

# AI for breast cancer detection

Tommaso Ballarin

tommaso.ballarin@studenti.unipd.it

Paolo Conti

paolo.conti@studenti.unipd.it

Filippo Demo

filippo.demo@studenti.unipd.it

Piersilvio Spicoli

piersilvio.spicoli@studenti.unipd.it

Department of Information Engineering, University of Padua

## Abstract

This paper explores the application of AI in the diagnosis and prevention of breast cancer, with a focus on several approaches, highlighting both the scientific advances and the clinical relevance of these technologies. In particular, machine learning and image-based methods will be analyzed, focusing on their strengths and weaknesses. More agents based on these methods can be included in a multi-agent system, so as to exploit their individual strengths, thus achieving a high accuracy of the diagnosis. The paper presents an example of a multi-agent system for breast cancer detection, integrated with blockchain technology, which makes the decision-making process accessible only to a few trusted agents. In this way, the privacy of sensitive data and the accuracy of the process can be ensured.

## 1 Introduction

In recent years, artificial intelligence (AI) has emerged as a transformative force in the field of healthcare, offering the potential to improve diagnostic accuracy, optimize treatment pathways, and improve overall patient outcomes. Using advanced algorithms capable of learning from vast and complex datasets, AI technologies are redefining the way medical professionals detect diseases, interpret medical images, manage clinical data, and make therapeutic decisions. Healthcare systems around the world are facing increasing pressure due to the increase in patient volumes, the aging population, and the growing complexity of chronic diseases. In this context, AI provides scalable solutions that can support clinicians in routine tasks and complex decision making, thus improving efficiency and reducing burnout. Applications range from natural language processing in electronic health records, predictive analytics for disease progression, and computer vision systems that assist in radiology, pathology, and dermatology. Among these domains, medical imaging represents one of the most promising areas for AI integration. Deep learning,

a subset of machine learning, has shown impressive performance in tasks such as image classification, segmentation, and anomaly detection. In radiology, AI models have achieved expert-level accuracy in the detection of various conditions, including lung nodules, brain tumors, diabetic retinopathy, and breast cancer.

The adoption of AI in healthcare also raises important challenges, including the need for robust validation, transparency, fairness, and integration within clinical workflows. Regulatory frameworks are evolving to address these issues and ensure that AI-driven solutions are safe, effective, and ethical. As the field continues to evolve, the synergy between human expertise and machine intelligence holds great promise for the future of medicine.

### 1.1 Background and Significance

Breast cancer is the most commonly diagnosed cancer among women worldwide, accounting for approximately 2.3 million new cases and nearly 685,000 deaths in 2022 alone. It represents about one in four cancer cases among women. In Italy, breast cancer similarly ranks as the most prevalent form of cancer in the female population. Recent estimates show that about 57,500 new cases were diagnosed in 2022, corresponding to 28.2% of all new female cancer diagnoses (Organization 2023). Despite significant advances in treatment, early detection remains critical in reducing mortality, making screening and diagnostic technologies essential in public health strategies.

### 1.2 Current Imaging Technologies for Detection and Diagnosis

Early detection through screening plays a vital role in improving the prognosis. Mammography is currently the gold standard for the screening of breast cancer, widely used in national screening programs. It uses low-dose X-rays to detect abnormalities such as masses or microcalcifications. Ultrasound is commonly used as an adjunct imaging tool, especially in women with dense breasts where mammography

sensitivity is lower. It is useful for characterizing palpable masses and distinguishing cystic from solid lesions. Magnetic Resonance Imaging (MRI), particularly contrast-enhanced breast MRI, offers the highest sensitivity among imaging modalities and is typically used for high-risk women. Although highly sensitive, magnetic resonance imaging is also associated with higher false positive rates and is more expensive, limiting its use in population-wide screening (Mann 2019).

### 1.3 AI in Imaging-Based Diagnosis

The increasing complexity and volume of medical imaging data have positioned artificial intelligence (AI), particularly deep learning (DL) techniques, as transformative tools in breast cancer diagnostics. AI-based systems, especially those that use convolutional neural networks (CNNs), have shown remarkable success in image classification, lesion detection, and segmentation tasks. These models are trained on large annotated datasets to recognize patterns indicative of malignancy that may not be apparent to human readers. Recent studies have shown that AI algorithms can achieve diagnostic performance comparable to, or in some cases exceeding, that of experienced radiologists. For example, a landmark study published in *Nature* (McKinney 2020) showed that a deep learning model reduced false positives and false negatives when applied to mammography screening, outperforming radiologists in several key metrics. The utility of AI extends beyond mammography. In ultrasound imaging, DL models are being developed to automate lesion characterization, reducing interobserver variability and potentially expediting diagnosis. In MRI, artificial intelligence techniques are used for automatic segmentation of the lesion, analysis of the kinetic curve, and evaluation of the background parenchymal enhancement. Such tools can improve diagnostic accuracy and efficiency, especially in complex or ambiguous cases.

