

# Detection of Breast Cancer Using Artificial Intelligence Algorithm: SVM Review

Abhirami Natarajan Chettiar, Saniya Sunil Gawas

Student, Department Of Information Technology, B. K. Birla College Of Arts, Science & Commerce (Autonomous), Kalyan, Maharashtra, India.

Student, Department Of Information Technology, B. K. Birla College Of Arts, Science & Commerce (Autonomous), Kalyan, Maharashtra, India.

Submitted: 10-08-2021

Revised: 25-08-2021

Accepted: 29-08-2021

## ABSTRACT

Breast cancer is one of the major types of cancer affecting the people from at the most every corner of the world. A survey 40000 woman loses their life because of breast cancer which means in each 13 minutes a woman dies. Screening mammography techniques helps in diagnosing a cancer and if the cancer is detected at a very early stage the rate of recovery will be faster. There are others methods like biopsy, blood analysis, MRI scans, etc which can be used to detect this type of cancer but the most commonly used one is mammography. The disadvantage of the screening mammography that is the traditional mammography is that there are high chances of false analysis. The first chance is of getting a false negative means the victim actually has a breast cancer but the report denies it and the second chance is of getting a false positive result means an unaffected person is tested positive for breast cancer. And it also takes a long time for the radiologists to conclude with the report. To overcome this artificial intelligence is introduced in

mammography. In these algorithms like KNN, Decision Tree Algorithm, Random Forest Classification, Logistic Regression, etc are used and the machine is also trained with many images of mammography with all the possible critical situations. In this abstract, the algorithm used is SVM – support vector machine classifier. In this method first the data is collected from various sources and the processed. Because of inclusion of artificial Intelligence in mammography the accuracy of the results has been increased. This used algorithm is 98% accurate and the false reports have been reduced almost up to 10%. Now a day's risk of the tumor can also be predicted by using artificial intelligence. Another advantage is that with the help of this technique the time gap between the test and result has been reduced due to which the patients can be given treatment from the earliest as possible which will increase their life span.

**KEYWORDS:** Breast Cancer, Mammography, Artificial Intelligence, SVM

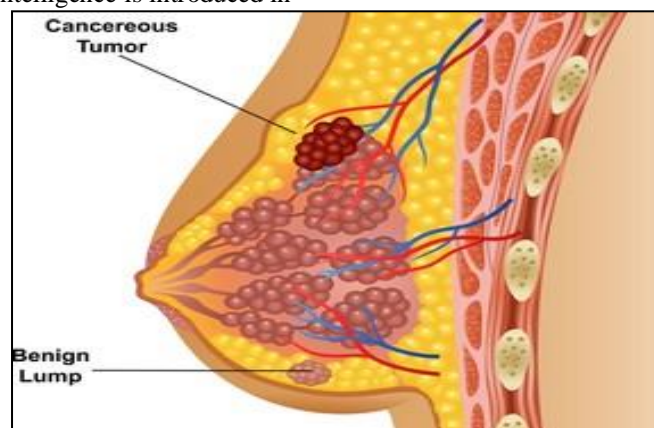


FIGURE 1[9]: Breast Infected With Cancerous Tumor

## I. INTRODUCTION

One of the deadly diseases among humans is Cancer. Cancer starts when cells begin to grow out of control. [1] Breast cancer is the most typical cancer among women but it is also seen in few men. In 2020 there were 2.3 million women diagnosed with breast cancer and 685000 deaths globally. As of the end of 2020 there were 7.8 million women alive who were diagnosed with breast cancer in the past 5 years, making it the world's most prevalent breast cancer. Most breast cancers begin within the ducts that carry milk to the nipple i.e. ductal cancers. Some begin within the glands that build breast milk i.e. lobular cancers. Different kinds of breast cancers are Ductal Carcinoma in-situ (DCIS), Invasive carcinoma, Phyllodes tumor and Angiosarcoma. Breast cancer can spread when the cancer cells get into the blood or lymph system and are carried to different elements of the body.

