is also necessary to select the correct criteria for assessing the quality of the evidence and any inherent biases in each study. To ensure the quality of the selected publications, the standard quality checklist questions (SCQ) designed in [48] are applied. To this end, following [49], we choose the articles that replied “yes” to at least seven questions. To ensure the findings significantly contribute to the review, the quality assessment will be taken into account alongside the data extraction [50]. Table 3 shows the SCQ used in the SLR.

**Table 3.** Quality Checklist.

No	Quality Question
SCQ1	Is the report clear and coherent?
SCQ2	Is the aim of the research clearly specified?
SCQ3	Is the data collection method obviously described?
SCQ4	Have the diversity contexts been well explored?
SCQ5	Are the findings of the study reliable?
SCQ6	Are there links between the data, interpretation and conclusion?
SCQ7	Are the methodology and experimentation process clear?
SCQ8	Are the research procedures documented adequately?
SCQ9	Are they important, if credible?
SCQ10	Could the research findings be replicated?

### 3.4. Data Extraction and Synthesis

Firs, we took note of the key details, including the title of the paper, its publication year, an author list and the publisher. Afterward, we added some data for running the SLR, such as the deep learning method used, the reported accuracy and the evaluation metrics. The data synthesis stage particularly examines the associated findings from the data extraction process, which can be used to answer the research questions. After gathering the data, we visualize and analyzed the data through different visualization techniques and tools, such as a bar chart, a pie chart, histograms, etc.

## 4. Research Questions

Choosing RQs is very important in defining the overall purpose and expected outcomes of a study. Therefore, to achieve the main purpose of our SLR, we design the following RQs:

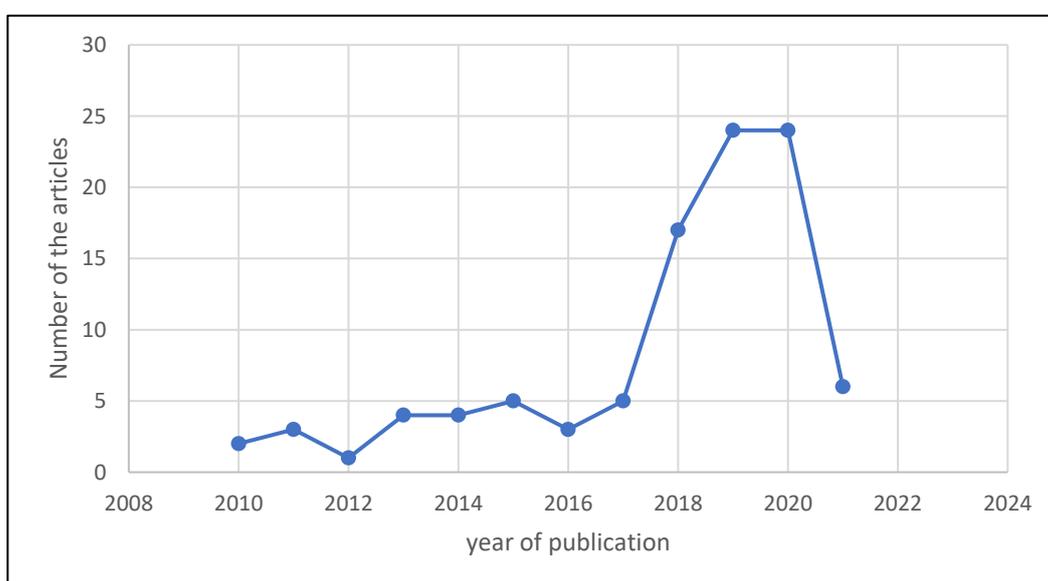
- RQ1: What are the most common deep learning-based methods applied for breast cancer detection?
- RQ2: What is the most effective deep learning-based method for breast cancer detection in terms of performance?
- RQ3: What are the commonly used performance evaluation metrics for deep learning-based breast cancer detection?
- RQ4: What are the common datasets used for deep learning-based breast cancer detection?
- RQ5: What are the challenges and future directions of the deep learning-based methods for cancer detection?

## 5. Results and Metanalysis

This section presents the metanalysis of the searched results of our SLR. It begins by presenting descriptions of the selected articles in this SLR and subsequently answers each of the RQs specified.

### 5.1. Description of the Selected Studies

The number of publications regarding deep learning-based breast cancer detection methods from 2010 until 2021 is illustrated in Figure 9, which shows a chronological summary of the studies that were selected for this SLR. The illustration on the graph reveals a growing trend of studies in this field over recent years, especially from 2013 to 2015, when the number of published papers started to show a significant increase. The number of articles considered for the study mostly increased after 2013. Specifically, the years 2019 and 2020 saw the highest number of published articles (24) within a study year, followed by 16 papers in 2018 and 4 papers in 2021. As only 1 paper was examined in the year 2010, it can be seen that there was a decreased rate of relevant publications in that year. Figure 9 depicts the different journals and the total number of relevant articles used in the review. Similarly, the different names of the journals and the number of their corresponding papers used in the study are shown in Figure 10. It should be noted that, due to the space limit, only 70% of the reviewed articles and their corresponding journal names are shown in Figure 10.



**Figure 9.** The number of papers with the corresponding years of publication.

### 5.2. RQ1: What Are the Common Deep Learning Methods for Breast Cancer?

To better provide effective cancer detection, a system needs to process 200 to 300 cells per frame, which is impossible through manual tracking [51]. Therefore, the development of effective technologies for breast cancer detection becomes necessary. In contrast, deep learning can be utilized to find patterns in unprocessed data. In recent times, deep learning is a common tool used to detect breast cancer. Deep learning methods have been demonstrated to be capable of diagnosing breast cancer up to 12 months earlier than those using conventional clinical procedures [16]. In addition, the techniques can be used to learn the most pertinent features to best tackle the issue. In recent times, different deep learning-based methods have been introduced for breast cancer diagnosis, which include CNN-, DNN-, RNN-, DBN- and AE-based approaches.

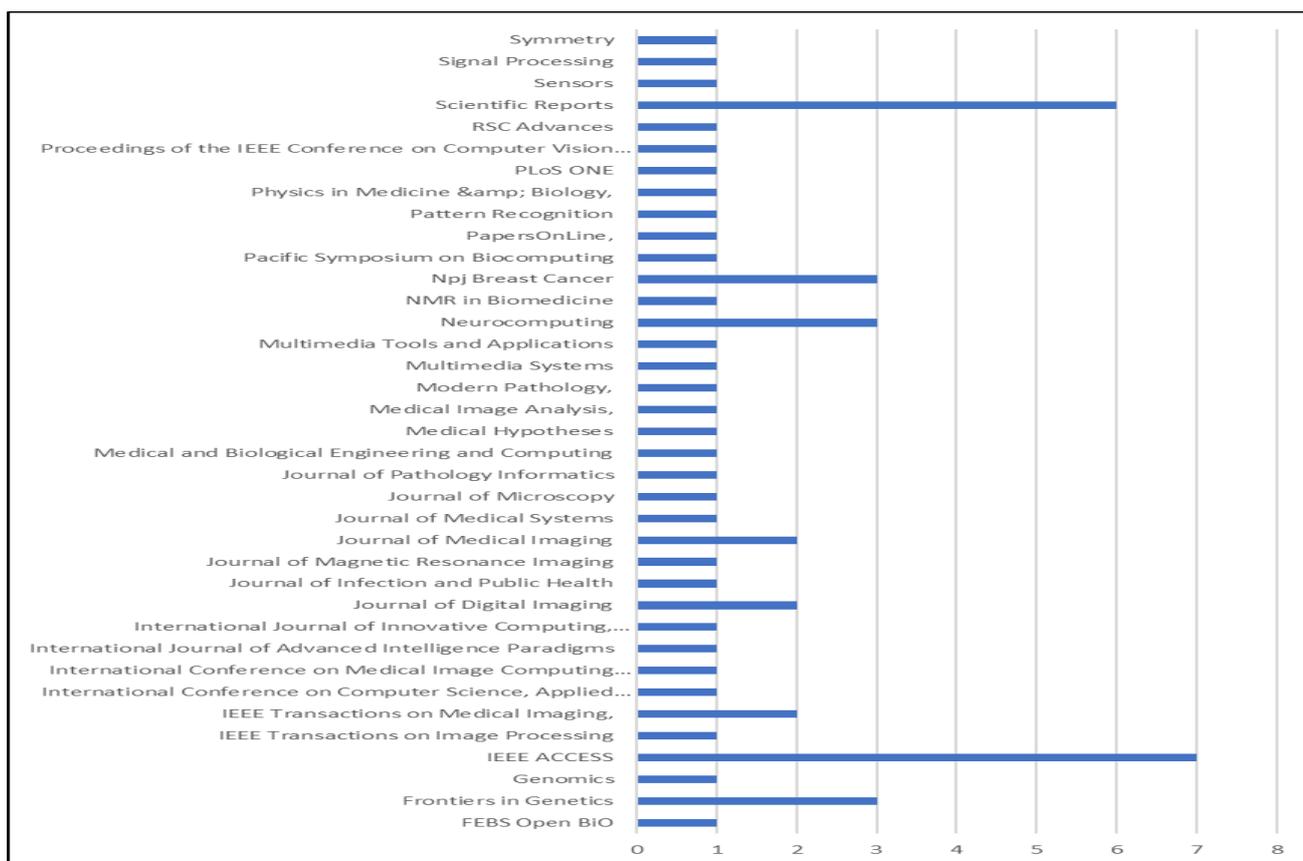


Figure 10. Number of articles per journal.

CNN is the most popular deep-learning technique that has been utilized in several studies for breast cancer detection [16]. The CNN is a deep learning model which represents hierarchical abstraction and consists of various layers which accept features as raw data [24]. The CNNs used for breast cancer diagnosis can be grouped into two: the transfer learning-based model and the de novo trained model [19]. The kinds of CNN models that were generated and trained from the scratch are referred to as “de novo models”. In contrast, the CNN techniques using previously trained neural networks such as AlexNet, residual neural network (ResNet), visual geometry group (VGG), etc. are called “transfer learning (TL)-based methods” [16]. Several methods used CNN-based methods for breast cancer diagnosis. The studies basically used the CNN model to extract various features based on the validated gene expression data in order to detect clinical outcomes in breast cancer [52,53]. Some authors use CNN to detect the mitosis process for the invasive breast cancer diagnosis based on histopathological imaging [54]. Others used the deep CNN method to classify and identify tumor-related stroma for breast cancer diagnosis [55–58]. In [59], a CNN-based method was utilized in combination with linear discriminative analysis and ridge regression based on image processing for breast cancer detection.

