JOURNAL OF SOUTHERN COMMUNICATION ENGINEERING
ISLAMIC AZAD UNIVERSITY BUSHEHR BRANCH
E-ISSN: 2980-9231
https://jce.bushehr.iau.ir
https://doi.org/10.30495/jce.2023.1993480.1227

Vol. 14/ No. 53/Autumn 2024

Research Article

Classification of Breast Cancer Masses using Non-Linear Quadratic Support Vector Machine and Comparison with Self-Organizing Neural Network

Soodeh Bakhshandeh, Assistant Professor [1*] | Seyede Monire Atyabi, MSc[2] | Sahar Saberi, Assistant Professor[3]

[1]Department of Computer Engineering, East Tehran Branch, Islamic Azad University, Tehran, Iran, Soodeh.Bakhshandeh@gmail.com, Soodeh.bakhshandeh@iau.ac.ir

[2]Department of Computer Engineering. South Tehran Branch Islamic Azad University. Tehran. Iran, donya.atyabi2017@gmail.com

[3]Department of Computer Engineering. East Tehran Branch, Islamic Azad University, Tehran, Iran, saharsaberi@gmail.com

Correspondence
Soodeh Bakhshandeh, Assistant Professor of Computer Engineering, East Tehran Branch, Islamic Azad University, Tehran, Iran
Soodeh.Bakhshandeh@gmail.com

Received: 1 October 2023
Revised: 12 November 2023
Accepted: 5 December 2023

Abstract

Breast cancer is the second most common cancer after lung cancer and the fifth leading cause of death in women. In less developed countries, breast cancer is the most important cause of death. In this disease, the cells of the breast tissue change and divide into multiple cells and cause a lump. If breast cancer is in the early stages, treatment is possible. There are many treatment methods such as surgery to remove the defective area, drug therapy, radiation therapy, chemotherapy, hormone therapy, and immunotherapy. These treatments have the potential to save lives when administered in the early stages. From the above explanations, it can be seen that early detection of breast cancer is very important and in this research, an attempt has been made to identify suspected cancer data with the quadratic support vector machine method and based on the features extracted from valid and numerous MRI images. Let's classify so that the process of diagnosing the disease in the early stages is easier and faster. The results showed that 356 out of 357 malignant data and 202 out of 211 benign data were correctly classified. The classification accuracy of malignant data was 99.7% and the classification accuracy of benign data was 97.5%, and finally the overall classification accuracy was 98.2%, which indicates the optimal performance of this method in breast cancer data classification.

Keywords: Breast Cancer, Wisconsin Data, Support Vector Machine, Quadratic Kernel.

Highlights

- Developing a method to diagnose breast cancer with the aim of diagnosing the disease in the early stages.
- Applying support vector machine with quadratic kernel with the aim of reducing classification time.
- Using the graphic feature selection method (SU-CFAM) with good speed and performance.
- Achieving 98.2% accuracy without using SU-CFAM, and 99.1% using it.

Citation: S. Bakhshandeh, S. Monire Atyabi, and S. Saberi, “Classification of Breast Cancer Masses using Non-Linear Quadratic Support Vector Machine and Comparison with Self-Organizing Neural Network,” *Journal of Southern Communication Engineering*, vol. 14, no. 53, pp. 1–14, 2024, doi:10.30495/jce.2023.1993480.1227, [in Persian].

1. Introduction

Cancer is a deadly disease, with approximately ten million deaths and 19.3 million reported cases in the past year. Breast cancer is the second most common cancer after lung cancer and the fifth leading cause of death among women. In less developed countries, breast cancer is the leading cause of mortality. In this disease, the cells of breast tissue change and divide into multiple cells, forming masses. Cancer begins in the ducts or lobules connected to the nipples. Most breast masses are benign and non-cancerous, resulting in fibromas, tenderness, thickening of the area, or lumps. Often, small breast tumors show no symptoms and are easily treatable; painless masses are indicative of abnormal cells. Family history, reproductive factors, individual characteristics, excess body weight, diet, alcohol, tobacco, environmental factors, and other risk factors, such as working night shifts, are all related to breast cancer. Breast cancer advances slowly in its early stages but affects other parts of the body over time. Many tests are recommended for diagnosing breast tumors, including mammography, magnetic resonance imaging (MRI), and ultrasound. Mammography is a cost-effective and low-radiation test suggested for early detection of breast tumors. MRI is an alternative test used to confirm the presence of tumors. During an MRI test, allergic reactions to contrast dye may occur, which is an unintended consequence of the procedure. In the early stages, the recommended test is mammography. If breast cancer is detected in its early stages, treatment is possible. There are many treatment methods available, such as surgery to remove the affected area, drug therapy, radiation therapy, chemotherapy, hormone therapy, and immunotherapy. These treatments, when conducted in the early stages, have the potential to save lives. The survival rate, when diagnosed in the early stage, is 90% in developed countries, 40% in South Africa, and 66% in India. Low-income countries have fewer resources, so early detection methods and treatments can be crucial for saving women's lives [1]. From the above explanations, it is evident that early detection of breast cancer is of great importance.

