

## **Breast cancer detection using AI: A literature review**

### **Introduction**

According to the National Breast Cancer foundation, in 2020, more than 276,000 new cases of invasive breast cancer and more than 48,000 non-invasive cases were diagnosed in the US. 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. These techniques become the need of an hour to cater the growing patient population and for the improvement in the Healthcare system. In this literature review, we aim to elucidate the insufficiencies of traditional methods in achieving optimal detection. Subsequently, we will delineate pertinent artificial intelligence (AI) (Deep learning (DL)) techniques, emphasizing the potential for integrating AI into cancer imaging, thereby showcasing their advantages over conventional approaches. Finally, we will explicate the outcomes yielded by these algorithms.[6]

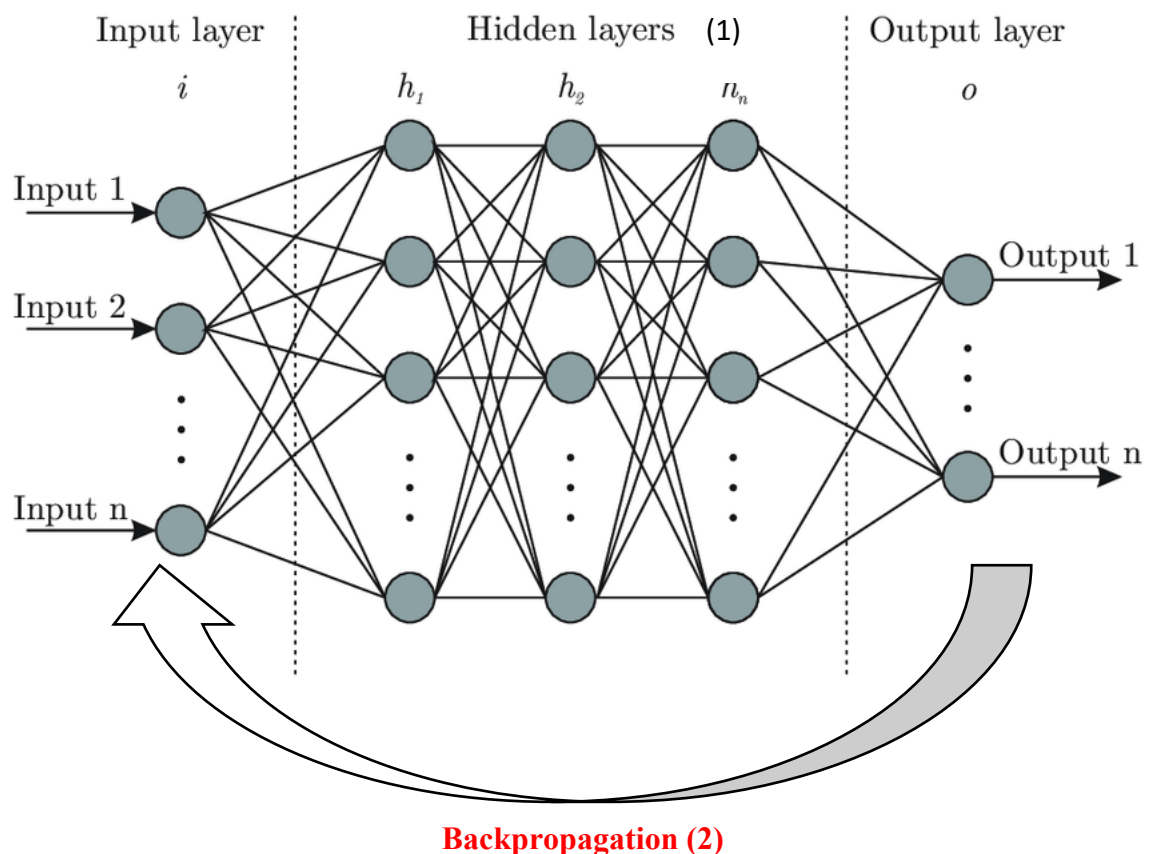
### **The Imperative to Go Beyond Traditional Approaches for Breast Cancer Detection**