DNN has been shown to be effective for breast cancer detection [25,35,41]. The layers that make up a DNN include an output layer, a convolution layer, a fully connected layer and a pooling layer. Among these layers, a convolution layer is used for learning high-level characteristics. The purpose of a fully connected layer is to learn pixel-level features. The size of the convolved features can be decreased by a pooling layer, which lowers the amount of computation needed. This layer is capable of performing average pooling and maximum pooling operations [25]. Several DNN-based methods have been introduced for breast cancer detection. For example, Che et al. [60] used an attentive-based model for breast cancer diagnosis. The authors utilize multi-NNF DNN based on multi-modality information to improve the performances of breast cancer detection and prognosis.

Lu et al. [61] applied a DNN technique for cancer detection. The authors used the deep learning-based feature representation of tumor-infiltrating lymphocytes in the cancer of the breast based on histopathological imaging for improving the detection process. Several other studies used the DNN model for detecting the subtypes of cancer by combining several types of transcriptomics data on breast cancer using deep learning techniques [62] and Identifying Differentially Expressed (DE) Biomarkers [63].

Another important deep learning-based method used for breast cancer detection is the RNN, which comprises many versions such as LSTM and GRU. RNN is a supervised deep learning method that is specifically used to process sequential data because it uses the loops and memories to keep track of previous computations when processing sequential inputs. Therefore, processing 3D volumetric images specifically as MRI image slices has been reported to be very helpful. Recently, several methods have applied the LSTM and GRU [64] for breast cancer detection. Some methods used the RNN technique for breast cancer detection. For example, a study in [64] introduced a gene-subcategory deep learning-based method that uses interaction-based learning for feature representations to improve the breast cancer subcategorical analysis based on gene expressions.

Autoencoders (AE) have also been applied for breast cancer detection [34]. A decoder rebuilds the images using the learned features to capture the essence of the raw features, and AE employs an encoder to properly transfer each image into a latent space. Several studies utilize AE for breast cancer detection. For example, Zhang et al. [65] used AE for improving deep learning for breast cancer detection. The authors employed Integrates Feature Selection and Feature Extraction to Predict the Clinical Outcome of Breast Cancer. In another study, Toğaçar, et al. [59] used the autoencoder-processed invasive breast cancer images to combine the CNN with linear discriminant analysis and ridge regression for breast cancer diagnosis. Xu et al. [66] introduced an approach based on histopathological images using a stacked sparse autoencoder (SSAE) to improve the model performances for breast cancer detection. A similar approach was introduced by Xu et al. [67], which used the SSAE model consisting of two SAEs for the detection of nuclei patches on breast cancer histopathology images to enhance breast cancer diagnosis.

GAN, which is a deep-learning-based generative model, has also been utilized for breast cancer detection. Shams et al. [68] designed a deep generative multi-tasking based on the combination of the GAN and CNN for reducing the death rate in breast cancer. The model uses strategies for achieving better accuracy in mammography diagnoses. Singh et al. [69] designed a GAN-based method for segmenting breast tumors inside the region of interest on a mammogram. The generative model creates a binary mask that defines the tumor region after learning to detect it. It motivates the generative network to produce binary masks that are as realistic as feasible. Additionally, GAN has been applied as image-augmentation methods recently to address the issue of limited data. Digital breast tomosynthesis data were used by the authors of [70] to detect anomalies and complete an image using GAN. The detection method described in this study produced encouraging results because it could locate suspicious areas without the need for training photos with anomalies. Fan et al. [70] used a generative adversarial technique with an improved deep network, bicubic interpolation and other techniques to create super-resolution images. GAN was employed by Guan and Loew [71] as a new mammographic image generator from the DDSM datasets, while CNN was utilized as GAN's discriminator. Compared to other image-augmentation methods, GAN performed better.

DBN comprises a stack of RBMs, which are seen as the visible layer, and multiple hidden layers (the top two hidden layers contain random relationships), from which the deep features of the visible layer are extracted using a model of the generative probability union distribution [34,38]. Several studies used DBM for breast cancer detection; for example, Smolander et al. [33] introduced a deep learning-based approach for classifying gene expression data from complicated diseases by comparing DBN with SVM. Salma used a metaheuristic approach based on the DBN for breast cancer classification. Table 4 illustrates the deep learning-based methods used for breast cancer detection.

**Table 4.** The summary of the deep learning methods reviewed in this SLR.

Model	Description	References	Frequency
CNN	It uses convolution and pooling operations for feature extraction.	[52–59,61,72–85]	23
DNN	It is trained to construct posterior distributions for all possible values, using the input distribution’s encoding as a model distribution.	[60–63,86–89]	8
DBN	Includes directed connections at the lower layers and undirected connections at the two top layers.	[33,90]	2
RNN	An RNN version that incorporates a memory block to overcome the backpropagation problem.	[64]	1
MLP	A deep feed-forward network that produces a collection of outputs from a set of inputs.	[64,77,84,91–100]	13
AE	The generative model that recreates the input data in the output layer.	[59,65–67,101,102]	6
GAN	Deep learning technique that is generatively semi-supervised.	[68–71,103–105]	8

5.3. RQ2: Which Deep Learning Models Perform Most Effectively?

This section summarizes the performances of the different deep learning-based methods for cancer detection identified in this SLR. Table 5 demonstrates the summarized performances of different methods. Essentially, deep learning-based methods used two different approaches to classify breast cancer: multiclass and binary classification, which involve two instances and multiple subtypes, respectively. From the table, it can be seen that the use of the multiclass classification generally leads to lower performances compared to the binary classification. From Table 4, it can also be noticed that the majority of the used binary classifications performed better in most of the methods. In Table 5, the best accuracy for multiclass categorization was indicated to be 95.7% accuracy [52]. Regarding the task of classifying breast cancer subtypes, the author of the research essentially compared deep learning with machine learning. There have been numerous models employed, including CNN, DNN with the attention mechanism and many other DL techniques.

**Table 5.** Models used for gene sequencing data in selected papers.

DL Used	Brief Description of the Method	Classification	Method	Accuracy	Reference
FFNN	Uses negative and positive classes for the breast cancer diagnosis	Binary	Genomic	95%	[87]
FFNN	Uses negative and positive classes for the breast cancer detection	Binary	Genomic	92%	[106]
CNN	Provides breast cancer subtype classification based on the feature extraction	Multiclass	Genomic	95.6%	[52]
DNN	Uses mitotic and non-mitotic processes for identifying the breast cancer	Binary	Genomic	87%	[60]
DNN	Identifies risk categories based on the four classes	Multiclass	Genomic	94%	[86]
MLP	Axillary prediction of the lymph node status in breast cancer	Multiclass	Genomic	84%	[91]
FFNN	Breast cancer detection based on the cancer subtypes	Binary	Genomic	98.3%	[107]
CNN	Breast cancer detection based on the presence or absence of a tumor	Binary	Genomic	96.7%	[53]

Table 5. Cont.

DL Used	Brief Description of the Method	Classification	Method	Accuracy	Reference
CNN	Subtype identification based on the seven cancer types	Multiclass	Genomic	84.7%	[72]
AE	Predicts clinical outcomes of breast cancer	Binary	Genomic	84%	[65]
CNN	Breast cancer detection based on multiple categories of cancer	Multiclass	Genomic	76.4%	[73]
CNN	Mitotic and non-mitotic	Binary	Genomic	73.6%	[54]
CNN	Identification and classification of tumor-associated stroma in diagnostic breast biopsies.	Binary	Genomic	92%	[55]
CNN	Epithelial and stromal	Binary	Genomic	88%	[56]
DNN	Breast cancer molecular subtype classification	Multiclass	Genomic	87%	[108]
MLP	It uses the feature selection method for the detection	Binary	Genomic	98%	[109]
PNN	Breast cancer early diagnosis	Binary	Genomic	96%	[110]
MLP	Breast cancer prognosis detection	Binary	Genomic	96%	[93]
CNN	Breast cancer survival according to different features	Multiclass	Genomic	90%	[74]
CNN	Uses the feature selection method for the detection.	Binary	Genomic	98.7%	[59]
CNN	Breast cancer image classification based on epithelial and non-epithelial	Binary	Genomic	94%	[75]
DNN	Uses nuclei probability for the breast cancer tumor detection	Binary	Genomic	93%	[88]
MLP	Uses malignant/benign classification for the detection	Binary	Genomic	96.5%	[96]
CNN	Uses malignant/benign classification to determine the breast cancer	Binary	Genomic	94.5%	[76]
CNN	It determines the low risk and high risk for breast cancer	Binary	Genomic	98%	[58]
MLP	Uses multiclass based on the malignant, benign and other subclasses	Binary	Genomic	98%	[77]
MLP	Uses malignant/benign classification to determine the breast cancer	Binary	Genomic	97%	[97]
MLP	It uses the clustering method to identify breast cancer tumors	Binary	Genomic	96%	[98]
DBN	It uses multi-categories for the breast cancer gene classification	Binary	Genomic	90%	[33]
DNN	It uses the clustering method to identify breast cancer tumors	Binary	Genomic	84%	[89]
MLP	Identifies risk categories based on the four classes	Binary	Genomic	95%	[89]
FNN	Uses malignant/benign classification to determine the breast cancer	Binary	Genomic	94%	[90]
CNN	It identifies breast cancer lymph nodes	Binary	Genomic	84%	[79]
MLP	It identifies breast cancer tumors	Multiclass	Genomic	95%	[100]
CNN	It identifies breast cancer tumors	Binary	Genomic	93%	[80]
RNN	It identifies breast cancer tumors	Multiclass	Genomic	82%	[64]
CNN	Uses feature selection for the breast cancer detection	Multiclass	Genomic	94.5%	[81]
CNN	Uses feature selection for the breast cancer detection	Binary	Genomic	95.6%	[52]
DNN	It uses negative/positive classification to detect breast cancer tumors	Multiclass	Genomic	92%	[61]
CNN	Uses multiple categories based on A Luminal B HER2	Multiclass	Imaging	70%	[82]
CNN	Predicts breast tumors and responses to chemotherapy	Multiclass	Imaging	88%	[57]
CNN	Uses negative and positive to diagnose the breast cancer	Binary	Imaging	97%	[83]

Table 5. Cont.