Zerouaoui and Idri, in their research, investigated decision-making systems based on machine learning and image processing for breast cancer imaging. Their article was a structured literature review using machine learning techniques and image processing for breast cancer imaging. A collection of 530 published articles between the years 2000 and August 2019 was selected and analyzed based on ten criteria. The criteria included year and channel of publication, type of experiment, type of research, medical work, machine learning techniques, datasets used, validation methods, performance metrics, and image processing techniques, which included image preprocessing, segmentation, feature extraction, and feature selection. The results showed that deep learning techniques are predominantly used for classification. Most selected studies relied on mammography as the imaging method instead of ultrasound or MRI. Regarding image processing techniques, in most of the selected studies, the input images were altered during preprocessing by noise reduction and color normalization, and some of them utilized segmentation for extracting the region of interest using thresholding methods [2].

Ekici and Jawzal, in their research, examined breast cancer detection using thermography and convolutional neural networks. In their study, a new algorithm was proposed for extracting characteristic features of breast tissue based on biological data, image analysis, and image statistics. These features were extracted from thermal images taken by a thermal camera and were used to classify breast images using convolutional neural networks optimized by a Bayesian algorithm. Using their proposed algorithm, an accuracy rate of 98.95% for thermal images from a dataset belonging to 140 individuals was achieved [3].

Sadhukhan and colleagues conducted a study on breast cancer detection using image processing and machine learning. They presented a computer-based method for early-stage cancer detection in a very short amount of time. Specifically, they used machine learning to train a model utilizing features predictive of cell nuclei. In their research, a comparative study of two different algorithms—K-nearest neighbors and SVM—was performed, measuring the accuracy of each classification. Subsequently, they analyzed a digital image of an aspirated breast tissue sample using image processing to understand the features of the cell nuclei. They then applied the feature values in their trained model to determine whether the developed tumor was benign or malignant [4].

Sahni and Mittal investigated breast cancer detection using image processing techniques in their research. In their proposed work, images obtained from mammography and MRI were used as input for the method focused on tumor detection. In one approach, the tumor-related part of the resulting image was extracted using various segmentation methods, such as edge detection and thresholding. Additionally, various operators were applied to the resulting image, and the performance of the method was validated quantitatively using entropy as a measuring parameter [5].

Adel and colleagues focused on breast cancer detection using image processing and machine learning for elastography images in their research. In this method, image processing techniques were applied to extract features from the obtained images. Furthermore, data preprocessing methods and principal component analysis were applied as a dimensionality reduction technique for the dataset. Model validation was conducted using K-fold cross-validation to ensure the generalization of the algorithm. Accuracy, confusion matrix, and logistic loss were then evaluated for the algorithm used. In this method, the maximum classification accuracy when using SVM with a radial basis function kernel was 94.12% [6].

Gardezi and colleagues conducted a study on breast cancer detection using mammography data. This review aimed to investigate traditional machine learning and deep learning literature with a specific application for breast cancer detection. The study also provides a brief insight into some well-known deep learning networks. In their research, they offered an overview of machine learning and deep learning techniques specifically applied to breast cancer [7].

Varlamis and colleagues conducted a study on "using data mining techniques and data analysis methods to measure mortality and morbidity data related to cancer in a regional cancer registry: the case of Crete, Greece." They believe that examining cancer statistics is very important, as it necessitates long-term planning, ongoing learning, and continuous observation of every cancer patient. The goal of this study was to demonstrate how data mining processes can be used to improve the results of statistical analysis of cancer registry data [8].

Tan and colleagues addressed the relationship between changes in mammographic image features and short-term breast cancer risk. They developed and tested a new computational model to predict the risk of breast cancer in the near future based on a quantitative study of the distinguishing features of bilateral mammographic images in a digital mammography dataset [9].

Yan and colleagues proposed using autoencoder neural networks as a classifier for detecting cancerous regions in prostate images from MRI. In their proposed method, an energy minimization approach for identifying mapping locations was introduced, taking into account the relationship between adjacent pixels [10].

