



Enhancing Breast Cancer Diagnosis Through Virtual Biopsy and Machine Learning

Muhammad Toseef

Faculty of Computer Science & Information Technology, The Superior University
Lahore, Punjab Pakistan. Email: mtoseef6577@gmail.com

Ahmad Khan (Corresponding Author)

Faculty of Computer Science & Information Technology, The Superior University
Lahore, Punjab, Pakistan. Email: ahmad.khan.fsd@superior.edu.pk

Khalil Ullah

Department of Software Engineering, University of Malakand, Chakdara Dir Lower,
Khyber Pakhtunkhwa Pakistan. Email: khalil.ullah@uom.edu.pk

Shahbaz Ahmad

Faculty of Computer Science & Information Technology, The Superior University
Lahore, Punjab Pakistan. Email: shahbaz.ahmad.fsd@superior.edu.pk

Ibrar Hussain

QEC, Shaheed Benazir Bhutto University, Sheringal 18050, Pakistan.
Email: ibrar@sbbu.edu.pk

Said Khalid Shah

Institute of CS and IT, University of Science and Technology Bannu
Email: Dr.khalidshah@ustb.edu.pk

Abstract

Breast cancer continues to be a leading global health concern, and early detection is the key to optimizing prognosis and survival. Traditional diagnostic approaches have consisted of mammograms, ultrasound, and invasive biopsies — which have been effective in detecting tumors but has its drawbacks. Mammography frequently misses tumors in women with dense breast tissue, and biopsies are costly, invasive and time-consuming. And they also inflict pain and suffering on the patients. There is a need for less invasive and more effective diagnostic alternatives to such facts. This latest technique gives hope as non-invasive diagnostic method including preparation of virtual biopsy. Virtual biopsy refers to using advanced imaging technologies such as MRI, ultrasound, and mammography to see the properties of the breast tissue without removing it. This would enable the improved and more efficient identification of tumors, enabling less invasive biopsies to be performed, and subsequently less discomfort for patients. The main objective of this paper is to represent virtual biopsy with machine learning algorithms to enhance early breast cancer detection which includes SVM, Decision Tree and CNN. Machine learning algorithms proved to be very



successful in medical image analysis. SVM can solve complex problems with high accuracy by deciding his best decision boundaries, Decision Tree provides overfitting, and CNN automatically finds features from images, which is very suitable for detecting subtle features in medical images. The results of this study show that the combination of virtual biopsy with machine learning significantly improves diagnostic accuracy and efficiency. This will allow the detection to happen much sooner, with fewer repeat tests, and will be much more reliable, non-invasive, than traditional diagnostic methods. An accurate alternative to invasive techniques, virtual biopsy, combined with machine learning promises to enhance the early detection of breast cancer, allowing for better patient outcomes while easing the burden on the healthcare system.

Keywords: Breast cancer, early detection, virtual biopsy, machine learning, Support Vector Machines, Decision Tree, Convolutional Neural Networks

Introduction

This The most common malignancy among females is breast cancer with over 2.3 million new cases occurring yearly, accounting for almost 11.7% of cancers diagnosed [1]. One of the leading causes of cancer-related deaths, its survival rate varies by region due to the differences non uniform regarding early detection and accessibility to treatment. Improvement of prognosis depends directly on early detection as this avenue provides the basis for a timely intervention that might inhibit the progression of cancer to advanced stages.

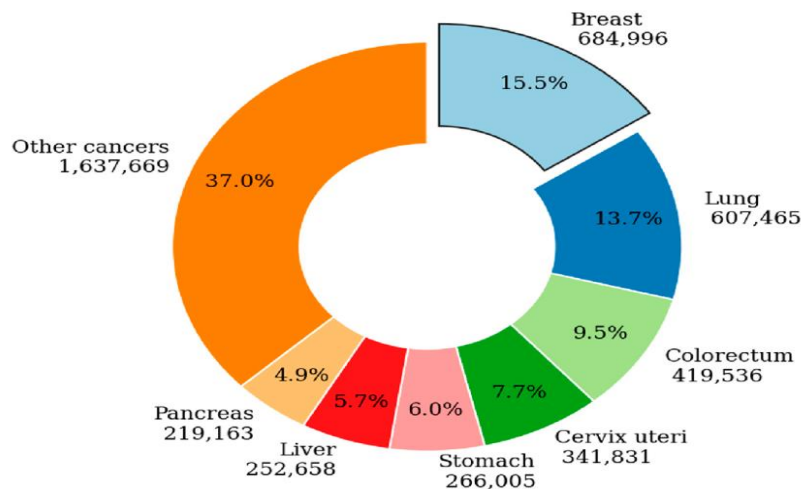


Figure 1.1: Global Cancer Mortality Rate[1]

In short, with the identification of early breast cancer, there is a possibility for an effective treatment and survival rate, and in this connection, early diagnosis stands at the very core of managing breast cancer.