DL Used	Brief Description of the Method	Classification	Method	Accuracy	Reference
MLP	Breast cancer subtype classification based on different categories	Multiclass	Imaging	90%	[84]
CNN	Uses negative and positive to diagnose the breast cancer	Binary	Imaging	71%	[85]
DBN	Uses the unsupervised method	Binary	Imaging	98%	[111]
GAN	Uses the generative unsupervised method for the breast cancer detection	Multiclass	Imaging	80%	[105]

It should be noted that, because not all of the published works in the reviewed articles adopted the confusion matrix parameters, only the accuracy metrics were considered for analysis in the reviewed articles. Based on the results from Table 5, CNN-based techniques seem to be most popular for both multiclass and binary classification. Generally, it can be seen that a high performance is obtained when imaging data are used for breast cancer based on binary classification [83]. However, when it comes to subtypes classification (multiple classifications), genetic sequencing exhibits better results than imaging data. It can also be observed that, generally, the use of CNN results in outstanding performances for both gene expression and imaging data. An example of this is paper [83], which produced 99% accuracy for binary classification. The models used hybrid models that combine deep learning methods and machine learning as well as standalone deep learning models.

In conclusion, the studies made extensive use of different algorithms. However, in both gene sequence and image data, MLP and CNN were the most often utilized algorithms. The majority of the articles employed CNN and MLP with different parameters and properties; however, several other algorithms, including DNN, were also used.

#### 5.4. RQ3: What Are the Evaluation Measures Commonly Used for Deep Learning-Based Breast Cancer Detection?

The evaluation metrics are essential for evaluating the models' performances. However, deep learning methods are not rigorously evaluated using any specific set of metrics. Specifically, many performance evaluation metrics, which include recall, precision, F1-measure, accuracy, area under the curve (AUC), false-negative rate (FNR), etc., have recently been used by various researchers. We outline the evaluation metrics used in the reviewed papers based on the chosen publications in this section. The different performance measure formulas are displayed in Table 6.

The precision assesses the accuracy of the model's positive predictions, whereas the accuracy indicates how many of the model's overall predictions were correct. The recall of a classifier, sometimes referred to as sensitivity, is the proportion of positive cases that the classifier correctly detects. The ratio of successfully identified negative samples to all negative samples serves as a proxy for the classifier's specificity. Although it makes sense to improve both precision and recall, there is an adverse relationship between the two metrics. By enforcing a higher precision, a lower recall may occur, and vice versa. This is often known as a trade-off for recall/precision measures. The harmonic mean of the recall and precision, which is referred to as the F-score, is a preferable metric to maximize. By adjusting the decision threshold value for a classifier, it is conceivable to observe the performance change in terms of the trade-off between particular metrics such as recall and precision. Parametric evaluation is the procedure of examining all confusion matrices that may emerge from altering the decision threshold.

**Table 6.** Performance evaluation measures.

Metrics	Description	Formula	References	Frequency
Accuracy	This is calculated by dividing the percentage of accurate predictions by the total number of predictions made by the model.	$\frac{(TN+TP)}{(TN+FN+FP+TP)}$	[33,52–54,56,57,59–62,65,72–77,79,81,82,87–91,93,95,97,98,100,106–110,112–116]	42
Precision	This is calculated by dividing the actual positive results by the true positive results.	$\frac{TP}{(TP+FP)}$	[54,59–62,72,73,75–77,79–81,88–90,96,97,100,109,114]	21
Recall/Sensibility	This is calculated as the ratio of actual positive samples that should have been discovered to true positive results.	$\frac{TP}{(TP+FN)}$	[33,54,59–62,72,75,79–81,89–91,96,97,100,109]	18
F1-score	This is measured as the model's accuracy in each class.	$2 \times \frac{Recall \times Precision}{(Recall + Precision)}$	[33,53,54,59,61,62,65,72,74,75,77,78,80,81,88–91,96,97,100,109,117]	23
Specificity (TNR)	This is how well the prediction model predicts the percentage of negative tuples.	$1 - FP$	[53,57,59,61,65,72,75,78,80,81,96,97,116,117]	14
AUC-ROC	It indicates the trade-off between the false positive rate (FPr) and the true positive rate (TPr).	Area under the ROC curve	[57,60,75,79,81,84,90,97,108,113,114,116,117]	13

The other metrics comprise precision and recall and the ROC curves. Classifiers can be compared based on the area under the ROC curve (AUC), with perfect models having an AUC value equal to 1 and fully random classifiers having an AUC equal to 0.5. The performance measure formulas, the references and the number of studies are shown in Table 5 for each metric. For the FPR and FNR, the lower the value of the performance metrics, the better the ability of models to generalize. The TP, FP, TN and FN in this table represent the true positive, false positive, true negative and false negative, respectively.

#### 5.5. RQ4: What Datasets Are Available for Breast Cancer Diagnosis?

Generally, deep learning methods require a large amount of data for training the model and achieving improved performance. Consequently, a major barrier to using deep learning algorithms for medical diagnosis is the lack of data. Recently, several datasets have been published for breast cancer diagnosis. These datasets can either be private or public. The private datasets can be acquired through different academic-comprising universities across the world [118]. There were hardly any freely accessible public datasets. The accessible datasets for both imaging and gene sequencing are included in Tables 7 and 8 under public and private. We discovered different public and private databases that contained gene expression data for healthy and sick individuals, even though gene expression data are less prevalent than imaging data.

Genome Atlas [119] includes the most popular datasets used for breast cancer analysis; it aims to detect the complete set of DNA changes in various kinds of cancer-related problems. It offers a large number of instances for researchers. Each participant's clinical data are provided, along with some generic data. Understanding these modifications could aid the study into how various cancer types develop [57]. For a variety of cancer types, including breast cancer, the Genome Atlas dataset contains gene data. The METABRIC dataset, which comprises the clinical characteristics, SNP genotypes, CNV profiles and expression derived through breast cancers acquired from participants in the METABRIC study, is the second most widely used dataset [3]. Additionally, for the purpose of studying and researching cancer, NCI Genomic Data Commons (GDC) [120] offers scholars a vast

collection of gene-related data. Genome and DNA sequencing information from studies on cancer detection can be found in the GEO database [118].

Compared to genetic data, imaging data are more readily available as datasets. The Wisconsin breast cancer (WBC) dataset, available from the UCI repository, is the most used imaging dataset [121]. It has features that were determined from a digital image of a breast mass which was sampled with a fine needle aspiration (FNA). A total of 11 attributes are listed for each of the 699 cases in this dataset. Clumps' thickness, uniform cell size and shapes, marginal adhesion, the size of a single epithelial cell, the absence of mitoses, naked nuclei, bland chromatin, normal nucleoli and marginal adhesion is the independent properties.

Another popular breast cancer dataset for imaging is the DDSM dataset [122], which can be used on its own. CBIS-DDSM, an enhanced and condensed version of the DDSM dataset for the evaluation of breast cancer diagnosis, was just published by Clark et al. [123]. It includes improved ROI-segmented images in addition to a readily usable dataset. The dataset contains 891 mass cases and 753 microcalcification instances, respectively.

Another widely used dataset is the Mammographic Image Analysis Society Digital Mammogram Database (MIAS) [124]. The MIAS database contains 322 digitized MLO images and 161 cases with a variety of findings, including benign and malignant tumors as well as regular images. Another well-liked dataset for the detection of breast cancer is IN breast, which contains mammography pictures from cases involving screening, diagnosis and follow-up. The work's most distinctive aspect is the expertly annotated ground truth that has been carefully associated with it. Tables 7 and 8 summarize the datasets, and Figure 11 visualizes the distribution of the datasets for the cancer detection.

**Table 7.** Genomic Datasets used for breast cancer diagnosis.

Datasets	Description	Links	Accessibility	Type	Instances	Reference
The Cancer Genome Atlas	It offers clinical data for every participant, along with some general data.	<a href="http://cancergenome.nih.gov">http://cancergenome.nih.gov</a>	Public	Genomic	11,429	[86]
METABRIC	Clinical characteristics, expression and SNP genotypes derived from breast cancers are included in the databases.	<a href="https://ega-archive.org/datasets/EGAD0010000268">https://ega-archive.org/datasets/EGAD0010000268</a>	Public	Genomic	543	[3]
Array Express Database	A database for high-throughput functional genomics data.	NA	Private	Genomic	NA	[125]
Geo Database	Data from high-throughput functional genomics investigations are stored in the functional genomics collection.	<a href="https://www.ncbi.nlm.nih.gov/geo/info/download.html">https://www.ncbi.nlm.nih.gov/geo/info/download.html</a>	Public	Genomic	404	[118]
STRING and BIOGRID	A proteomic database involving the networks and interactions of proteins in a wide array of species.	NA	Private	Genomic	NA	[114]
GDC	It offers many gene-related data to researchers for use in the study and analysis of cancer.	<a href="https://gdc.cancer.gov/">https://gdc.cancer.gov/</a>	Public	Genomic	9114	[120]
Spark Dataset	A public dataset which uses gene sequence data for breast cancer diagnosis.	<a href="https://drive.google.com/file/d/1yd1gwk2o">https://drive.google.com/file/d/1yd1gwk2o</a>	Public	Genomic	106	[112]