There are five stages of breast cancer they are stage 0, stage 1, stage 2, stage 3, stage 4, and stage 5. In stage 0 the cancer has been diagnosed early. It started within the breast ducts or milk

glands and has stayed there. In stage 1 Starting at this level, breast cancer is called invasive, meaning it has broken free to attack healthy tissue. In stage 2 the cancer has grown, spread, or both. In stage 3 the cancer has not spread to bones or organs, but it's considered advanced, and it's harder to fight. In stage 4 Breast cancer cells have spread far away from the breast and lymph nodes right around it. The foremost sites are the bones, lungs, liver, and brain. This stage is represented as "metastatic," meaning it has spread beyond the region of the body where it was first found. The main aim of the paper is how breast cancer is detected using artificial intelligence algorithm i.e. SVM (Support Vector Machine).

A Support Vector Machine (SVM) is a binary linear classification whose decision boundary is explicitly constructed to minimize generalization error. It is a very powerful and versatile Machine Learning model, capable of performing linear or nonlinear classification, regression and even outlier detection. SVM is well suited for classification of complex but small or medium sized datasets.

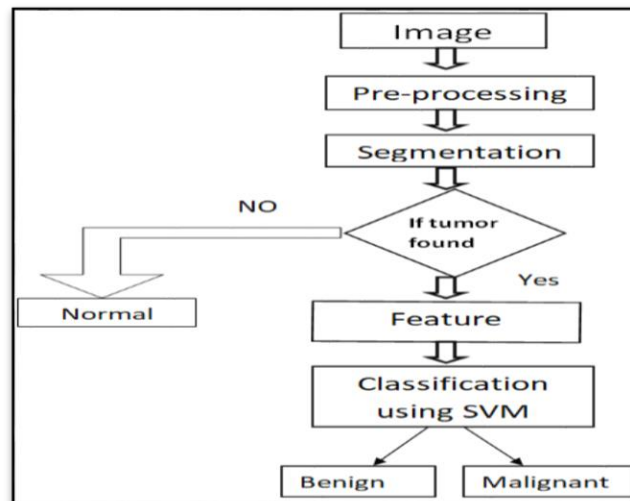


Figure 2: Flow chart of SVM algorithm

## STEPS FOR DETECTING BREAST CANCER

### I. PRE-PROCESSING

In this process data and mammogram of various infected and non infected persons are collected by various methods. In this step the data cleaning is done in which either the missing values of a person are filled or the whole details of the person are removed so that there is no inconsistency while training the model. The mammogram is then modified by image enhancement techniques which helps in not only removing noise from the collected data but also to improve accuracy of segmentation results. The

medical Images are pre-processed that ensures a proper diagnosis.[3]It also improves accuracy of the segmentation results. It reduces the complexity involved in color processing, resizing of image to the standard size and removing noise from image. [3]

### II. SEGMENTATION

Segmentation is an essential stage of the image recognition system, because it collects the objects of our interest, for further processing such as description or recognition. The pre-processed images are converted to binary images through thresholding at different values. The images are

filtered again with Gaussian smoothing filter to eliminate noise. [4] Morphological technique is used to select the region of interest. [3]The thresholding is an important step to improve the detection of breast cancer segmentation subdivides an image into its constituent regions. This involves two levels, first the normal image is dilated and it is separated from the normal image and the resultant image is the affected area.

### III. FEATURE EXTRACTION

Feature extraction is a part of the dimensionality reduction process, in which, an initial set of the knowledge is split and decreased to lot of teams. The most important characteristic of these big data sets is that they have a variety of variables. The feature extraction is used to measure the properties from the segmented image are, Area, Centroid, major axis length, minor axis length, eccentricity, orientation, filled area, extrema, solidity, equidiameter.[4] The area is the scalar value; it computes the particular range of pixels within the region. Then the centroid is the vector and it computes the centre of the tumor region [4]. Feature extraction helps to decrease the quantity of redundant data from the information set. In the end, the reduction of the information helps to build the model with less machine's efforts and also increase the speed of learning and generalization steps in the machine.