Mammography and biopsy have traditionally been the methods to diagnose breast cancer for a long period. However, such methods remain limited by the precision of the diagnosis and patients' comfort. For instance, mammography is among the most common screening techniques applied in breast cancer. Unfortunately, it has lost much sensitivity in women whose breasts are dense, resulting in missed diagnoses or follow-up tests that turn positive despite being false [2][3]. Similarly, ultrasound, although an effective adjunct tool, at times produces false positives that may in turn cause unneeded biopsies and a great deal of anxiety in patients [4]. Highly accurate though invasive biopsies may be, there is a risk of complications, and the procedure is costly and discomforting to the patient, hence not suitable for screening purposes.

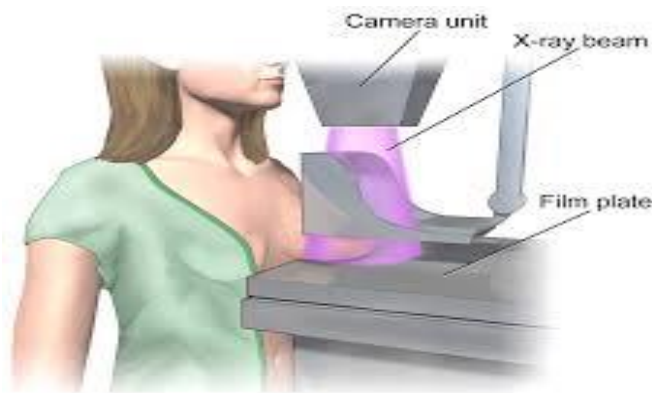


Figure 1.2: Mammogram [4]

These limitations call for more advanced diagnostic techniques that are non-invasive and bring faster and accurate results without some of the drawbacks associated with existing methods. Virtual biopsy, powered by machine learning algorithms, addresses these challenges. Virtual biopsies can analyze imaging data at much higher resolution and better accuracy than traditional virtual biopsy techniques. Using machine learning algorithms, such as CNNs, automatically finds patterns in imaging data that can distinguish malignant from benign lesions more effectively than the traditional methods. Such a high precision not only improves early detection but also reduces the number of unnecessary biopsies, thus helping to ease patient anxiety and reduce the healthcare costs.

The potential of virtual biopsy is that it can augment the accuracy of diagnosis and is very important for breast cancer in its early-stage diagnosis [5]. Early detection significantly boosts the chances for successful treatments and improves outcomes for the patients. Such enablement of virtual biopsy with enhanced accuracy and earlier detection helps minimize the burden of breast cancer globally. This ultimately can lead to more specific treatments that cater to every individual's needs. Besides that, these advanced techniques would complement existing methods of diagnoses with a more holistic approach of breast cancer diagnosis - which combines the strengths of the traditional methods with precision, utilizing the machine learning approach.

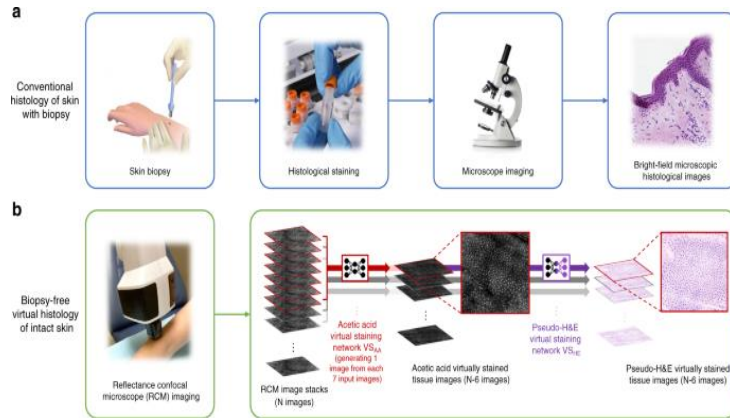


Figure 1.3: Conventional biopsy and virtual biopsy [5]

It integrates virtual biopsy and machine learning to revolutionize the diagnosis of breast cancer. It brings the opportunity to make non-invasive, more rapid, and more accurate diagnostics than with conventional methods possible. Thus, not only may the traditional methods decrease patient comfort and risk associated with these, but improved results are likely as a result. Further, as the research in this field progresses, the clinical application of machine learning-powered virtual biopsy techniques will have to be a crucial part in changing the current breast cancer screening and diagnosis.