**Table 8.** Imaging-based datasets used for breast cancer diagnosis.

Datasets	Description	Links	Accessibility	Type	Instances	Reference
WBCD	It comprises 699 records from the FNA of human breast tissue. Each record has nine attributes.	<a href="https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+">https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+</a>	Public	Imaging	569	[121]
Helsinki University	A customized private dataset designed by the Cancer Institute and Helsinki University.	NA	Private	Imaging	NA	[56]
MRI Data	The datasets are collected provisionally under an HIPAA-compliant approval from the institutional board.	<a href="https://wiki.cancerimagingarchive.net/display/Public/RIDER+Breast+MRI#2251275749b786f1af5747c39abd8eda0d12e2b7">https://wiki.cancerimagingarchive.net/display/Public/RIDER+Breast+MRI#2251275749b786f1af5747c39abd8eda0d12e2b7</a>	Public	Imaging	1500	[6]
University of Vermont Medical Center	A customized private dataset designed by the University of Vermont Medical Center.	NA	Private	Imaging	NA	[9]
DDSM	It is a combination of digitalized benign without callback, benign, normal and cancer volumes that were carefully chosen.	<a href="https://wiki.cancerimagingarchive.net/display/Public/CBIS-DDSM#2251662">https://wiki.cancerimagingarchive.net/display/Public/CBIS-DDSM#2251662</a>	Public	Imaging	10,239	[122]
Stanford Tissue Microarray Database	A database that offers tools for tissue microarrays, design, image scoring and annotation to researchers.	NA	Private	Imaging	-NA	[126]
MIAS	It is available on a 2.3 GB 8 mm (Exabyte) tape and has 322 digitized films.	<a href="http://peipa.essex.ac.uk/benchmark/databases/index.html">http://peipa.essex.ac.uk/benchmark/databases/index.html</a>	Public	Imaging	322	[124]
IN breast	It consists of images captured between 2008 and 2010 at the Breast Center in CHSJ, Porto.	<a href="http://medicalresearch.inescporto.pt/breastresearch/index.php/Get_INbreast_Database">http://medicalresearch.inescporto.pt/breastresearch/index.php/Get_INbreast_Database</a>	Public	Imaging	410	[127]
CBIS-DDSM	It provides better ROI-segmented images in addition to a readily useable dataset.	<a href="https://wiki.cancerimagingarchive.net/display/Public/CBIS-DDSM">https://wiki.cancerimagingarchive.net/display/Public/CBIS-DDSM</a>	Public	Imaging	1644	[123]

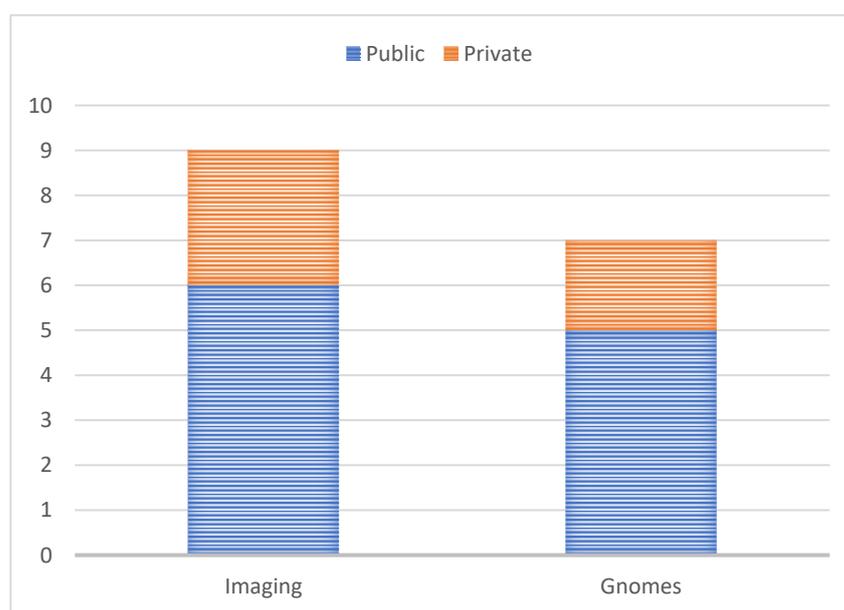
### 5.6. RQ5: What Are the Research Gaps, Challenges and Future Directions?

This section aims to answer RQ5 by presenting the research gaps and the future direction of deep learning-based methods for breast cancer detection. Although the existing studies have substantially contributed some basis for the deep learning-based methods for breast cancer detection, there are some gaps and future research directions in the field. The future research directions in this area and some of the important problems that have yet to be resolved are outlined as follows:

#### 5.6.1. Balanced Dataset

Deep learning methods have been demonstrated to be effective and promising in data mining. However, the evaluated datasets are virtually imbalanced for the deep learning-based method in the breast cancer area. First, there are not enough publicly

available imaging and genomes databases that comprise the pathological heterogeneity and coexisting benign malignancy in different populations. Additionally, private datasets are arbitrary in terms of size, number and format. The work of annotation, which can be labor- and time-intensive, is another major difficulty for which clinical radiologists are not always available. An uneven breast cancer diagnosis is caused by an imbalanced training network dataset. In addition to creating adequate public breast cancer databases, future work will also focus on creating private datasets, which will primarily address issues with insufficient medical images, complex annotations and gene sequencing data. Moreover, to solve the problem of imbalanced data, some works applied oversampling methods [128,129]. However, it was shown in [130] that the oversampling method can lead to choice-based sample biases. Additionally, SMOTE (Synthetic Minority Oversampling Technique) was the only oversampling technique that was used [131]. This led to the conclusion that future studies might take into account using both under- and oversampling methods.



**Figure 11.** Distribution of the datasets for the cancer detection.

### 5.6.2. Interpretable Deep Architecture

Designing deep models that are appropriate for the medical data would be challenging in addition to the complex radiological images. Hence, deep learning networks that are difficult to interpret pose a common problem in the detection of breast cancer. Along with diagnostic and histological reports, the ability of deep learning to handle heterogeneous datasets increases the possibilities for interpretable DL in breast cancer. Additionally, image captioning, which combines CV and NLP, has made significant progress [132]. The paradigm basically uses a standard encoder–decoder architecture to provide descriptions of given images based on RNN using visual attributes derived by CNN. The automatic creation of medical imaging reports appears in the medical domain, inspired by picture captioning [133]. This hybrid deep learning architecture can more accurately understand breast cancer diagnostic data because it uses text features to express radiomic features [134]. The first step is to pay attention to the learning model’s attributes and weights on the breast cancer images. The second is to allow for the understanding of the factors impacting the learning process involved within images of breast cancer.

### 5.6.3. Clinical Application

For the management and treatment of patients, medical decisions are crucial. There is a need to put more emphasis on the deep learning-based method’s value in real-world applications in order to increase the accuracy of medical reports. Despite their best efforts,

most studies did not fully utilize the relevant data that were available, despite the fact that they aimed to enhance the ultimate performance. There is a need to take breast cancer's characteristics into account when using various imaging techniques (such as detailed malignant lesion features on ultrasound, hypoechogenicity, angular margin, posterior shadowing and internal vascularity). Additionally, there are several risk factors for breast cancer, such as age, family history and genetics, that are often taken into account by doctors but are overlooked in deep learning-based models. It can be another direction we can explore.

As a result, there is a need to pay closer attention to the deep features obtained using DLR, which might be used as information or data for additional research. In addition, multitask learning is a promising future direction. Through many tasks that gather the shared learning features one at a time, radionics features in a deep neural framework can reduce overfitting.

## 6. Limitations of the Study

In this SLR, various deep learning-based methods for breast cancer detection are identified. By developing our protocols, we aim to maximize both internal and external validity while addressing the RQs. There are still some restrictions and challenges to the validity of this argument, which can be encountered and explained in this section.

- This SLR is solely restricted to journal and conference materials that discuss breast cancer detection in DL. Several irrelevant research publications were found and eliminated from this review in the early stages of the study by using our search method. This guarantees that the chosen research papers met the requirements for the investigation. However, it is believed that incorporating other sources—such as extra sourcebooks, for instance—would have improved this review.
- We limited our search to items written in English. Due to the possibility of related publications in this area of study existing in other languages, this leads to linguistic bias. Thankfully, all of the papers gathered for this study were written in English. We are not language-biased as a result.
- Although the primary databases were considered when looking through the study articles, it is possible that other digital libraries with pertinent studies were disregarded. To overcome this limitation, we compared the search phrases and keywords to a well-known collection of research studies. However, when looking for the keywords, certain synonyms might be missed. To solve this issue, the SLR protocol has been updated to make sure no crucial phrases are omitted.

## 7. Conclusions

In this research work, we presented the SLR to systematically review the existing research works on breast cancer diagnosis based on the deep learning-based methods involving genetic sequencing or the histopathological imaging process. The SLR specifically adopts the PRISMA approach. Particularly, the applicability of various techniques in deep learning-based methods for breast cancer detection is studied. Several studies were searched and gathered, and after the eligibility screening and quality evaluation, 95 articles were identified. The results of the systematic review showed that the Convolutional Neural Network (CNN) is the most accurate and widely used model for breast cancer detection, and the accuracy metrics are the most popular method used for performance evaluation. Moreover, the datasets used for breast cancer diagnosis and the performance of different algorithms are also explored. Finally, the challenges and future research directions on breast cancer detection using deep learning techniques are also investigated to help researchers and practitioners acquire in-depth knowledge of and insight into the area. The widespread application of the CNN algorithms to data on both MRI images and gene expressions is a significant breakthrough. Comparing these models with other algorithms, they often produce positive results. It could be interesting for researchers to carry out additional research and apply more hybrid algorithms with CNN.

In addition, it was found that most of the studies on classifying images have not properly utilized the attention mechanism. This has provided researchers with the opportunity to employ attention mechanisms in the future to increase the precision of the deep learning methods. Given that this field is broad and that there is always the potential for more research and findings, researchers have recently been focusing on gene sequence data. Future researchers will have various chances to make a further contribution by combining different gene sequencing datasets to predict more results with larger datasets. More studies should emphasize deriving important aspects from gene expression data to improve outcomes and improve accuracy by employing confusion matrix parameters. This opens up the prospect for future studies to address a range of related issues, including determining risk levels and predicting the likelihood of recurrence. Future studies could focus on leveraging genetic data to create multiclass predictors, for example. Breast cancer detection and survival likelihood were the main focuses of the majority of studies, which only used genetic sequencing data with binary categorization. Furthermore, large-scale, thorough and fully labeled WSI datasets are currently lacking. Consequently, the creation of sizable public databases is crucial for future research.