Zhao and colleagues conducted a study on the automatic screening of cervical cells using block image processing. In this study, a computer-aided analytical approach was proposed for identifying suspicious cells in a complete cervical cell image. The main difference of their method from the conventional algorithm is that the image is divided into fixed-size blocks instead of segmented cell images, which can significantly reduce computational complexity. Through data analysis, some texture features and histogram color differences showed significant variations between blocks with and without suspicious cells. Therefore, these features can be used as input for an SVM classifier [11].

In [12], Siwakami and Sarawathi investigated large data extraction: predicting breast cancer using a combined DT-SVM model. This study was conducted using the Wisconsin breast cancer dataset. The accuracy achieved in the proposed method was 91% with an error rate of 2.58%. This paper also compared the proposed algorithm with other methods, including other classification algorithms such as IBL, SMO, and Naïve Bayes. For IBL, the achieved accuracy was 85.23% with an error rate of 12.63%. For SMO, the accuracy was 72.56% with an error rate of 5.96%. For Naïve Bayes, the achieved accuracy was 89.48% with an error rate of 9.89%. Thus, this comparative study showed that DT-SVM outperforms all other classification algorithms.

2. Innovation and contributions

In this research, we aimed to classify data suspected of cancer using machine learning methods such as Support Vector Machines (SVM), based on features extracted from valid MRI images, to make the disease detection process faster and easier in the early stages. The proposed method utilizes Support Vector Machines with a quadratic kernel, aiming to reduce classification time. This kernel is notable for its high-speed nature alongside adequate accuracy compared to most other kernels in this field. Moreover, considering the importance of leveraging feature selection methods in the preprocessing phase, the SU-CFAM method has been employed. This feature selection method is graph-based and offers suitable speed and performance, and additionally, it can be used in both supervised and unsupervised modes.

Among the innovations of this paper, the following can be stated:

- Developing a method to diagnose breast cancer to diagnose the disease in the early stages.
- Applying a support vector machine with a quadratic kernel to reduce classification time.
- Using the graphic feature selection method (SU-CFAM) with good speed and performance.
- Achieving 98.2% accuracy without using SU-CFAM, and 99.1% using it.

3. Materials and Methods

Two main approaches for dimensionality reduction of a dataset are presented: feature extraction and feature selection. In feature extraction, the original feature space is mapped to a smaller space. Essentially, this approach creates fewer features by combining existing features, so that these new features contain all (or most) of the information available in the original features. On the other hand, feature selection involves selecting a subset of the original features. Feature selection is a significant and widely used technique in data preprocessing that increases the speed of machine learning algorithms and improves classifier performance.

In this paper, a symmetric uncertainty map for class-feature association was used for feature selection. In this method, a graphical representation of the features is extracted at the start. After creating the graph in the previous step, the Louvain algorithm is applied to it, resulting in various clusters of nodes (genes) based on the weights of the edges. This way, all nodes within a cluster exhibit the highest similarity with each other, while nodes outside that cluster show the least similarity with nodes within the same cluster. At this stage, the created clusters are referred to as subgraphs, which are considered the input for the next step [13].

In the step of selecting the best feature in each cluster, adjacency matrices, which are binary matrices, are first constructed. Subsequently, to select the best features, all CFAMs in the maximal independent set are chosen.

In machine learning, SVM, also known as Support Vector Machine, is a supervised learning technique that performs classification by finding a hyperplane that maximizes the margin between two classes. The main steps of the SVM algorithm are outlined as follows:

1. Define an optimal hyperplane: maximize the margin.
2. Extend the above definition for non-linearly separable cases: define a penalty for misclassification.
3. Map the data to a high-dimensional space where classification with linear decision surfaces is easier: reformulate the problem so that data is implicitly mapped to this space.

To separate two groups, the simplest approach is through a straight line (1D), a flat plane (2D), or an N-dimensional hyperplane. However, in some cases, a nonlinear region can separate the data more effectively. Therefore, where a linear hyperplane cannot be used to separate two groups, SVM employs two solutions: one is the "kernel function" and the other is "mapping the data to a high-dimensional space" to facilitate linear separation.

When mapping data in a high-dimensional space becomes very complex, a kernel function is used as an alternative solution. Various kernel functions have been proposed, including Gaussian, polynomial, quadratic, and RBF. It is important to note that each kernel function has its own advantages and disadvantages. Therefore, the choice of kernel type and its parameters is left to the user based on the problem at hand [14].

In this study, the quadratic kernel is utilized. The quadratic kernel function (second degree) requires fewer computations compared to the Gaussian kernel function and can serve as an alternative solution when using the Gaussian kernel becomes overly complex.