## 2 AI Methodologies for Breast Cancer Diagnosis

The purpose of this review is to analyze and evaluate the main AI-based methodologies proposed by the scientific community to support breast cancer diagnosis. In particular, attention is paid to solutions based on Machine Learning techniques and those that make use of image processing; the intent is to compare these methodologies, the datasets used and the performances obtained, discussing among other things the relative advantages and disadvantages.

### 2.1 Models based on Machine Learning

In this section, we have chosen to analyze two case studies published in 2018 and 2021 that apply a supervised learning algorithm for the automated diagnosis of breast cancer, formulated as a binary classification problem between benign and malignant tumors. In the first case study, focuses on the analysis of the Support Vector Machine (SVM) trained and tested on different versions of Wisconsin Breast Cancer Dataset (WBCD), following two different pipelines based on

their proposed approach. The second case study try to employs K-Means for pre-clustering and PCA for dimensionality reduction, in order to support the application of SVM during the training phase, clearly using the same dataset. Our goal is to compare the final performance achieved by the classifier with the supervised and unsupervised approach, evaluated in terms of accuracy and sensitivity to the data.

**Diagnosis with supervised learning** Supervised Learning (SL) is one of the most widely used machine learning techniques. It consists of learning a function that maps input data to known labels, exploiting a labeled dataset. The goal is to build a classification model that can accurately predict outcomes for new data. The SL process includes two main phases: training and validation. During training, the model is optimized to minimize the error; in the validation phase, a separate subset of the data allows to evaluate the performance avoiding overfitting. However, SL requires large amounts of labeled data, often expensive to obtain, such as in the medical context. To address this limitation, the study “*Comparing Supervised and Semi-Supervised Machine Learning Models on Diagnosing Breast Cancer*” (Al-Azzam and Shatnawi 2021) compared SL and Semi-Supervised Learning (SSL) in breast cancer diagnosis, using the Wisconsin Diagnostic Breast Cancer (WDBC) dataset. WDBC includes 569 samples (357 benign and 212 malignant) where each of these contains a patient ID and 30 features extracted from digital images of fine needle aspirates (FNA) of breast masses. In the SL pipeline, 80 per cent of the data was used for training and 20 per cent for testing. This data was used to build the various classification models, followed by a hyperparameter tuning and validation phase to optimize their performance. The final model was then evaluated on the test set to assess its accuracy, precision, recall, and F1 score. In contrast, the SSL approach introduces a fundamental change: after splitting 80 per cent of the data for training, this subset was further split into 50 per cent labeled and 50 per cent unlabeled data. Initially, the various models were trained only on the labeled half. Subsequently, it was used to predict labels for the unlabeled data. These pseudo-labeled instances were combined with the original labeled set to retrain the model on an expanded dataset. The final model was then re-evaluated on the same test set. Specifically focusing on the Support Vector Machine (SVM) model, both linear and radial field (RBF) kernels were employed. In the SL setting, the SVM was trained on the fully labeled training data to find an optimal hyperplane that separated benign from malignant cases. In the SSL approach, the SVM was initially trained on only the labeled half of the training data, then used to predict labels for the unlabeled half, which were added back to retrain the model, improving its generalization ability. The results showed that SVM achieved high accuracy in both settings.

**Diagnosis with unsupervised learning** Unsupervised learning, which differs from supervised learning in the absence of labels in the data, finds application in breast cancer diagnosis mainly as a tool to support feature selection and reduction. In particular, techniques such as **K-Means** (for clustering) and **Principal Component Analysis (PCA)** (for

dimensionality reduction) are often employed to improve data quality before using supervised classification models.

A significant example of this approach is illustrated in the study “*Dimensionality Reduction using PCA and K-Means Clustering for Breast Cancer Prediction*” (Jamal et al. 2018) (2018), which uses the Wisconsin Breast Cancer Dataset (WBCD) to evaluate the combined effectiveness of PCA and K-Means. The purpose of this paper is to implement a prediction system based on a binary classification task to distinguish between benign and malignant breast cancer cases. In contrast to the first-mentioned method, which employs supervised learning without utilizing the actual WBCD dataset, the authors propose a different approach that integrates a data engineering phase prior to model training and testing. This phase involves the application of two unsupervised learning techniques like PCA and K-means, each applied separately to evaluate their effectiveness in terms of accuracy, specificity, and sensitivity, once the classification models have been trained and tested. The prediction system is structured around two main phases: training and testing. Both phases incorporate a data engineering sub-phase in which PCA and K-means are independently applied to generate refined datasets for model input. Regarding the two unsupervised learning methods, each produces a new dataset derived from the raw data:

- **PCA** is used to retain the most relevant features from the original dataset, functioning as a feature selection technique.
- **K-means**, on the other hand, identifies correlations among features in the original dataset and extracts new, relevant features by clustering similar features together.