### IV. SVM CLASSIFICATION

A Support Vector Machine is a supervised algorithm which will classify cases by finding a separator. SVM works by mapping information to a high-dimensional feature area so that information can be sorted, even when the information is not

otherwise linearly separable. Then, a separator is estimated for the data. The data should be transformed in such a way that a separator could be drawn as a hyperplane. Basically, mapping data into a higher dimensional space is called kernelling. The mathematical function used for the transformation is termed as the kernel function, and may be of various sorts, such as: Linear, Polynomial, Radial basis function, and Sigmoid. Every of those functions have its own characteristics, its pros and cons, and its equation, but the good news is that you don't need to know them, as most of them are already implemented in libraries of data science programming languages.

The hyperplane is learned from training data using an optimization procedure that maximizes the margin; and like many other problems, this improvement drawback can also be resolved by gradient descent. You can make classifications using this estimated line. It is enough to connect input values into the line equation, then, you can calculate whether an unknown point is on top or below the line. If the equation returns a value greater than 0, then the point belongs to the first class, which is on top of the line, and vice versa. However, primarily, it is used for Classification problems in Machine Learning. The goal of the SVM algorithm is to form the most effective line or decision boundary that may segregate n-dimensional space into categories so that we can easily put the new information point in the correct category in the future. This best decision boundary is termed as hyperplane. These extreme points/vectors that help in creating the hyperplane. These extreme cases are called as support vectors, and hence algorithm is called as Support Vector Machine (SVM).

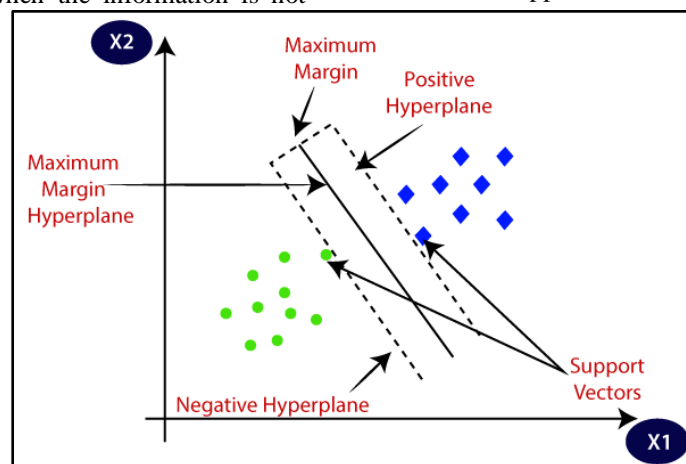


Figure 3[10]: Graphical Representation of SVM Classifier

## V. CONCLUSION

Artificial intelligence has provided the medical sector tools for analyzing the data related to diseases. Thus machine learning algorithms play an important role in early detection of diseases. This paper provides a review on how machine learning algorithm i.e. Support Vector Machine (SVM) is used for detecting breast cancer. Early detection of breast cancer can reduce the death rates of women on a very large scale leading to a good life.

To conclude, the algorithm's accuracy depends on the size of the dataset and the number of features, and the results of the ensemble model are better than using only a single model. In future the research can be carried out to improve the detection