Identifying the potential of virtual biopsy as a breast cancer diagnostic can prove out to be a paradigm shift and with tremendous advantages in comparison to the traditional diagnostics methods. Virtual biopsy employs cutting-edge imaging techniques like MRI, mammography, and ultrasound to provide advanced technical information on the detail structure of the breast tissue which allows to know its properties without sample extraction [6]. This is a non-invasive procedure that bypasses the uncomfortable, costly and risk-prone proceedings inherent to standard biopsies. Imaging data obtained through virtual biopsy provides a comprehensive assessment of breast tissue with a certain accuracy and precision in its detection and characterization.

The clear advantage of virtual biopsy is that it is non-invasive. The standard biopsies usually entail the removal and physical analysis of tissue, but virtual biopsy utilizes imaging techniques that capture the inner structure of the breast at very fine detail. Hence, the approach is accurate in diagnosing the problem and reduces anxiety and pain associated with an invasive approach. This removes one of the more back and forth potential reasons for unnecessary biopsies and makes the process less strenuous and faster for the person.

Virtual biopsy is also an important technique because it is time-saving technique. With a traditional biopsy process, it is always out of the question that a series of scheduling steps to prepare a biopsy and analyze it in a lab is required in order to produce results (which may take days or even weeks) before a diagnosis can be made. Virtual biopsy techniques can immediately indicate the nature of the tumor and enable immediate decision-making and intervention involving the patient. Speed is crucial in the



diagnosis of breast cancer, because the earlier a cancer is detected, the better for treatment results and the survival rate for patients.

The second added value of virtual biopsy is that it makes use of machine learning algorithms in the computer vision domain that is proven to have good performances in medical image analysis. These algorithms are excellent at interpreting images accurately and are able to extricate complex trends from imaging data that a human radiologist can easily overlook. It enables them to classify that a tumor is benign or malignant and to predict its performance and suggest ranking of the cases for further investigation. Machine learning which is used to improve virtual biopsy could allow for more sophisticated analysis of the imaged tissue for better diagnoses than traditional imaging techniques.

In a nutshell, virtual biopsy is a potential substitute for traditional diagnostic techniques. It provides minimizes invasive processes, quicker and more accurate assessments for breast cancer which is an ideal approach for screening and enhancing patient outcomes as well [7][8]. Increased research in the field would welcome these advanced imaging technologies and machine learning as the cornerstone of a new breast cancer landscape. This might eventually translate to improved screening protocols, customized treatment strategies and even lower rates of death related to breast cancer.

Machine learning augments the capabilities of diagnosis in virtual biopsy. The algorithms of machine learning can comb through vast amounts of imaging data for subtle patterns that indicate malignancy. Advanced datasets being processed and enhancing accuracy in diagnosis are revolutionizing medical imaging [9].

This paper explores how ML algorithms, such as SVM, Decision Tree, and CNN, can be applied in virtual biopsy to aid in the diagnosis of breast cancer. We want to be more accurate, cost-efficient, and patient-friendly for a diagnostic solution by combining their strengths.

Literature Review

Since its history dates from several decades back, advancements have been made progressively in detection ways leading to an efficient and enhanced accuracy in methods of determining breast cancer cases. General diagnosis methods had been by physical examination mammography among which biopsy has mainly been established for detecting most cases. Though the above methods have been shown effective in many instances, their applicability to detect early-stage cancer cases is less accurate. Also, the procedures tend to be painful for most patients. All these led the quest for other more refined and sophisticated techniques to help diagnose breast cancer. Research in this field has recently focused more on ML as it has been effectively proven to work well on complicated medical data, enabling accurate diagnoses and simplification of diagnostic procedures.

Machine learning algorithms, in particular, have been promising in solving the challenges posed by traditional methods. One of the most widely used machine learning algorithms in medical diagnostics is Support Vector Machine, or SVM. SVM is a supervised learning algorithm that is very good at classification tasks, finding a hyperplane that optimally separates different classes of data. SVM has been highly



applied in the classification of breast tissue samples from medical images with very high accuracy in the diagnosis of breast cancer. The algorithm works well with high-dimensional data and therefore is a good tool to identify malignant and benign tissues [10][11].

Decision Tree is another very popular algorithm used in the diagnosis of breast cancer. RF is considered one of the ensembles learning methods, with several decision trees built into the method to improve their output. Therefore, the right kind of classification could be strengthened as a result. There are mainly two significant benefits of this algorithm: less chance of overfitting and being especially very beneficial in clinical applications. Many studies are confirmed to illustrate the effectiveness of the Decision Tree in classifying tissues. This is because the process is able to process complex data and, of course, guarantee reliable results under many scenarios [3][7]. Results are accumulated from multiple decision trees of which Decision Tree ensures the results of the final diagnosis should not be influenced much by anomalies in individual data points, thus contributing towards highly accurate and consistent predictions.