**Author Contributions:** Writing—original draft, conceptualization, methodology, formal analysis: M.N.; writing—review and editing, supervision, project administration, funding acquisition: U.K.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Research Creativity and Management Office (RCMO), and the School of Computer Sciences at the Universiti Sains Malaysia (USM).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** The authors would like to thank Fatima Alsharjabi for proofreading and editing the article.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Aggarwal, R.; Sounderajah, V.; Martin, G.; Ting, D.S.W.; Karthikesalingam, A.; King, D.; Ashrafian, H.; Darzi, A. Diagnostic accuracy of deep learning in medical imaging: A systematic review and meta-analysis. *Npj Digit. Med.* **2021**, *4*, 23. [[CrossRef](#)]
2. Yassin, N.I.; Omran, S.; Houbay, E.M.; Allam, H. Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review. *Comput. Methods Programs Biomed.* **2018**, *156*, 25–45. [[CrossRef](#)]
3. El-Nabawy, A.; El-Bendary, N.; Belal, N.A. A feature-fusion framework of clinical, genomics, and histopathological data for METABRIC breast cancer subtype classification. *Appl. Soft Comput.* **2020**, *91*, 20. [[CrossRef](#)]
4. Nassif, A.B.; Abu Talib, M.; Nasir, Q.; Afadar, Y.; Elgendy, O. Breast cancer detection using artificial intelligence techniques: A systematic literature review. *Artif. Intell. Med.* **2022**, *127*, 13. [[CrossRef](#)]
5. Yao, H.; Zhang, X.; Zhou, X.; Liu, S. Parallel Structure Deep Neural Network Using CNN and RNN with an Attention Mechanism for Breast Cancer Histology Image Classification. *Cancers* **2019**, *11*, 1901. [[CrossRef](#)]
6. Ha, R.; Chang, P.; Mutasa, S.; Karcich, J.; Goodman, S.; Blum, E.; Kalinsky, K.; Ms, M.Z.L.; Jambawalikar, S. Convolutional Neural Network Using a Breast MRI Tumor Dataset Can Predict Oncotype Dx Recurrence Score. *J. Magn. Reson. Imaging* **2019**, *49*, 518–524. [[CrossRef](#)]
7. Bhatt, C.; Kumar, I.; Vijayakumar, V.; Singh, K.U.; Kumar, A. The state of the art of deep learning models in medical science and their challenges. *Multimed. Syst.* **2021**, *27*, 599–613. [[CrossRef](#)]
8. Fujioka, T.; Mori, M.; Kubota, K.; Oyama, J.; Yamaga, E.; Yashima, Y.; Katsuta, L.; Nomura, K.; Nara, M.; Oda, G.; et al. The Utility of Deep Learning in Breast Ultrasonic Imaging: A Review. *Diagnostics* **2020**, *10*, 1055. [[CrossRef](#)]
9. Mullooly, M.; Bejnordi, B.E.; Pfeiffer, R.M.; Fan, S.; Palakal, M.; Hada, M.; Vacek, P.M.; Weaver, D.L.; Shepherd, J.A.; Fan, B.; et al. Application of convolutional neural networks to breast biopsies to delineate tissue correlates of mammographic breast density. *Npj Breast Cancer* **2019**, *5*, 11. [[CrossRef](#)]
10. Saba, T. Recent advancement in cancer detection using machine learning: Systematic survey of decades, comparisons and challenges. *J. Infect. Public Health* **2020**, *13*, 1274–1289. [[CrossRef](#)]
11. Shah, S.M.; Khan, R.A.; Arif, S.; Sajid, U. Artificial intelligence for breast cancer analysis: Trends & directions. *Comput. Biol. Med.* **2022**, *142*, 15. [[CrossRef](#)]

12. Houssein, E.H.; Emam, M.M.; Ali, A.A.; Suganthan, P.N. Deep and machine learning techniques for medical imaging-based breast cancer: A comprehensive review. *Expert Syst. Appl.* **2021**, *167*, 20. [[CrossRef](#)]
13. Bharati, S.; Podder, P.; Mondal, M.R.H. Artificial neural network based breast cancer screening: A comprehensive review. *Int. J. Comput. Inf. Syst. Ind. Manag. Appl.* **2020**, *12*, 125–137.
14. Bai, J.; Posner, R.; Wang, T.; Yang, C.; Nabavi, S. Applying deep learning in digital breast tomosynthesis for automatic breast cancer detection: A review. *Med. Image Anal.* **2021**, *71*, 19. [[CrossRef](#)]
15. Rautela, K.; Kumar, D.; Kumar, V. A Systematic Review on Breast Cancer Detection Using Deep Learning Techniques. *Arch. Comput. Methods Eng.* **2022**, *31*, 4599–4629. [[CrossRef](#)]
16. Nemade, V.; Pathak, S.; Dubey, A.K. A Systematic Literature Review of Breast Cancer Diagnosis Using Machine Intelligence Techniques. *Arch. Comput. Methods Eng.* **2022**, *29*, 4401–4430. [[CrossRef](#)]
17. Huang, X.W.; Qian, S.S.; Fang, Q.; Sang, J.T.; Xu, C.S. CSAN: Contextual Self-Attention Network for User Sequential Recommendation. In Proceedings of the 26th ACM Multimedia Conference (MM), Seoul, Republic of Korea, 22–26 October 2018; pp. 447–455.
18. Zhou, X.; Li, C.; Rahaman, M.; Yao, Y.; Ai, S.; Sun, C.; Wang, Q.; Zhang, Y.; Li, M.; Li, X.; et al. A Comprehensive Review for Breast Histopathology Image Analysis Using Classical and Deep Neural Networks. *IEEE Access* **2020**, *8*, 90931–90956. [[CrossRef](#)]
19. Pang, T.; Wong, J.H.D.; Ng, W.L.; Chan, C.S. Deep learning radiomics in breast cancer with different modalities: Overview and future. *Expert Syst. Appl.* **2020**, *158*, 113501. [[CrossRef](#)]
20. Fatima, N.; Liu, L.; Hong, S.; Ahmed, H. Prediction of Breast Cancer, Comparative Review of Machine Learning Techniques, and Their Analysis. *IEEE Access* **2020**, *8*, 150360–150376. [[CrossRef](#)]
21. Lu, Y.; Li, J.Y.; Su, Y.T.; Liu, A.A. A Review of Breast Cancer Detection in Medical Images. In Proceedings of the 33rd IEEE International Conference on Visual Communications and Image Processing (IEEE VCIP), Taichung, Taiwan, 9–12 December 2018.
22. Gupta, N.P.; Malik, P.K.; Ram, B.S. A Review on Methods and Systems for Early Breast Cancer Detection. In Proceedings of the International Conference on Computation, Automation and Knowledge Management, ICCAKM 2020, Dubai, United Arab Emirates, 9–10 January 2020; pp. 42–46. [[CrossRef](#)]
23. Oyelade, O.N.; Ezugwu, A.E.-S. A State-of-the-Art Survey on Deep Learning Methods for Detection of Architectural Distortion From Digital Mammography. *IEEE Access* **2020**, *8*, 148644–148676. [[CrossRef](#)]
24. Roslidar, R.; Rahman, A.; Muharar, R.; Syahputra, M.R.; Arnia, F.; Syukri, M.; Pradhan, B.; Munadi, K. A Review on Recent Progress in Thermal Imaging and Deep Learning Approaches for Breast Cancer Detection. *IEEE Access* **2020**, *8*, 116176–116194. [[CrossRef](#)]
25. Schmidhuber, J. Deep Learning in Neural Networks: An Overview. *Neural Netw.* **2015**, *61*, 85–117. [[CrossRef](#)]
26. Hatcher, W.G.; Yu, W. A Survey of Deep Learning: Platforms, Applications and Emerging Research Trends. *IEEE Access* **2018**, *6*, 24411–24432. [[CrossRef](#)]
27. Da'U, A.; Salim, N. Aspect extraction on user textual reviews using multi-channel convolutional neural network. *PeerJ Comput. Sci.* **2019**, *16*, e191. [[CrossRef](#)]
28. Da'U, A.; Salim, N. Sentiment-Aware Deep Recommender System with Neural Attention Networks. *IEEE Access* **2019**, *7*, 45472–45484. [[CrossRef](#)]
29. Deng, L. Deep Learning: Methods and Applications. *Found. Trends Signal Process.* **2013**, *7*, 197–387. [[CrossRef](#)]
30. Xue, P.; Wang, J.; Qin, D.; Yan, H.; Qu, Y.; Seery, S.; Jiang, Y.; Qiao, Y. Deep learning in image-based breast and cervical cancer detection: A systematic review and meta-analysis. *Npj Digit. Med.* **2022**, *5*, 19. [[CrossRef](#)]
31. Shibata, H.; Takama, Y.; Ieee. Behavior Analysis of RBM for Estimating Latent Factor Vectors from Rating Matrix. In Proceedings of the 6th International Conference on Informatics, Electronics and Vision (ICIEV)/7th International Symposium in Computational Medical and Health Technology (ISCMHT), Univ Hyogo, Himeji Engn, Himeji, Japan, 1–3 September 2017.
32. Deng, L. Three classes of deep learning architectures and their applications: A tutorial survey. *PSIPA Trans. Signal Inf. Process.* **2012**, *57*, 58.
33. Smolander, J.; Dehmer, M.; Emmert-Streib, F. Comparing deep belief networks with support vector machines for classifying gene expression data from complex disorders. *FEBS Open Bio* **2019**, *9*, 1232–1248. [[CrossRef](#)]
34. Salas, J.; Vidal, F.D.B.; Martinez-Trinidad, F. Deep Learning: Current State. *IEEE Lat. Am. Trans.* **2019**, *17*, 1925–1945. [[CrossRef](#)]
35. Pouyanfar, S.; Sadiq, S.; Yan, Y.; Tian, H.; Tao, Y.; Reyes, M.P.; Shyu, M.-L.; Chen, S.-C.; Iyengar, S.S. A Survey on Deep Learning: Algorithms, Techniques, and Applications. *ACM Comput. Surv.* **2019**, *51*, 36. [[CrossRef](#)]
36. Hossain, Z.; Sohel, F.; Shiratuddin, M.F.; Laga, H. A Comprehensive Survey of Deep Learning for Image Captioning. *ACM Comput. Surv.* **2019**, *51*, 36. [[CrossRef](#)]
37. Alom, M.Z.; Taha, T.M.; Yakopcic, C.; Westberg, S.; Sidike, P.; Nasrin, M.S.; Hasan, M.; Van Essen, B.C.; Awwal, A.A.S.; Asari, V.K. A State-of-the-Art Survey on Deep Learning Theory and Architectures. *Electronics* **2019**, *8*, 292. [[CrossRef](#)]
38. Yu, X.; Zhou, Q.; Wang, S.; Zhang, Y. A systematic survey of deep learning in breast cancer. *Int. J. Intell. Syst.* **2022**, *37*, 152–216. [[CrossRef](#)]
39. Zhao, Y.; Hu, B.; Wang, Y.; Yin, X.; Jiang, Y.; Zhu, X. Identification of gastric cancer with convolutional neural networks: A systematic review. *Multimed. Tools Appl.* **2022**, *81*, 11717–11736. [[CrossRef](#)]
40. Binder, A.; Bach, S.; Montavon, G.; Müller, K.-R.; Samek, W. Layer-Wise Relevance Propagation for Deep Neural Network Architectures. In *Lecture Notes in Electrical Engineering*; Springer: Singapore, 2016; Volume 376, pp. 913–922. [[CrossRef](#)]

