ScienceDirect

Artificial Intelligence in Medicine

Volume 127, May 2022, 102276

# Breast cancer detection using artificial intelligence techniques: A systematic literature review

Ali Bou Nassif, Manar Abu Talib, Qassim Nasir, Yaman Afadar, Omar Elgendy

Show more

Share Cite

https://doi.org/10.1016/j.artmed.2022.102276 ↗

Get rights and content ↗

## Highlights

- Systematic review on detection breast cancer (BC) using genetic sequencing imaging
- 80 papers were extracted related to machine and deep learning algorithms.
- Deep learning models are very popular for breast cancer detection.
- RNN and GAN models have not been widely used in the field.
- Challenges and opportunities in BC are proposed as potential research directions.

## Abstract

Cancer is one of the most dangerous diseases to humans, and yet no permanent cure has been developed for it. Breast cancer is one of the most common cancer types. According to the National Breast Cancer Foundation, in 2020 alone, more than 276,000 new cases of invasive breast cancer and more than 48,000 non-invasive cases were diagnosed in the US. To put these figures in perspective, 64% of these cases are diagnosed early in the disease's cycle, giving patients a 99% chance of survival. Artificial intelligence and machine learning have been used effectively in detection and treatment of several dangerous diseases, helping in early diagnosis and treatment, and thus increasing the patient's chance of survival. Deep learning has been designed to analyze the most important features affecting detection and treatment of serious diseases. For example, breast cancer can be detected using genes or histopathological imaging. Analysis at the genetic level is very expensive, so histopathological imaging is the most common approach used to detect breast cancer. In this research work, we systematically reviewed previous work done on detection and treatment of breast cancer using genetic sequencing or histopathological imaging with the help of deep learning and machine learning. We also provide recommendations to researchers who will work in this field.

## Introduction

Breast cancer is one of the major causes of death in women around the world. According to the American cancer society, 41,760 women and more than 500 men died from breast cancer recently.[1] Breast cancer occurs in four main types: normal, benign, in-situ carcinoma and invasive carcinoma [1]. A benign tumor involves a minor change in the breast structure. It is not harmful and does not classify as a harmful cancer. In cases of in-situ carcinoma, the cancer is only in the mammary duct lobule system and does not affect other organs. This type is not dangerous and can be treated if diagnosed early. Invasive carcinoma is considered to be the most dangerous type of breast cancer, as it can spread to all other organs. According to the authors in [1], breast cancer can be detected using several methods including X-ray mammography, ultrasound (US), computed tomography (CT), portion emission tomography (PET), magnetic resonance imaging (MRI) and breast temperature measurement. Usually, the golden standard is a pathological diagnosis for detecting breast cancer. This involves an image analysis of the removed tissue, which is stained in the lab to increase visibility. Hematoxylin and eosin (H&E) are commonly used for the staining process. Breast cancer can be diagnosed using one of two approaches: histopathological image analysis or genomics. Histopathological images are microscopic images of breast tissue that are extremely useful in early treatment of the cancer. As for genomics, the authors in [2] stated that radio-genomics is an

emerging research field focusing on multi-scale associations between medical imaging and gene expression data.

Radio-genomics provide both radiological and genetic features that can enhance diagnosis. It can analyze tissues at the molecular level, helping with prediction and early detection of cancer. The main difference between imaging information and radio-genomics is the critical knowledge gap between imaging at the tissue level and analyzing the underlying molecular and genetic disease biomarkers. As imaging is less precise, it may lead to over- or under-treatment. While radio-genomics is much more effective than histopathological imaging, it is rarely used because the process involves datasets that are very expensive and require high computational power. As a result, a limited number of labs conduct radio-genomics experiments [2]. This research paper addresses the following research questions and highlights the deep learning models, looking at their performance, the datasets used and possibilities for breast cancer classification and detection.

1. **Which deep learning models perform most effectively?**

   We will compare deep learning models with classical machine learning models to compare their performance. We will also list the performance metrics used.

2. **What are the most used features for breast cancer classification? How are these features selected and extracted?**

   We will observe the most important features that contribute to breast cancer classification, and the methods used to extract these features.

3. **What datasets are available for both gene sequencing and MRI? What feature selection and extraction methods are used?**

   We will list and discuss all public and private datasets for gene sequencing and MRI imaging data. We will also list some of the methods used to select and extract the features.

4. **Comparing gene sequence data with image data, for breast cancer detection problem, what are the drawbacks, challenges, and advantages?**

   We will compare imaging and gene sequencing as they relate to breast cancer detection, using a tabular presentation to highlight the main differences between the two approaches.

The remaining of this paper is structured as follows: Section 2 presents related work which includes surveys conducted in breast cancer area. Section 3 explains the methodology used to conduct this research. Section 4 presents the obtained results and related discussions. Lastly, Section 5 concludes the paper and suggests future research directions.