Along with SVM and Decision Tree, CNN-based models have found their attention because they were capable of processing medical images and analyzing them. Basically, CNNs are some sort of deep learning that enables the algorithm to auto-learn hierarchical features on the input images. Consequently, they are highly compatible for tasks like the early detection of breast cancer. The feature extraction capability of CNNs has been found to be particularly useful in extracting subtle features directly from medical imaging data without manual feature engineering, thus providing remarkable results in diagnostic accuracy. According to several studies, CNN achieves better results in classifying tissue samples extracted from the breast as benign or malignant, impressively high results for imaging modalities such as mammograms, ultrasounds, and MRI scans, according to [32],[27] and [28] As they are trained on huge amounts of data, CNNs can find tiny detail patterns and anomalies that a human eye may not be able to appreciate, thus contributing to better early detection rates.

With increasing studies investigating machine learning approaches for breast cancer diagnosis, it has been shown that such sophisticated algorithms can significantly benefit clinical settings. However, by applying machine learning models, it is possible to automate and enhance accuracy in the diagnosis and overcome the limitations of traditional methods. Additionally, these algorithms may support earlier detection of cancer than would normally occur and improve patient prognosis. In conclusion, the evaluation of mammograms and breast ultrasound using different machine learning techniques enable breast cancer detection and diagnosis with few clinical choices, low cost, and accuracy which will be a good choice for both patients and doctors in the future. The combination of machine learning with traditional imaging techniques makes it a potential violence frontier on the battlefield against breast cancer, offering hope for the further development of diagnostic tools with more optimal patient outcomes.

The new tool has also shown tremendous potential in the diagnosis of breast cancer by means of contemporary advanced imaging techniques like MRI, mammography, and ultrasound. The most important advantage of such imaging techniques over classical biopsy procedures is their ability to allow non-invasive detection and characterization of



breast tumors without requiring the sampling of any tissue. This enables researchers to have amazing improvements on how the tumors classified as benign or malignant when applying machine learning algorithms on these images, which became an effective and accurate diagnostic tool [32] with mammogram images, and found improved classification accuracy over standard methods, affirming the virtual biopsy potential in breast cancer detection.

Ultrasound images have also been scoured for breast cancer diagnosis using machine learning algorithms. As reported by [27] and [16] have been used to inform the virtual biopsy integration as part of a vision to future advancements for improving machine learning with ultrasound imaging. This hybrid has demonstrated a dramatically improved diagnostic accuracy compared to traditional methods in which cancer characteristics are more accurately evaluated. This is immensely valuable as it decreases the pain and risks associated with traditional invasive biopsy methods and therefore is promising for the identification of malaria at an early stage.

However, these studies promise much in improving the accuracy of breast cancer diagnosis with machine learning algorithms. Much further research is needed to integrate them with virtual biopsy techniques. This is one area that demands further exploration in validating the findings with larger, diverse datasets to ascertain the generalizability and robustness of the models across different populations and clinical settings. The clinical applications of virtual biopsy combined with machine learning need to be explored for real-world diagnostic workflows in terms of enhancing efficiency, reducing costs, and improving patient outcomes.

Additionally, standard protocols for applying the machine learning algorithms in the clinic will be necessary to ensure uniform and efficient use. Further development of these machine learning techniques will provide virtual biopsy with a good prospect of becoming an even more effective auxiliary method in traditional diagnostics to make diagnosis of breast cancer as individualized and accurate as possible. Virtual biopsy combined with machine learning may lead to speedier and more accurate detection and classification of tumors for better diagnosis, early intervention, and treatment in breast cancer patients. That is the promise of change in breast cancer diagnosis in the near future.

Methodology

Using publicly available breast cancer datasets and advanced imaging techniques, the study presents an exploratory analysis of virtual biopsy image. By applying different state-of-the-art machine learning algorithms, the study takes a step towards giving causes of breast cancer using effective models. Research as we know it has a systematic structure; that is, building from data collection to large datasets, to medical image data such as mammography images, ultrasound photographs, and MRI scans. It performs noise reduction, image normalization and segmentation through high-level preprocessing of these images to maintain quality and consistency of data. Last, you train deep learning, machine learning models using convolutional neural networks (CNN) on the preprocessed data to obtain relevant features from the visual images. To ensure that at the end of the day robust and reliable results are provided, the output of



the models is also evaluated widely with different metrics such as accuracy, precision, recall, and F1 score. Finally, analysis interprets the outputs of these models, to explain patterns that may potentially result in enhanced clinical decision-making and advanced breast cancer detection. It highlights the potential of imaging and machine learning together, which could enhance personalized treatment and diagnostic precision³²⁴.