41. Mambou, S.J.; Maresova, P.; Krejcar, O.; Selamat, A.; Kuca, K. Breast Cancer Detection Using Infrared Thermal Imaging and a Deep Learning Model. *Sensors* **2018**, *18*, 2799. [\[CrossRef\]](#)
42. Egger, J.; Gsaxner, C.; Pepe, A.; Pomykala, K.L.; Jonske, F.; Kurz, M.; Li, J.; Kleesiek, J. Medical deep learning—A systematic meta-review. *Comput. Methods Programs Biomed.* **2022**, *221*, 22. [\[CrossRef\]](#)
43. Miotto, R.; Wang, F.; Wang, S.; Jiang, X.; Dudley, J.T. Deep learning for healthcare: Review, opportunities and challenges. *Brief. Bioinform.* **2018**, *19*, 1236–1246. [\[CrossRef\]](#)
44. Shreir, L.L. (Ed.) Front Matter. In *Corrosion*, 2nd ed.; Newnes: London, UK, 1976; p. iii.
45. Bentur, Y.; Layish, I.; Krivoy, A.; Berkovitch, M.; Rotman, E.; Bar Haim, S.; Yehezkelli, Y.; Kozler, E. Civilian Adult Self Injections of Atropine—Trimedoxime (TMB4) Auto-Injectors. *Clin. Toxicol.* **2006**, *44*, 301–306. [\[CrossRef\]](#)
46. Wang, Q.Y.; Yin, H.Z.; Hu, Z.T.; Lian, D.F.; Wang, H.; Huang, Z. Neural Memory Streaming Recommender Networks with Adversarial Training. In Proceedings of the 24th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD), London, UK, 19–23 August 2018; pp. 2467–2475.
47. Page, M.J.; McKenzie, J.E.; Bossuyt, P.M.; Boutron, I.; Hoffmann, T.C.; Mulrow, C.D.; Shamseer, L.; Tetzlaff, J.M.; Akl, E.A.; Brennan, S.E.; et al. The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews. *BMJ Br. Med. J.* **2021**, *372*, n71. [\[CrossRef\]](#)
48. Keele, S. *Guidelines for Performing Systematic Literature Reviews in Software Engineering*, Technical Report, version 2.3 ESBE Technical Report; ESBE; Keele University: Keele, UK, 2007.
49. Genc-Nayebi, N.; Abran, A. A systematic literature review: Opinion mining studies from mobile app store user reviews. *J. Syst. Softw.* **2017**, *125*, 207–219. [\[CrossRef\]](#)
50. Kitchenham, B.; Brereton, O.P.; Budgen, D.; Turner, M.; Bailey, J.; Linkman, S. Systematic literature reviews in software engineering—A systematic literature review. *Inf. Softw. Technol.* **2009**, *51*, 7–15. [\[CrossRef\]](#)
51. Mridha, M.F.; Hamid, A.; Monowar, M.M.; Keya, A.J.; Ohi, A.Q.; Islam, R.; Kim, J.-M. A Comprehensive Survey on Deep-Learning-Based Breast Cancer Diagnosis. *Cancers* **2021**, *13*, 6116. [\[CrossRef\]](#)
52. Mostavi, M.; Chiu, Y.-C.; Huang, Y.; Chen, Y. Convolutional neural network models for cancer type prediction based on gene expression. *BMC Med. Genom.* **2020**, *13*, 44. [\[CrossRef\]](#)
53. Lancashire, L.J.; Powe, D.G.; Reis-Filho, J.S.; Rakha, E.; Lemetre, C.; Weigelt, B.; Abdel-Fatah, T.M.; Green, A.; Mukta, R.; Blamey, R.; et al. A validated gene expression profile for detecting clinical outcome in breast cancer using artificial neural networks. *Breast Cancer Res. Treat.* **2010**, *120*, 83–93. [\[CrossRef\]](#)
54. Paul, A.; Mukherjee, D.P. Mitosis Detection for Invasive Breast Cancer Grading in Histopathological Images. *IEEE Trans. Image Process.* **2015**, *24*, 4041–4054. [\[CrossRef\]](#)
55. Bejnordi, B.E.; Mullooly, M.; Pfeiffer, R.M.; Fan, S.; Vacek, P.M.; Weaver, D.L.; Herschorn, S.; Brinton, L.A.; Van Ginneken, B.; Karssemeijer, N.; et al. Using deep convolutional neural networks to identify and classify tumor-associated stroma in diagnostic breast biopsies. *Mod. Pathol.* **2018**, *31*, 1502–1512. [\[CrossRef\]](#)
56. Simidjievski, N.; Bodnar, C.; Tariq, I.; Scherer, P.; Andres-Terre, H.; Shams, Z.; Jamnik, M.; Liò, P. Variational Autoencoders for Cancer Data Integration: Design Principles and Computational Practice. *Front. Genet.* **2019**, *10*, 14. [\[CrossRef\]](#)
57. Ha, R.; Chin, C.; Karcich, J.; Liu, M.Z.; Chang, P.; Mutasa, S.; Pascual Van Sant, E.; Wynn, R.T.; Connolly, E.; Jambawalikar, S. Prior to initiation of chemotherapy, can we predict breast tumor response? Deep learning convolutional neural networks approach using a breast MRI tumor dataset. *J. Digit. Imaging* **2019**, *32*, 693–701. [\[CrossRef\]](#)
58. Turkki, R.; Byckhov, D.; Lundin, M.; Isola, J.; Nordling, S.; Kovanen, P.E.; Verrill, C.; von Smitten, K.; Joensuu, H.; Lundin, J.; et al. Breast cancer outcome prediction with tumour tissue images and machine learning. *Breast Cancer Res. Treat.* **2019**, *177*, 41–52. [\[CrossRef\]](#)
59. Toğaçar, M.; Ergen, B.; Cömert, Z. Application of breast cancer diagnosis based on a combination of convolutional neural networks, ridge regression and linear discriminant analysis using invasive breast cancer images processed with autoencoders. *Med. Hypotheses* **2020**, *135*, 10. [\[CrossRef\]](#)
60. Chen, H.; Gao, M.; Zhang, Y.; Liang, W.; Zou, X. Attention-Based Multi-NMF Deep Neural Network with Multimodality Data for Breast Cancer Prognosis Model. *BioMed Res. Int.* **2019**, *2019*, 11. [\[CrossRef\]](#)
61. Lu, Z.; Xu, S.; Shao, W.; Wu, Y.; Zhang, J.; Han, Z.; Feng, Q.; Huang, K. Deep-Learning-Based Characterization of Tumor-Infiltrating Lymphocytes in Breast Cancers from Histopathology Images and Multiomics Data. *JCO Clin. Cancer Inform.* **2020**, *4*, 480–490. [\[CrossRef\]](#)
62. Guo, Y.; Shang, X.; Li, Z. Identification of cancer subtypes by integrating multiple types of transcriptomics data with deep learning in breast cancer. *Neurocomputing* **2019**, *324*, 20–30. [\[CrossRef\]](#)
63. Lim, J.; Bang, S.; Kim, J.; Park, C.; Cho, J.; Kim, S. Integrative Deep Learning for Identifying Differentially Expressed (DE) Biomarkers. *Comput. Math. Methods Med.* **2019**, *2019*, 10. [\[CrossRef\]](#)
64. Sur, C. GSIAR: Gene-subcategory interaction-based improved deep representation learning for breast cancer subcategorical analysis using gene expression, applicable for precision medicine. *Med. Biol. Eng. Comput.* **2019**, *57*, 2483–2515. [\[CrossRef\]](#)
65. Zhang, D.; Zou, L.; Zhou, X.; He, F. Integrating Feature Selection and Feature Extraction Methods with Deep Learning to Predict Clinical Outcome of Breast Cancer. *IEEE Access* **2018**, *6*, 28936–28944. [\[CrossRef\]](#)
66. Xu, J.; Xiang, L.; Liu, Q.; Gilmore, H.; Wu, J.; Tang, J.; Madabhushi, A. Stacked Sparse Autoencoder (SSAE) for Nuclei Detection on Breast Cancer Histopathology Images. *IEEE Trans. Med. Imaging* **2016**, *35*, 119–130. [\[CrossRef\]](#)