Once the two improved datasets are obtained after the data engineering phase by PCA and K-means, the training and testing phase is carried out for each dataset, applying SVM on a split of 67-33% respectively of training and testing, measuring the final performances in terms of accuracy, sensitivity and specificity.

## 2.2 Model based on images

Among the various diagnostic techniques, ultrasound imaging is one of the most widely used for detecting breast abnormalities. However, the automatic processing of ultrasound images remains a significant challenge. This section compares two approaches: one based on traditional Computer Vision techniques (2020), and a more recent one based on Deep Learning networks (2024). The results are compared in the final section in terms of accuracy and correctly processed images.

**Diagnosis with Computer Vision** Ultrasound scans often contain speckle noise, which hinders the segmentation of Regions of Interest (ROIs). In the study “*Fully Automatic Segmentation of Gynecological Abnormality Using a New Viola-Jones Model*” (Hussein et al. 2020), a method is proposed that adapts the Viola-Jones algorithm, originally designed for facial recognition.

The process involves:

- A Wiener filter to reduce noise while preserving edges and fine details;

- Haar features to capture spatial relationships between regions;
- Weak classifiers in cascade to detect abnormalities;
- Active contour models to iteratively refine the ROI boundaries;
- Support Vector Machines (SVM) to classify the detected areas as benign or malignant.

This combination enables fully automatic segmentation with good accuracy, although it remains sensitive to image quality.

**Diagnosis with Deep Learning** Although traditional methods have improved the analysis of ultrasound images, they still suffer from issues such as false positives and the growing complexity of large datasets. In recent years, Deep Learning models have emerged as a valuable aid to radiologists in the diagnostic process.

The study “*Deep Learning in Breast Cancer Imaging: State of the Art and Recent Advancements in Early 2024*” (Carriero et al. 2024) outlines several approaches:

- GAN (Generative Adversarial Networks): generate synthetic images to expand datasets and simulate disease progression;
- CNN (Convolutional Neural Networks): identify tumor masses and estimate their severity;
- LLM (Large Language Models): integrate patient data and generate diagnostic reports.

CNNs appear to be the most promising. In “*Boosting Breast Cancer Detection Using CNN*” [3], three architectures are tested on a Kaggle dataset (162 H&E) containing 275,000 labeled medical image patches, classified as either benign or malignant. Each CNN operates across three levels:

- Convolution layer, to extract visual patterns;
- Pooling layer, to reduce dimensionality while retaining key features;
- Fully connected layer, for the final classification.

The study shows that, by combining well-designed CNNs with data augmentation techniques, it is possible to achieve higher accuracy than traditional methods.

Finally, newer architectures such as Vision Transformers (ViT) and ConvNeXt propose notable improvements: the former processes images globally, while the latter enhances CNNs with more advanced activation and normalization mechanisms.

## 3 Agent-based approach

Multi-agent systems have been widely used in the healthcare field, not only for decision-making purposes in patient monitoring and care, but also for medical diagnosis.

A multi-agent system is a distributed software system where data to be processed is decentralized between the different agents. These agents work independently from each other in order to find partial solutions, and then communicate together to aggregate them into a global solution for the original problem.

It is essential that only a few “trusted” agents are involved in

the working process of a multi-agent system. In this way, data cannot be altered by suspicious agents, which could lead to a wrong decision during the process or a violation of privacy, especially serious in the presence of sensitive information. A solution for this issue is represented by the integration of the blockchain technology to the multi-agent systems. This technology, based on cryptographic methods, introduces an authentication system for the agents, so that only the trusted ones can participate to the decision-making process of the multi-agent system and can communicate safely to each other.

In the following section, we focus on an example of blockchain integrated multi-agent learning system (BMAES), which is useful for our goal.

### 3.1 BMAES for breast cancer diagnosis

In the paper "*Blockchain integrated multi-agent system for breast cancer diagnosis*" (Ramanathan, Hossen, and Sayeed 2022), the authors propose an efficient BMAES that supports breast cancer diagnosis in a safe and efficient way. The architecture of this system is composed of five classification agents, an intermediate agent and a blockchain. The classification agents work independently to provide partial outputs, which are stored in the blockchain. In the end, the intermediate agent computes an average of these outputs, which corresponds to the final prediction for the diagnosis.