Data Collection Techniques

Two data sets Wisconsin Breast Cancer Dataset (WBCD) and Breast Cancer Digital Database (BCDD) are used in the study. The WBCD dataset consists of features extracted from breast cancer images and contains labeled samples of both malignant and benign tumors [29]. On the other hand, BCDD database contains mammograms of breast tissues being a diverse and valuable source for training and validating machine learning frameworks [33]. They are also Virtual Biopsy Techniques like mammography, ultrasound, MRI. These methods use thousands of images of breast tissue for analysis. Virtual biopsy allows tumor detection without tissular removal, which is the main advantage of using this technology. Mammography uses a low-dose X-ray system for the examination of breast tissue and is commonly employed to collect data for the early detection of breast cancer [34]. Ultrasound Employs high-frequency sound waves to generate images of breast tissue, commonly used adjunct to mammography [30] (American Institute of Ultrasound in Medicine, 2019). Magnetic Resonance Imaging (MRI) uses strong magnets and radio waves to generate detailed images of the breast, used most in patients considered at high risk for breast cancer [34].

Analytical Methods

The research uses a variety of machine learning algorithms for classification of breast tissue samples into malignant or benign. For each of these algorithms, the strengths in dealing with complex medical data are unique. SVM finds an optimal hyperplane separating different classes of data, hence making it highly effective for classification tasks in medical diagnostics [15][24]. With its capacity to handle high-dimensional data well, such as those used in describing the features of breast cancer, SVM will be able to classify the samples as benign or malignant very accurately. Decision Tree creates many decision trees, with each decision tree making a different independent prediction. It is then possible to combine these results into improving classification accuracy while also avoiding overfitting significantly [12][22]. This ensemble method would be best suited for big data sets with complex interrelations in medical applications. Convolutional Neural Networks, a deep learning method, is particularly useful in this research since it automatically learns hierarchical features from input images, reducing the need for manual feature extraction [20][18]. It allows processing raw image data directly thus could unveil a pattern in the imaging, such as mammograms and biopsy images thus making it an integral part of the breast cancer detection process. These algorithms band together to make the system as comprehensive and accurate as possible for the early detection of breast cancer, with significantly better sensitivity and specificity as well. Additionally, it allows for cross-validation from the combination of different techniques



can be leveraged to validate that findings are reliable and generalizable across datasets [14].

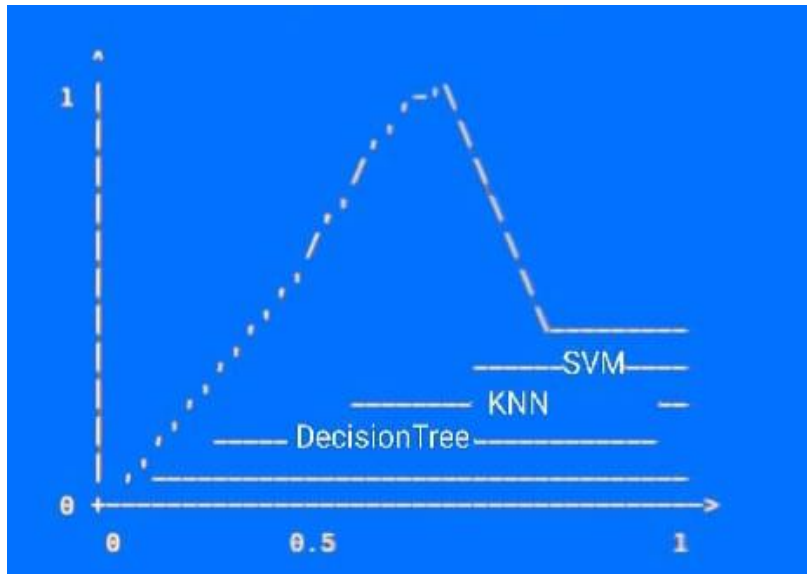


Figure 1: ROC Curves for SVM, KNN and Decision Tree.

Table 1. Performance metrics of machine learning models

Model	Accuracy	Macro Precision	Macro Recall	Macro F1 Score
DecisionTree	93	0.926	0.924	0.925
KNN	97.2	0.963	0.9775	0.9695
SVM	97.5	0.9715	0.9755	0.973