67. Xu, J.; Xiang, L.; Hang, R.L.; Wu, J.Z. Stacked Sparse Autoencoder (SSAE) Based Framework for Nuclei Patch Classification on Breast Cancer Histopathology. In Proceedings of the 11th IEEE International Symposium on Biomedical Imaging (ISBI), Beijing, China, 29 April–2 May 2014; pp. 999–1002.
68. Shams, S.; Platania, R.; Zhang, J.; Kim, J.; Lee, K.; Park, S.J. Deep Generative Breast Cancer Screening and Diagnosis. In Proceedings of the 21st International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)/8th Eurographics Workshop on Visual Computing for Biology and Medicine (VCBM)/International Workshop on Computational Diffusion MRI (CDMRI), Granada, Spain, 16–21 September 2018; pp. 859–867.
69. Singh, V.K.; Rashwan, H.A.; Romani, S.; Akram, F.; Pandey, N.; Sarker, M.K.; Saleh, A.; Arenas, M.; Arquez, M.; Puig, D.; et al. Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network. *Expert Syst. Appl.* **2020**, *139*, 14. [[CrossRef](#)]
70. Fan, M.; Liu, Z.; Xu, M.; Wang, S.; Zeng, T.; Gao, X.; Li, L. Generative adversarial network-based super-resolution of diffusion-weighted imaging: Application to tumour radiomics in breast cancer. *NMR Biomed.* **2020**, *33*, 12. [[CrossRef](#)]
71. Guan, S.Y.; Loew, M. Breast cancer detection using synthetic mammograms from generative adversarial networks in convolutional neural networks. *J. Med. Imaging* **2019**, *6*, 10. [[CrossRef](#)]
72. Rawat, R.R.; Ruderman, D.; Macklin, P.; Rimm, D.L.; Agus, D.B. Correlating nuclear morphometric patterns with estrogen receptor status in breast cancer pathologic specimens. *Npj Breast Cancer* **2018**, *4*, 7. [[CrossRef](#)]
73. Cruz-Roa, A.; Gilmore, H.; Basavanthally, A.; Feldman, M.; Ganesan, S.; Shih, N.; Tomaszewski, J.; Madabhushi, A.; González, F. High-throughput adaptive sampling for whole-slide histopathology image analysis (HASHI) via convolutional neural networks: Application to invasive breast cancer detection. *PLoS ONE* **2018**, *13*, 23. [[CrossRef](#)]
74. Turkki, R.; Linder, N.; Kovanen, P.E.; Pellinen, T.; Lundin, J. Antibody-supervised deep learning for quantification of tumor-infiltrating immune cells in hematoxylin and eosin stained breast cancer samples. *J. Pathol. Inform.* **2016**, *7*, 38. [[CrossRef](#)]
75. Khameneh, F.D.; Razavi, S.; Kamasak, M. Automated segmentation of cell membranes to evaluate HER2 status in whole slide images using a modified deep learning network. *Comput. Biol. Med.* **2019**, *110*, 164–174. [[CrossRef](#)]
76. Murtaza, G.; Shuib, L.; Mujtaba, G.; Raza, G. Breast Cancer Multi-classification through Deep Neural Network and Hierarchical Classification Approach. *Multimed. Tools Appl.* **2020**, *79*, 15481–15511. [[CrossRef](#)]
77. Bardou, D.; Zhang, K.; Ahmad, S.M. Classification of Breast Cancer Based on Histology Images Using Convolutional Neural Networks. *IEEE Access* **2018**, *6*, 24680–24693. [[CrossRef](#)]
78. Tellez, D.; Balkenhol, M.; Otte-Holler, I.; van de Loo, R.; Vogels, R.; Bult, P.; Wauters, C.; Vreuls, W.; Mol, S.; Karssemeijer, N.; et al. Whole-Slide Mitosis Detection in H&E Breast Histology Using PHH3 as a Reference to Train Distilled Stain-Invariant Convolutional Networks. *IEEE Trans. Med. Imaging* **2018**, *37*, 2126–2136. [[CrossRef](#)]
79. Luo, J.; Ning, Z.; Zhang, S.; Feng, Q.; Zhang, Y. Bag of deep features for preoperative prediction of sentinel lymph node metastasis in breast cancer. *Phys. Med. Biol.* **2018**, *63*, 11. [[CrossRef](#)]
80. Elbashir, M.K.; Ezz, M.; Mohammed, M.; Saloum, S.S. Lightweight Convolutional Neural Network for Breast Cancer Classification Using RNA-Seq Gene Expression Data. *IEEE Access* **2019**, *7*, 185338–185348. [[CrossRef](#)]
81. Saha, M.; Arun, I.; Ahmed, R.; Chatterjee, S.; Chakraborty, C. HscoreNet: A Deep network for estrogen and progesterone scoring using breast IHC images. *Pattern Recognit.* **2020**, *102*, 11. [[CrossRef](#)]
82. Ha, R.; Mutasa, S.; Karcich, J.; Gupta, N.; Van Sant, E.P.; Nemer, J.; Sun, M.; Chang, P.; Liu, M.Z.; Jambawalikar, S. Predicting Breast Cancer Molecular Subtype with MRI Dataset Utilizing Convolutional Neural Network Algorithm. *J. Digit. Imaging* **2019**, *32*, 276–282. [[CrossRef](#)] [[PubMed](#)]
83. Zheng, J.; Lin, D.; Gao, Z.; Wang, S.; He, M.; Fan, J. Deep Learning Assisted Efficient AdaBoost Algorithm for Breast Cancer Detection and Early Diagnosis. *IEEE Access* **2020**, *8*, 96946–96954. [[CrossRef](#)]
84. Tewary, S.; Arun, I.; Ahmed, R.; Chatterjee, S.; Chakraborty, C. AutoIHC-scoring: A machine learning framework for automated Allred scoring of molecular expression in ER- and PR-stained breast cancer tissue. *J. Microsc.* **2017**, *268*, 172–185. [[CrossRef](#)]
85. Fang, Y.; Zhao, J.; Hu, L.; Ying, X.; Pan, Y.; Wang, X. Image classification toward breast cancer using deeply-learned quality features. *J. Vis. Commun. Image Represent.* **2019**, *64*, 7. [[CrossRef](#)]
86. Danaee, P.; Ghaeini, R.; Hendrix, D.A. A Deep Learning Approach for Cancer Detection and Relevant Gene Identification. In Proceedings of the 22nd Pacific Symposium on Biocomputing (PSB), Kohala Coast, 4–8 January 2017; pp. 219–229.
87. Punitha, S.; Amuthan, A.; Joseph, K.S. Enhanced Monarchy Butterfly Optimization Technique for effective breast cancer diagnosis. *J. Med. Syst.* **2019**, *43*, 14. [[CrossRef](#)]
88. Romo-Bucheli, D.; Janowczyk, A.; Gilmore, H.; Romero, E.; Madabhushi, A. Automated Tubule Nuclei Quantification and Correlation with Oncotype DX risk categories in ER+ Breast Cancer Whole Slide Images. *Sci. Rep.* **2016**, *6*, 9. [[CrossRef](#)]
89. Zemouri, R.; Omri, N.; Morello, B.; Devalland, C.; Arnould, L.; Zerhouni, N.; Fnaiech, F. Constructive Deep Neural Network for Breast Cancer Diagnosis. *IFAC-Pap.* **2018**, *51*, 98–103. [[CrossRef](#)]
90. Salma, M.U.; Doreswamy, N. Hybrid BATGSA: A metaheuristic model for classification of breast cancer data. *Int. J. Adv. Intell. Paradig.* **2020**, *15*, 207–227. [[CrossRef](#)]
91. Karakış, R.; Tez, M.; Kılıç, Y.; Kuru, Y.; Güler, I. A genetic algorithm model based on artificial neural network for prediction of the axillary lymph node status in breast cancer. *Eng. Appl. Artif. Intell.* **2013**, *26*, 945–950. [[CrossRef](#)]
92. Shimizu, H.; Nakayama, K.I. A 23 gene-based molecular prognostic score precisely predicts overall survival of breast cancer patients. *EBioMedicine* **2019**, *46*, 150–159. [[CrossRef](#)] [[PubMed](#)]