- [1]. Kumar, M. V., Ram, S. A., Nageswari, C. S., Raveena, C., & Rajan, S. (2021b). Early-Stage Detection of Cancer in Breast Using Artificial Intelligence. *Revista Gestão Inovação e Tecnologias*, 11(2), 2016–2028. <https://doi.org/10.47059/revistageintec.v11i2.1821>
- [2]. Huang M-W, Chen C-W, Lin W-C, Ke S-W, Tsai C-F (2017) SVM and SVM Ensembles in Breast Cancer Prediction. *PLoS ONE* 12(1): e0161501. <https://doi.org/10.1371/journal.pone.0161501>
- [3]. Breast Cancer Classification using SVM Classifier. (2019). *International Journal of Recent Technology and Engineering*, 8(4), 527–529. <https://doi.org/10.35940/ijrte.d7227.118419>
- [4]. Y.Ireaneus Anna Rejani+, Dr.S.Thamarai Selvi, EARLY DETECTION OF BREAST CANCER USING SVM CLASSIFIER TECHNIQUE. *International Journal on Computer Science and Engineering* Vol.1 (3), 2009, 127-13.
- [5]. Abdollahi, J., Keshandehghan, A., Gardaneh, M., Panahi, Y., & Gardaneh, M. (2020). Accurate Detection of Breast Cancer Metastasis Using a Hybrid Model of Artificial Intelligence Algorithm. *Archives of Breast Cancer*, 18–24. <https://doi.org/10.32768/abc.20207118-24>
- [6]. Yoon, J. H., & Kim, E. K. (2021). Deep Learning-Based Artificial Intelligence for Mammography. *Korean Journal of Radiology*, 22(8), 1225. <https://doi.org/10.3348/kjr.2020.1210>
- [7]. Ibrahim, I., & Abdulazeez, A. (2021). The Role of Machine Learning Algorithms for Diagnosing Diseases. *Journal of Applied*

techniques using artificial intelligence field algorithms. Early breast cancer detection can dramatically save lives particularly within the developing world. This technique will improve by combining computer vision and machine learning and deep learning techniques to classify cancer directly using tissue images.

**ACKNOWLEDGEMENT:**I would like to thank the Department of Information Technology, B. K. Birla College of Arts, Science & Commerce (Autonomous), Kalyan and Professor Swapna Nikale for their support that made this research possible.

## REFERENCES:

- Science and Technology Trends, 2(01), 10–19. <https://doi.org/10.38094/jastt20179>.
- [8]. Frazer, H. M., Qin, A. K., Pan, H., & Brotchie, P. (2021). Evaluation of deep learning-based artificial intelligence techniques for breast cancer detection on mammograms: Results from a retrospective study using a BreastScreen Victoria dataset. *Journal of Medical Imaging and Radiation Oncology*, 65(5), 529–537. <https://doi.org/10.1111/1754-9485.13278>
- [9]. [https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.mountelizabeth.com.sg%2Ffacilities-services%2Fcentre-excellence%2Fcancer%2Fbreast-cancer&psig=AOvVaw0AcTiy5bVZnLRCVg8Xmb8T&ust=1629972352978000&source=images&cd=vfe&ved=0CAsQjRxqFwoTCNixjK\\_2y\\_ICFQAAAAAdAAAAABAF](https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.mountelizabeth.com.sg%2Ffacilities-services%2Fcentre-excellence%2Fcancer%2Fbreast-cancer&psig=AOvVaw0AcTiy5bVZnLRCVg8Xmb8T&ust=1629972352978000&source=images&cd=vfe&ved=0CAsQjRxqFwoTCNixjK_2y_ICFQAAAAAdAAAAABAF)
- [10]. [https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.javatpoint.com%2Fmachine-learning-support-vector-machine-algorithm&psig=AOvVaw2gVtIa-cJGDB6Boebf\\_aKY&ust=1629972177745000&source=images&cd=vfe&ved=0CAsQjRxqFwoTCMCC2db1y\\_ICFQAAAAAdAAAAABAD](https://www.google.com/url?sa=i&url=https%3A%2F%2Fwww.javatpoint.com%2Fmachine-learning-support-vector-machine-algorithm&psig=AOvVaw2gVtIa-cJGDB6Boebf_aKY&ust=1629972177745000&source=images&cd=vfe&ved=0CAsQjRxqFwoTCMCC2db1y_ICFQAAAAAdAAAAABAD)