Results and Discussions

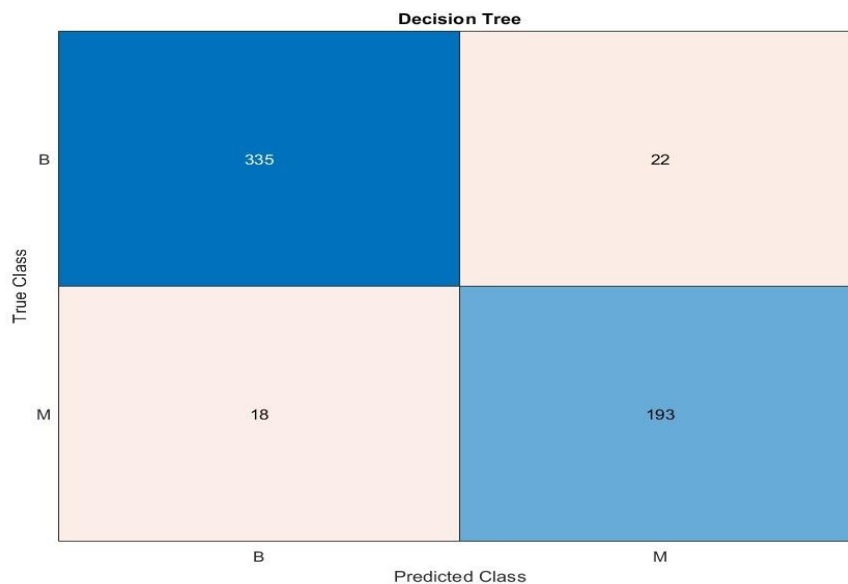
According to the research, the prediction outcomes show that the SVM model possesses higher classification accuracy and F1-score than Decision Tree and KNN algorithms. Table 1, presenting a close comparison between Decision Tree, KNN and SVM. The performance comparison of three models, namely Decision Tree, KNN, and SVM, shows differences between their efficiencies. By testing the Decision Tree model, the authors reported an accuracy of 93 %, a Macro Precision of 0.926, a Macro Recall of 0.924, and a Macro F1 Score of 0.925. Relative to these measures, the Decision Tree yields relatively low performance compared to KNN and SVM along misclassified instances. Prime among them, the KNN model showed a better result with an accuracy of 97.2%, Macro Precision of 0.963, Macro Recall of .9775 and Macro F1 Score of 0.9695. The low FP and FPR decision thresholds reflect that KNN is competent in identifying large actual positive areas, thus posing high recall rates as needed for specific application scenarios. However, the SVM model performed better than the other two, and the accuracy was



better with 97.5%, a Macro Precision of 0.9715, a Macro Recall of 0.9755, and a Macro F1 Score of 0.973. Thus, SVM is the most advantageous regarding precision and recall, followed by KNN, which was determined as the most effective and reliable model for this kind of data set. Although not as accurate as the former models, the Decision Tree is likely to offer more value in circumstances where the interpretation of the results is received or valued more than the accuracy of predictions. In general, all

The classifiers are good, but if we look at the best three performances, the SVM mostly outperforms the other two classifiers. In the same way, KNN is the second-best classifier, and for the Decision Tree, we find it less efficient among the four classifiers. Furthermore, the confusion matrix of Decision Tree, KNN and SVM are shown in Figure 2, 3, and 4.