Medical imaging is typically performed manually by skilled doctors, such as sinologists, radiologists, or pathologists using several methods including X-ray mammography, ultrasound...[3] However, manual histopathology image analysis presents several challenges. Firstly, there is a scarcity of expert pathologists in low-income and developing countries. Secondly, the process of multi-class classification with image analysis is time-consuming and burdensome for pathologists. Thirdly, pathologists may experience reduced attention and fatigue during image analysis. Lastly, the accurate identification of breast cancer subtypes relies on the domain knowledge and professional experience of skilled pathologists. These issues, especially in early-stage breast cancer, can lead to misdiagnosis. [1]

### **The introduction of methods based on deep learning (ANN)**

Artificial intelligence (AI), which involves the development of machines or tools capable of emulating human thinking and behavior, is profoundly transforming various scientific fields, including healthcare. Within AI, deep learning stands out as a machine learning approach that automatically learns feature representations from input data, eliminating the need for human-engineered features to achieve optimal

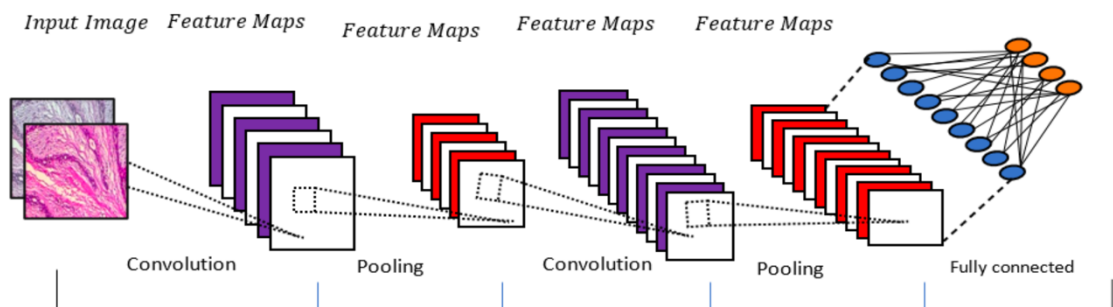
performance. At the core of deep learning are artificial neural networks (ANNs), fundamental building blocks that enable the creation of powerful models capable of extracting complex patterns and representations from data. In the context of breast cancer detection, an ANN is constructed by gathering relevant data that includes characteristics such as patient age, tumor size, and lymph node status. This data is then preprocessed to ensure quality and consistency, involving tasks like removing missing values, and splitting it into training and testing sets. The ANN comprises interconnected layers (fig 1(1)) of artificial neurons, with the input layer receiving the preprocessed data and subsequent hidden layers processing and transforming the information. [3],[5] The final output layer provides the prediction or diagnosis. Training the ANN involves the iterative process of Backpropagation (fig 1(2)), where the network's parameters, or weights, are initially randomly assigned. The training data is fed through the network, and the output is compared to the known correct output to calculate the error. This error is then used to adjust the weights in a way that minimizes it, gradually improving the network's predictive capabilities. Once training is completed, the ANN is tested using a separate dataset to evaluate its performance, comparing predicted results with known correct results to assess accuracy and generalization ability. Once trained and validated, the ANN can be utilized to predict the presence or absence of breast cancer. [4],[7]