---

## Section snippets

### Related work

Many studies have been conducted about breast cancer detection through imaging or through genomics. However, to the best of our knowledge, no research has been conducted including both techniques.

The authors in [1] summarized the various techniques used to classify breast cancer using histopathological image analysis (HIA) based on different architectures of artificial neural networks (ANN). The authors grouped their work according to the applied dataset. They arranged it in ascending …

### Methodology

Our target topic is breast cancer detection using deep learning. We ended up using around 80 of the most recent papers related to breast cancer treatment and diagnosis. Some of the papers examined only deep learning, while others used a combination of machine learning and deep learning.

In our search process, we mainly used the Scopus database to obtain the articles. This is to exclude non-refereed publications However, in Fig. 1, we state the distribution of selected papers among the existing …

### Results and discussion

In this systematic study, our initial search turned up 1000 conference and journal papers. After eliminating duplicated papers and unrelated studies that were “purely medical or about cancer in general”, we ended up with 80 papers related to both ML and DL.

We wanted to focus on DL approaches or DL-ML hybrid models, so only papers related to DL were selected. Fig. 3 explains our search methodology. …

## Conclusions and future research directions

Most papers published in the field of breast cancer detection and subtype classification use machine learning techniques. However, deep learning models have not been heavily investigated in this domain. This presents researchers with opportunities to use various deep learning mechanisms to predict patient status such as LSTM, GAN and RNN, as these types of research have not yet been conducted in the field.

Moreover, most papers focus only on the accuracy metric to evaluate their performance and …

## Compliance with ethical standards

The authors thank the University of Sharjah for supporting this work.

Funding is through the competitive project “Applications of Machine Learning in Metastatic Breast Cancer Detection”. …

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper. …

Recommended articles

## References (94)

T. Pang *et al.*
Deep learning radiomics in breast cancer with different modalities: overview and future
Expert Syst Appl (2020)

P. Jiang *et al.*
Deep graph embedding for prioritizing synergistic anticancer drug combinations
Comput Struct Biotechnol J (2020)

A. El-Nabawy *et al.*

A feature-fusion framework of clinical, genomics, and histopathological data for METABRIC breast cancer subtype classification

Appl Soft Comput (2020)

R. Karakış *et al.*

A genetic algorithm model based on artificial neural network for prediction of the axillary lymph node status in breastcancer

Eng Appl Artif Intel (2013)

M.A. Alsaleem *et al.*

A novel prognostic two-gene signature for triple negative breast cancer

Mod Pathol (2020)

B. Ehteshami Bejnordi *et al.*

Using deep convolutional neural networks to identify and classify tumor-associated stroma in diagnostic breast biopsies

Mod Pathol (2018)

H. Shimizu *et al.*

A 23 gene–based molecular prognostic score precisely predicts overall survival of breast cancer patients

EBioMedicine (2019)

J. Xu *et al.*

A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images

Neurocomputing (2016)

S. Simsek *et al.*

A hybrid data mining approach for identifying the temporal effects of variables associated with breast cancer survival

Expert Syst Appl (2020)

R. Jafari-Marandi *et al.*

An optimum ANN-based breast cancer diagnosis: bridging gaps between ANN learning and decision-making goals

Appl Soft Comput J (2018)

View more references

## Cited by (229)

### A novel breast cancer detection architecture based on a CNN-CBR system for mammogram classification

2023, Computers in Biology and Medicine

Show abstract

### BreastScreening-AI: Evaluating medical intelligent agents for human-AI interactions

2022, Artificial Intelligence in Medicine

> *Citation Excerpt :*
>
> ...In breast cancer, medical imaging systems allow the end-user to diagnose several modalities, such as MammoGraphy (MG), UltraSound (US) or Magnetic Resonance Imaging (MRI), from the retrieval of medical imaging data [58,127]. A wide range of Clinical Decision Support Systems (CDSS) exist, providing clinicians with the knowledge to enhance the clinical workflow [28,55,64], from systems that supply potential information for medical decision-making to those that make diagnostic decisions [77,111]. Experts may resist using a system if it does not capture the nuances of their mental models or provides relevant context [84,87,141]....

Show abstract

### Machine learning-based models for the prediction of breast cancer recurrence risk

2023, BMC Medical Informatics and Decision Making

Show abstract

### A Comprehensive Review on Breast Cancer Detection, Classification and Segmentation Using Deep Learning

2023, Archives of Computational Methods in Engineering

### Object Detection Using Deep Learning, CNNs and Vision Transformers: A Review

2023, IEEE Access

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

2023, Diagnostics

View all citing articles on Scopus ↗

View full text