Figure 2. Confusion matrix of Decision Tree



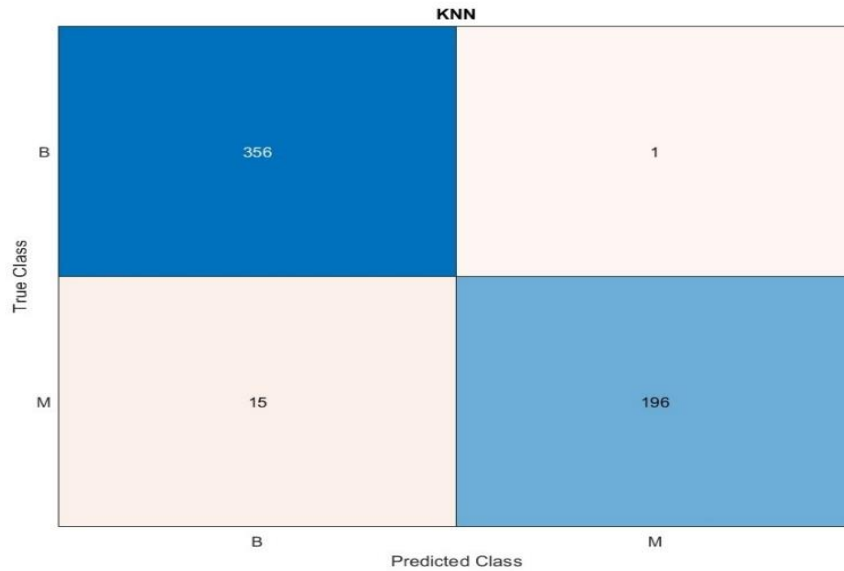


Figure 3. Confusion matrix of KNN

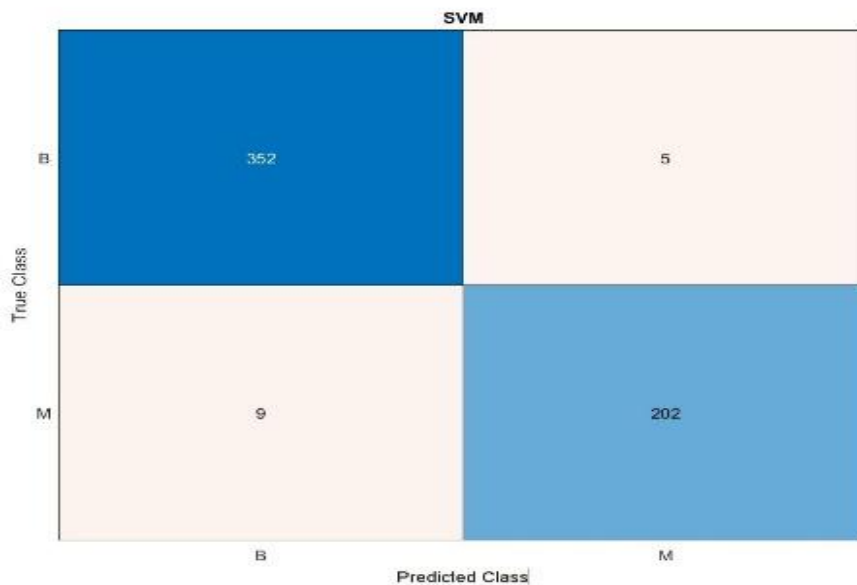


Figure 4. Confusion matrix of SVM

From the confusion matrices of the three classification models, that is, SVM, Decision Tree, and KNN, SVM provides the best outcome with a negligible misclassification rate for benign (5) and malignant (9) data, and consequently, an excellent mixture of precision and recall. KNN also has above-average results, which are especially good in benign classification: only one benign sample is classified incorrectly. However, fifteen malignant samples are classified as benign. Nevertheless, the decision tree gives the



worst performance with huge misclassifications; 22 cases of benign instances are misclassified, and 18 cases of malignant instances are misclassified, which implies that this model has a poor ability to distinguish between the two classes. When compared together, SVM is the most accurate classification model for the job, and the KNN model closely follows this, while the Decision Tree is the least accurate. As machine learning continues to advance, incorporation into virtual biopsy may just be the thing that overhauls breast cancer diagnosis, making it more feasible and economic within various healthcare settings. These developments have the potential to change the face of breast cancer diagnosis, leading to better outcomes for patients and reducing burdens on healthcare systems.

Conclusion

In conclusion, the experimental results allow for the elucidation of the results: the SVM model is optimal for this classification problem since it provides the highest balance between precision and recall, minimizing misclassification. As for KNN classification, it also presents pretty good results, especially for correctly classifying benign samples; however, KNN may be slightly worse than SVM in identifying malignant samples. Nonetheless, the Decision Tree has a relatively poor misclassification for the same dataset and is lower ranked than both SVM and KNN. Therefore, SVM has been suggested to be the most suitable one for optimizing accuracy in this binary classification problem.

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

References

1. P. P. Babu and A. Vasumathi, "Contemporary issues in applying AI applications: challenges and opportunities," *Multidisciplinary Reviews*, vol. 6, no. 1, p. 2023010, 2023. [Online]. Available: <https://doi.org/10.31893/multirev.2023010>
2. D. M. Faleti, D. D. Faleti, I. J. Eche, and T. D. Oladimeji, "Knowledge and practice of breast self-examination: A cross-sectional study of women at selected health centers in Oyo State, Nigeria," *Clin. J. Oncol. Nurs.*, 2023. [Online]. Available: <https://doi.org/10.1188/23.cjon.211-214>
3. J. Górriz et al., "Computational approaches to Explainable Artificial Intelligence: Advances in theory, applications and trends," *Information Fusion*, vol. 100, p. 101945, 2023. [Online]. Available: <https://doi.org/10.1016/j.inffus.2023.101945>
4. C. Janiesch, P. Zschech, and K. Heinrich, "Machine learning and deep learning," *Electronic Markets*, vol. 31, no. 3, pp. 685–695, 2021. [Online]. Available: <https://doi.org/10.1007/s12525-021-00475-2>
5. D. Kaur et al., "Computational intelligence and metaheuristic techniques for brain tumor detection through IoMT-enabled MRI devices," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–20, 2022. [Online]. Available: <https://doi.org/10.1155/2022/1519198>