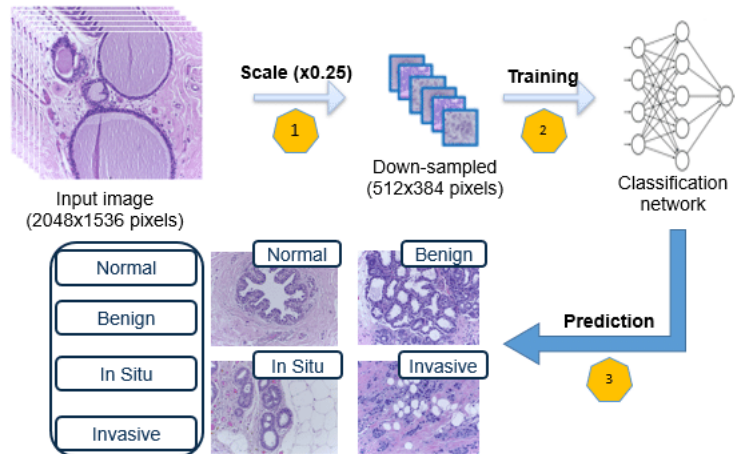
**Figure 1: ANN architecture**

## Convolution neural network (CNN)

The mostly used algorithm for breast cancer detecting and subtype classification is convolutional neural network (CNN). Basically, three layers make up the CNN: a convolutional layer, a pooling layer (apply filters and reduce the images) and fully connected layers (neural network) (Fig 2). These layers are stacked to automatically extracting the features from breast cancer data (Fig 3). [3],[7]



**Figure 2:** CNN architecture



**Figure 3:** Functioning of the CNN

In image or breast data classification, two types of CNN models are commonly used: de novo trained models and transfer learning-based models [3], [4]. De novo trained models are built from scratch, which requires the total design of a CNN architecture and training on a large labeled tumor image dataset. This approach requires

considerable resources and expertise. In contrast, transfer learning-based models rely on pre-trained CNNs, which have learned general visual knowledge from large-scale datasets such as ImageNet. These models use pre-trained weights and features as a starting point, reducing the need for intensive training from scratch. Models based on transfer learning are more efficient.

## Results and limitation

Studies have shown that CNNs can achieve comparable or even superior performance to radiologists in breast cancer detection. For example, a study published in the journal Nature in 2019 reported that a CNN achieved a sensitivity of 90.2% and a specificity of 90.8% in detecting breast cancer from mammographic images [4]. However, creating a large dataset of breast cancer images requires a significant number of images and expert physicians to accurately label them. Building such datasets is therefore time-consuming and challenging. That's why researchers often prefer to utilize existing public datasets. However, models trained on public datasets may not necessarily provide the optimal solution. [3]

## Conclusion

In conclusion, the advancements in utilizing convolutional neural networks (CNN) for breast cancer detection have shown great promise. CNN models have demonstrated comparable or even superior performance to radiologists in identifying breast cancer from mammographic images. These findings highlight the potential of deep learning techniques to enhance the accuracy and efficiency of breast cancer detection.

## References

- [1] Charnpreet Kaur, Urvashi Garg, Artificial intelligence techniques for cancer detection in medical image processing <https://www.sciencedirect.com/science/article/abs/pii/S2214785321031618>, 2021
- [2] Ali Bou Nassif\*, Manar Abu Talib, Qassim Nasir, Yaman Afadar, Omar Elgendy Breast cancer detection using artificial intelligence techniques <https://arxiv.org/pdf/2203.04308.pdf>, 2020
- [3] Subrato Bharati<sup>1</sup>, Prajoy Podder<sup>2</sup>, M. Rubaiyat Hossain Mondal<sup>3</sup> Artificial Neural Network Based Breast Cancer Screening <https://arxiv.org/pdf/2006.01767.pdf>, 2023
- [4] Maged Nasser, Umi Kalsom Yusof Deep Learning Based Methods for Breast Cancer Diagnosis, 2023

[5] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton. ImageNet Classification with Deep Convolutional Neural Networks, 2012.

[6] Dr Jean Seely <https://www.youtube.com/watch?v=Fs6JKxFkLtq>, 2023

[7] Axel Carlier <https://moodle-n7.inp-toulouse.fr/course/view.php?id=1476> 2023