**Classification agents** The five classification agents use different machine learning algorithms (decision tree, KNN, SVM, random forest, Naïve Bayes), which are suitable for working on structured datasets to obtain a prediction of breast cancer disease. Since these algorithms classify data based on specific target classes of the training set, the relative classification agents are hybridized with three-layered parallel fuzzy systems. In fact, clinical data are often uncertain or incomplete, so by integrating these fuzzy systems, we can avoid the issue related to their management by machine learning algorithms.

Moreover, by using a parallel fuzzy logic approach rather than a single one fuzzy system, we can solve the issue of dimensionality problem related to the use of a single fuzzy system. Therefore, the output of a classification agent is represented by the average of the outputs computed by its sub-fuzzy systems.

The sub-fuzzy systems are defined starting from the training set, which is split into many subsets containing data values of at most three input attributes and the target attribute. So, the number of layers of the fuzzy system depends on the number of dataset groups which are defined from the training dataset.

In the paper under consideration, the authors use the WDBC dataset, which we have already introduced in the previous chapter. The feature selection process selects eight main features from the 569 samples of the dataset: radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry and fractal dimension of the cell nuclei. By applying the previous reasoning, three dataset groups are created, so a three-layered parallel fuzzy system is developed for each classification agent.

**Blockchain** As previously introduced, only some authenticated agents can join the decision-making process and consequently to store data on blockchain to make them not alterable.

The agents' authentication is based on some cryptography algorithm, such as the RSA method, which is an asymmetric cryptography algorithm used by the BMAES of our case study. According to this protocol, each agent has a private key and a public one. To make an agent authenticated with the blockchain, the security system encrypts a random number by using its public key, so that it can be decrypted only by this agent through its private key. If this happens, the agent can access the blockchain.

In addition, agents might have to satisfy some consensus algorithms to create new blocks of data on the blockchain. In the considered BAES architecture, two consensus algorithms are involved: the proof of IO limits and the proof of agent performance. The first checks whatever input and output values given by an agent are within the expected limits, while the second checks whatever accuracy of the agent's classification is above some minimum threshold, such as 70%. If the five classification agents pass these two tests, then five new blocks are created on the blockchain, where the results of the respective classifications can be stored.

## 4 Discussion

The analysis of the methodologies chosen in this review highlights how Machine Learning techniques, those based on image detection and the studied multi-agent system, have achieved significant results in tumor diagnosis despite their different applicability regarding their type of process and accuracy. The aim, therefore, is to critically compare each method analyzed in the previous sections by evaluating their performance, suitability to the available data and clinical applicability. The discussion was structured on three topic-specific levels: (1) comparison and analysis and (3) final considerations and integration.

### 4.1 Analysis and comparison of Machine Learning models

Both case studies described in section 2.1 used the Wisconsin Breast Cancer Dataset, but with radically different strategies:

- The first one relies on supervised and semi-supervised models (SVM with RBF and linear kernels), using labeled data for training and obtaining high accuracy and sensitivity, proving robust to overfitting. However, this approach is dependent on data labeling to obtain an accurate prediction, and such labeling will be difficult to obtain in a clinical context, in addition to its limited scalability with such data
- The second one applies unsupervised techniques (PCA and K-Means) to improve the data before the model training phase, showing how the engineering of the input data affects the performance of the final classifier, obtaining very similar results compared to the first case study.

By analyzing the two works in a comparative way, we can notice how the presence of PCA/K-Means does not re-

place the process of obtaining the final prediction with supervised learning, but rather, further enhances it a priori. It is also demonstrated that even if there is such an application on a small amount of data in the clinical context taken into consideration, it is still possible to obtain high accuracy and specificity on such data thanks to semi-supervised learning, thus reducing the dependence between the prediction and the data. Finally, we can notice how in both cases, the SVM model turns out to be a competitive classifier if inserted in well-designed pipelines.

#### 4.2 Analysis and comparison of image detection models

The two studies analyzed in section 2.2 have profoundly different approaches:

1. The first is based on classical Computer Vision techniques, adapting the Viola-Jones algorithm to automatically segment ROIs in ultrasound images. The pipeline includes a Wiener filter to reduce noise, Haar features to highlight spatial relationships, cascaded weak classifiers and an SVM classifier. Although the accuracy is high, this approach remains sensitive to image quality and is not applicable to large datasets.
2. The second method uses Deep Learning networks, including CNNs, GANs and LLMs, with better accuracy levels. CNNs, in particular, are trained on very large datasets thanks to data augmentation. Very promising results, moreover, come from more recent architectures such as Vision Transformers and ConvNeXt.