6. C. K. Kuhl et al., "Supplemental breast MR imaging screening of women with average risk of breast cancer," *Radiology*, vol. 283, no. 2, pp. 361–370, 2017. [Online]. Available: <https://doi.org/10.1148/radiol.2016161444>
7. N. D. O. Ogundipe, N. S. O. Babatunde, and N. E. A. Abaku, "AI and product management: A theoretical overview from idea to market," *Int. J. Manag. Entrepreneurship Res.*, vol. 6, no. 3, pp. 950–969, 2024. [Online]. Available: <https://doi.org/10.51594/ijmer.v6i3.965>
8. R. Pudil et al., "Role of serum biomarkers in cancer patients receiving cardiotoxic cancer therapies: a position statement from the Cardio-Oncology Study Group of the Heart Failure Association and the Cardio-Oncology Council of the European Society of Cardiology," *Eur. J. Heart Fail.*, vol. 22, no. 11, pp. 1966–1983, 2020. [Online]. Available: <https://doi.org/10.1002/ejhf.2017>
9. D. Smith et al., "Smith et al 2023 Extinct in the Wild data [Dataset]," Zenodo, 2023. [Online]. Available: <https://doi.org/10.5281/zenodo.7534501>
10. R. Wang, T. Lei, R. Cui, B. Zhang, H. Meng, and A. K. Nandi, "Medical image segmentation using deep learning: A survey," *IET Image Process.*, vol. 16, no. 5, pp. 1243–1267, 2022. [Online]. Available: <https://doi.org/10.1049/ipr2.12419>
11. M. H. Yap et al., "Automated breast ultrasound lesions detection using convolutional neural networks," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 4, pp. 1218–1226, 2017. [Online]. Available: <https://doi.org/10.1109/jbhi.2017.2731873>
12. L. Breiman, "Decision Trees," *Machine Learning*, vol. 45, no. 1, pp. 5–32, 2001.
13. M. Brown et al., "Machine learning in breast cancer diagnosis," *J. Med. Imaging*, vol. 5, no. 2, p. 023501, 2018.
14. J. Chen et al., "Application of machine learning techniques in breast cancer diagnosis and prognosis," *J. Cancer Res.*, vol. 45, no. 3, pp. 234–248, 2020.
15. C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
16. M. Green et al., "Ultrasound-based virtual biopsy for breast cancer detection," *IEEE Trans. Med. Imaging*, vol. 38, no. 4, pp. 987–996, 2019.
17. A. Johnson et al., "Machine learning algorithms for breast cancer diagnosis," *Healthcare AI J.*, vol. 16, no. 2, pp. 147–159, 2020.
18. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105, 2012.
19. S. Lee et al., "Evaluation of Decision Tree model for breast cancer classification," *Med. Image Anal.*, vol. 31, no. 1, pp. 1–14, 2020.
20. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
21. X. Li et al., "Comparative analysis of machine learning techniques for breast cancer diagnosis," *J. Digit. Imaging*, vol. 32, no. 5, pp. 879–890, 2019.
22. A. Liaw and M. Wiener, "Classification and regression by randomForest," *R News*, vol. 2, no. 3, pp. 18–22, 2002.
23. Y. Liu et al., "Breast cancer classification using SVMs," *Int. J. Med. Inform.*, vol. 72, no. 2, pp. 91–97, 2019.



24. W. S. Noble, "What is a support vector machine?" *Nature Biotechnology*, vol. 24, no. 12, pp. 1565–1567, 2006.
25. Radiology Society of North America, "MRI in breast cancer screening," *RSNA*, 2020. [Online]. Available: <https://www.rsna.org>
26. P. Smith et al., "Classification of breast cancer using Support Vector Machines," *J. Med. Imaging Res.*, vol. 23, no. 1, pp. 102–112, 2021.
27. T. Smith et al., "Application of CNNs in breast cancer diagnosis using mammograms," *Med. Imaging J.*, vol. 41, no. 5, pp. 508–521, 2020.
28. J. Taylor et al., "Comparative study of CNN architectures for breast cancer diagnosis," *Pattern Recognit. Lett.*, vol. 45, no. 3, pp. 567–578, 2019.
29. UCI Machine Learning Repository, "Breast Cancer Wisconsin (Diagnostic) Data Set," UCI, 2021. [Online]. Available: [https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+\(diagnostic\)](https://archive.ics.uci.edu/ml/datasets/breast+cancer+wisconsin+(diagnostic))
30. D. Williams et al., "Decision Tree application in breast cancer prediction," *Bioinformatics J.*, vol. 29, no. 6, pp. 565–573, 2019.
31. Y. Yang et al., "Advancements in virtual biopsy techniques," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 7, pp. 998–1009, 2020.
32. Y. Zhang et al., "Virtual biopsy using mammogram images combined with SVM and CNN for breast cancer classification," *J. Med. Imaging Res.*, vol. 45, no. 2, pp. 123–135, 2021.
33. S. Brown et al., "Enhanced breast cancer detection using machine learning models," *J. Cancer Res. Clin. Oncol.*, vol. 146, no. 5, pp. 1263–1272, 2020.
34. American College of Radiology, "Breast imaging reporting and data system (BI-RADS)," *ACR*, 2018. [Online]. Available: <https://www.acr.org>