93. Jafari-Marandi, R.; Davarzani, S.; Gharibdousti, M.S.; Smith, B. An optimum ANN-based breast cancer diagnosis: Bridging gaps between ANN learning and decision-making goals. *Appl. Soft Comput.* **2018**, *72*, 108–120. [[CrossRef](#)]
94. Tripathy, R.K.; Mahanta, S.; Paul, S. Artificial intelligence-based classification of breast cancer using cellular images. *RSC Adv.* **2014**, *4*, 9349–9355. [[CrossRef](#)]
95. Giordano, A.; Giuliano, M.; De Laurentiis, M.; Eleuteri, A.; Iorio, F.; Tagliaferri, R.; Hortobagyi, G.N.; Pusztai, L.; De Placido, S.; Hess, K.; et al. Artificial neural network analysis of circulating tumor cells in metastatic breast cancer patients. *Breast Cancer Res. Treat.* **2011**, *129*, 451–458. [[CrossRef](#)]
96. Mert, A.; Kılıç, N.; Bilgili, E.; Akan, A. Breast Cancer Detection with Reduced Feature Set. *Comput. Math. Methods Med.* **2015**, *2015*, 11. [[CrossRef](#)] [[PubMed](#)]
97. Hajiabadi, H.; Babaiyan, V.; Zabihzadeh, D.; Hajiabadi, M. Combination of loss functions for robust breast cancer prediction. *Comput. Electr. Eng.* **2020**, *84*, 12. [[CrossRef](#)]
98. Agrawal, U.; Soria, D.; Wagner, C.; Garibaldi, J.; Ellis, I.O.; Bartlett, J.M.; Cameron, D.; Rakha, E.A.; Green, A.R. Combining clustering and classification ensembles: A novel pipeline to identify breast cancer profiles. *Artif. Intell. Med.* **2019**, *97*, 27–37. [[CrossRef](#)]
99. Denkiewicz, M.; Saha, I.; Rakshit, S.; Sarkar, J.P.; Plewczynski, D. Identification of Breast Cancer Subtype Specific MicroRNAs Using Survival Analysis to Find Their Role in Transcriptomic Regulation. *Front. Genet.* **2019**, *10*, 19. [[CrossRef](#)]
100. Rawat, R.R.; Ortega, I.; Roy, P.; Sha, F.; Shibata, D.; Ruderman, D.; Agus, D.B. Deep learned tissue “fingerprints” classify breast cancers by ER/PR/Her2 status from H&E images. *Sci. Rep.* **2020**, *10*, 7275. [[CrossRef](#)]
101. Rouhi, R.; Jafari, M.; Kasaei, S.; Keshavarzian, P. Benign and malignant breast tumors classification based on region growing and CNN segmentation. *Expert Syst. Appl.* **2015**, *42*, 990–1002. [[CrossRef](#)]
102. Cheng, J.-Z.; Ni, D.; Chou, Y.-H.; Qin, J.; Tiu, C.-M.; Chang, Y.-C.; Huang, C.-S.; Shen, D.; Chen, C.-M. Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans. *Sci. Rep.* **2016**, *6*, 13. [[CrossRef](#)] [[PubMed](#)]
103. Thuy, M.B.H.; Hoang, V.T. Fusing of Deep Learning, Transfer Learning and GAN for Breast Cancer Histopathological Image Classification. In Proceedings of the 6th International Conference on Computer Science, Applied Mathematics and Applications (ICCSAMA), Hanoi, Vietnam, 19–20 December 2019; pp. 255–266.
104. Swiecicki, A.; Buda, M.; Saha, A.; Li, N.; Ghate, S.V.; Walsh, R.; Mazurowski, M.A. Generative adversarial network-based image completion to identify abnormal locations in digital breast tomosynthesis images. In Proceedings of the Conference on Medical Imaging—Computer-Aided Diagnosis, Houston, TX, USA, 16–19 February 2020.
105. Tien, H.-J.; Yang, H.-C.; Shueng, P.-W.; Chen, J.-C. Cone-beam CT image quality improvement using Cycle-Deblur consistent adversarial networks (Cycle-Deblur GAN) for chest CT imaging in breast cancer patients. *Sci. Rep.* **2021**, *11*, 12. [[CrossRef](#)] [[PubMed](#)]
106. Umamaheswari, T.S.; Sumathi, P. Enhanced firefly algorithm (EFA) based gene selection and adaptive neuro neutrosophic inference system (ANNIS) prediction model for detection of circulating tumor cells (CTCs) in breast cancer analysis. *Clust. Comput.* **2019**, *22*, 14035–14047. [[CrossRef](#)]
107. Zhang, F.; Chen, J.; Wang, M.; Drabier, R. A neural network approach to multi-biomarker panel discovery by high-throughput plasma proteomics profiling of breast cancer. *BMC Proc.* **2013**, *7*, S10. [[CrossRef](#)] [[PubMed](#)]
108. Jaber, M.I.; Song, B.; Taylor, C.; Vaske, C.J.; Benz, S.; Rabizadeh, S.; Soon-Shiong, P.; Szeto, C.W. A deep learning image-based intrinsic molecular subtype classifier of breast tumors reveals tumor heterogeneity that may affect survival. *Breast Cancer Res.* **2020**, *22*, 10. [[CrossRef](#)] [[PubMed](#)]
109. Inan, O.; Uzer, M.S.; Yilmaz, N. A New Hybrid Feature Selection Method Based on Association Rules and PCA for Detection of Breast Cancer. *Int. J. Innov. Comp. Inf. Control* **2013**, *9*, 727–739.
110. Mert, A.; Kılıç, N.; Akan, A. An improved hybrid feature reduction for increased breast cancer diagnostic performance. *Biomed. Eng. Lett.* **2014**, *4*, 285–291. [[CrossRef](#)]
111. Ronoud, S.; Asadi, S. An evolutionary deep belief network extreme learning-based for breast cancer diagnosis. *Soft Comput.* **2019**, *23*, 13139–13159. [[CrossRef](#)]
112. Chen, R.; Yang, L.; Goodison, S.; Sun, Y. Deep-learning approach to identifying cancer subtypes using high-dimensional genomic data. *Bioinformatics* **2020**, *36*, 1476–1483. [[CrossRef](#)]
113. Liu, Q.; Hu, P. Association Analysis of Deep Genomic Features Extracted by Denoising Autoencoders in Breast Cancer. *Cancers* **2019**, *11*, 494. [[CrossRef](#)]
114. Jiang, P.; Huang, S.; Fu, Z.; Sun, Z.; Lakowski, T.M.; Hu, P. Deep graph embedding for prioritizing synergistic anticancer drug combinations. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 427–438. [[CrossRef](#)] [[PubMed](#)]
115. Cruz-Roa, A.; Gilmore, H.; Basavanthally, A.; Feldman, M.; Ganesan, S.; Shih, N.N.; Tomaszewski, J.; González, F.A.; Madabhushi, A. Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent. *Sci. Rep.* **2017**, *7*, 14. [[CrossRef](#)] [[PubMed](#)]
116. Cui, X.; Li, Z.; Zhao, Y.; Song, A.; Shi, Y.; Hai, X.; Zhu, W. Breast cancer identification via modeling of peripherally circulating miRNAs. *PeerJ* **2018**, *6*, e4551. [[CrossRef](#)] [[PubMed](#)]
117. Beykikhoshk, A.; Quinn, T.P.; Lee, S.C.; Tran, T.; Venkatesh, S. DeepTRIAGE: Interpretable and individualised biomarker scores using attention mechanism for the classification of breast cancer sub-types. *BMC Med. Genom.* **2020**, *13*, 20. [[CrossRef](#)]

118. Lin, C.-Y.; Ruan, P.; Li, R.; Yang, J.-M.; See, S.; Song, J.; Akutsu, T. Deep learning with evolutionary and genomic profiles for identifying cancer subtypes. *J. Bioinform. Comput. Biol.* **2019**, *17*, 15. [[CrossRef](#)]
119. Tomczak, K.; Czerwińska, P.; Wiznerowicz, M. Review the Cancer Genome Atlas (TCGA): An immeasurable source of knowledge. *Contemp. Oncol.* **2015**, *2015*, 68–77. [[CrossRef](#)]
120. Luo, P.; Ding, Y.; Lei, X.; Wu, F.-X. deepDriver: Predicting Cancer Driver Genes Based on Somatic Mutations Using Deep Convolutional Neural Networks. *Front. Genet.* **2019**, *10*, 13. [[CrossRef](#)]
121. Rahman, A.; Muniyandi, R.C. An Enhancement in Cancer Classification Accuracy Using a Two-Step Feature Selection Method Based on Artificial Neural Networks with 15 Neurons. *Symmetry* **2020**, *12*, 271. [[CrossRef](#)]
122. Chougrad, H.; Zouaki, H.; Alheyane, O. Multi-label transfer learning for the early diagnosis of breast cancer. *Neurocomputing* **2020**, *392*, 168–180. [[CrossRef](#)]
123. Clark, K.; Vendt, B.; Smith, K.; Freymann, J.; Kirby, J.; Koppel, P.; Moore, S.; Phillips, S.; Maffitt, D.; Pringle, M.; et al. The Cancer Imaging Archive (TCIA): Maintaining and Operating a Public Information Repository. *J. Digit. Imaging* **2013**, *26*, 1045–1057. [[CrossRef](#)]
124. Hamidinekoo, A.; Denton, E.; Rampun, A.; Honnor, K.; Zwigelaar, R. Deep learning in mammography and breast histology, an overview and future trends. *Med. Image Anal.* **2018**, *47*, 45–67. [[CrossRef](#)] [[PubMed](#)]
125. Xiao, Y.; Xu, C.; Ping, Y.; Guan, J.; Fan, H.; Li, Y.; Li, X. Differential expression pattern-based prioritization of candidate genes through integrating disease-specific expression data. *Genomics* **2011**, *98*, 64–71. [[CrossRef](#)] [[PubMed](#)]
126. Majeed, H.; Nguyen, T.H.; Kandel, M.E.; Kajdacsy-Balla, A.; Popescu, G. Label-free quantitative evaluation of breast tissue using Spatial Light Interference Microscopy (SLIM). *Sci. Rep.* **2018**, *8*, 9. [[CrossRef](#)] [[PubMed](#)]
127. Moreira, I.C.; Amaral, I.; Domingues, I.; Cardoso, A.; Cardoso, M.J.; Cardoso, J.S. INbreast: Toward a Full-field Digital Mammographic Database. *Acad. Radiol.* **2012**, *19*, 236–248. [[CrossRef](#)] [[PubMed](#)]
128. Zhang, D.; Zhou, L. Discovering Golden Nuggets: Data Mining in Financial Application. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **2004**, *34*, 513–522. [[CrossRef](#)]
129. Pejic-Bach, M. Profiling Intelligent Systems Applications in Fraud Detection and Prevention: Survey of Research Articles. In Proceedings of the UKSim/AMSS 1st International Conference on Intelligent Systems, Modelling and Simulation, Liverpool, UK, 27–29 January 2010; pp. 80–85.
130. Wang, X.G.; Wu, H.; Yi, Z.C. Research on Bank Anti-fraud Model Based on K-means and Hidden Markov Model. In Proceedings of the 3rd IEEE International Conference on Image, Vision and Computing (ICIVC), Chongqing, China, 27–29 June 2018; pp. 780–784.
131. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* **2002**, *16*, 321–357. [[CrossRef](#)]
132. Tan, Y.H.; Chan, C.S. Phrase-based image caption generator with hierarchical LSTM network. *Neurocomputing* **2019**, *333*, 86–100. [[CrossRef](#)]
133. Zhang, Z.Z.; Xie, Y.P.; Xing, F.Y.; McGough, M.; Yang, L. MDNet: A Semantically and Visually Interpretable Medical Image Diagnosis Network. In Proceedings of the 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 3549–3557.
134. Wang, H.; Feng, J.; Zhang, Z.; Su, H.; Cui, L.; He, H.; Liu, L. Breast mass classification via deeply integrating the contextual information from multi-view data. *Pattern Recognit.* **2018**, *80*, 42–52. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.