Comparing the two approaches, neural networks are the best solution, especially with large datasets. However, using CNNs still presents challenges: the decisions taken are not directly verifiable, making clinical validation difficult. Furthermore, maintaining such models requires expensive infrastructure and a lot of computing power, which is impossible for most institutions to obtain. Delegating processing to external servers solves the problem, but opens up privacy issues on patient data. In conclusion, although Deep Learning represents a promising evolution for automatic diagnosis, its integration into the clinical world opens up many challenges and discussions.

#### 4.3 Analysis of the multi-agent system

The BMAES we have seen is based on hybrid agents (ML+fuzzy logic), which are more suitable than classical ML classifiers for our goal. In fact, they can handle the ambiguity/uncertainty/news of data, which frequently occurs in the case of clinical data, so without strongly depending on the classes included in the training set.

By testing the proposed BMAES on the WDCB database, we note that with a system of this type, we can achieve high accuracy (96.49%) and high sensibility (86.67%), even if the latter is 6% lower than the best among those of the individual agents. This could be due to what type of aggregation the system uses. In particular, the proposed BMAES aggregates the results of the single agents by taking the average of them. This may reduce the total performance of the system since all agents have the same weight in the decision of the final

output, independently of their accuracy/sensibility. An idea to overcome this issue could be represented by assigning a fuzzy weight to each agent based on its performance. In this way, the more performing agents have an higher weight than the less ones, and the sensibility of the system improves.

Finally, the performance of a BMAES can also be evaluated by the computational time needed to obtain the final output. From this point of view, the performance is affected by the presence of the blockchain, as it requires a long computational time for the agent authentication process.

#### 4.4 Conclusion and future developments

The results of this review demonstrate the ability of a multi-agent system to exploit the strengths of different agents, in order to get an accurate decision for the breast cancer diagnosis. In particular, hybrid agents are the most suitable for clinical purposes as they can handle ambiguity, uncertainty, and news of clinical data. Moreover, the integration of a multi-agent system with the blockchain technology leads to an improvement of the breast cancer diagnosis, guaranteeing at the same time the data security and the integrity of the decision-making process.

For all these reasons, the BMAES approach represents a suitable solution not only for diagnosis, but also for other healthcare applications. For example, in telemedicine, it can support remote patient monitoring and consequently the identification of the most critical cases to be treated in the specialized facilities. In this way, the workload of the healthcare personnel and the waste of useful resources/spaces can be reduced.

Currently, the challenge lies in understanding how to integrate agents with different architectures within the same multi-agent framework. To do this, suitable protocols for communication between agents must be defined, while ensuring high accuracy of results and data privacy. Finally, it is essential that the models are not defined from only a few databases, since they must be as generic as possible to represent several populations with different characteristics.

#### References

- Al-Azzam, N.; and Shatnawi, I. 2021. Comparing Supervised and Semi-Supervised Machine Learning Models on Diagnosing Breast Cancer. *Annals of Medicine and Surgery*, 62: 53–64.
- Carriero, A.; Groenhoff, L.; Vologina, E.; Basile, P.; and Albera, M. 2024. Deep Learning in Breast Cancer Imaging: State of the Art and Recent Advancements in Early 2024. *Diagnostics*, 14(8): 848.
- Hussein, I. J.; Burhanuddin, M. A.; Mohammed, M. A.; El-hoseny, M.; Garcia-Zapirain, B.; and Maashi, M. S. 2020. Fully Automatic Segmentation of Gynaecological Abnormality Using a New Viola–Jones Model. *Computers, Materials & Continua*, 66(3): 2611–2627.
- Jamal, A.; Handayani, A.; Septiandri, A. A.; Ripmiatin, E.; and Effendi, Y. 2018. Dimensionality Reduction Using PCA and K-Means Clustering for Breast Cancer Prediction. *Lontar Komputer: Jurnal Ilmiah Teknologi Informatika*, 9(2): 98–106.

Mann, R. e. a. 2019. Breast MRI: state of the art. *Radiology*.

McKinney, S. e. a. 2020. International evaluation of an AI system for breast cancer screening.

Organization, W. H. 2023. Global cancer statistics 2022.

Ramanathan, T.; Hossen, M.; and Sayeed, M. 2022. Blockchain Integrated Multi-Agent System for Breast Cancer Diagnosis. *Indonesian Journal of Electrical Engineering and Computer Science (IJECS)*, 26(2): 998–1008.