The second-degree kernel is used to transform the input data into a higher-dimensional space, allowing for the creation of nonlinear decision boundaries. It calculates the dot product between two input vectors, as well as the square of the dot product. This means that it considers not only the linear combination of input variables but also their interactions. By applying the training dataset to the model and using the above formula as the kernel function, we aim to classify the data with a high degree of accuracy.

In this paper, the support vector machine leverages the quadratic kernel, while also recognizing the importance of using feature selection methods in the preprocessing phase. The SU-CFAM method has been employed in the preprocessing phase. This feature selection method is efficient in terms of speed and performance due to its graph-based nature, and it can also be used in both supervised and unsupervised settings.

4. Results and Discussion

In this section, we implement and utilize the SU-CFAM method in the preprocessing phase and the quadratic SVM method in the classification phase using MATLAB scripting. The dataset used in this paper is the breast cancer dataset from the University of Wisconsin, which is considered one of the most reputable datasets in this field, with details provided in the subsequent section.

By applying the mentioned method, we will examine the results related to the classification accuracy of the system trained by this method. In the retrieved data, cancerous tissues are labeled as benign (M) and malignant (B), and various features have also been recorded, which serve as training and testing data for the designed quadratic SVM to perform classification.

In this paper, the breast cancer data is obtained from the University of Wisconsin, known as the WDBC dataset, and it is one of the most valid and comprehensive datasets in this field. The cancerous tissues are marked as benign and malignant, with various features registered and provided as training and testing data for the SVM to perform classification [15].

In the first step, we will investigate the proposed method without utilizing the feature selection method SU-CFAM and with its application. Tables 2 and 3 examine the number of correct and incorrect diagnoses, as well as the classification accuracy using the quadratic SVM method, both without and with the use of SU-CFAM. As can be seen, the proposed method achieves an accuracy of 98.2% without using SU-CFAM, while it reaches an accuracy of 99.1% with the use of SU-CFAM in the preprocessing phase, which is a very favorable accuracy for diagnosing this disease.

Table 4 provides a comparison of the use of various functions in the support vector machine employed in this paper. In this comparison, the results of accuracy, precision, and recall are presented for all proposed methods in two conditions: without and with the use of SU-CFAM. It can be observed that even in this comparison, the proposed method achieves satisfactory accuracy compared to other methods.

Table 5 compares the feature selection methods PCA and SU-CFAM in the feature selection phase. Principal Component Analysis (PCA) transforms data in vector space, allowing the analysis of large datasets with numerous dimensions or features, enhancing data interpretability while preserving the maximum amount of information and visualizing multidimensional data. PCA is essentially a statistical technique for reducing the dimensions of a dataset. This transformation involves linear mapping of the data to a new coordinate system, where (most of) the variations in the data can be described with fewer dimensions than the original data. Many studies use the first two principal components to represent the data in two dimensions and visually identify clusters of closely related data points.

In another comparison, the results obtained from the self-organizing mapping method and the method presented in this paper were examined. In both methods, the feature selection method SU-CFAM was applied. The comparative results of these two methods are provided in Tables 6 and 7.

By closely examining the numbers in the above tables and comparing them with the output figures from the quadratic SVM method, it becomes clear that the final classification accuracy percentage using the method of this study is approximately 14% higher than that of the Self-Organizing Mapping (SOM) method. This significant difference is primarily due to the classification of benign cases, which could be related to the lower number of benign data compared to malignant data. This also indicates that the implemented quadratic SVM method maintains a high level of accuracy in classification, even when fewer data points are available.

Finally, Table 8 presents a comparison between the proposed method and other methods in the field of breast cancer diagnosis based on data from the University of Wisconsin. Nguyen and colleagues [16] achieved an accuracy of 97.50% using feature extraction and collective voting as one of the machine learning algorithms for classifying breast cancer data. In the same domain, using another algorithm called collective learning with the Radial Basis Function (RBF), Osman and colleagues [17] reached an accuracy of 97.00%. Additionally, using the Ellipsoidal algorithm, Dumitru and colleagues [18] achieved an accuracy of 74.24%, and with the multilayer perceptron (MLP) collective learning algorithm, Kaushik and colleagues [19] reached an accuracy of 83.50%.

In a study by Mert and colleagues [20], which also utilized the Wisconsin data for breast cancer classification, the performance of KNN, ANN, and RBF-NN methods was evaluated. The ANN method was classified with an accuracy of 97.53%, the KNN method with 91.00%, and finally, the RBF-NN method with an accuracy of 88%. Furthermore, Moon and colleagues [21] attempted to classify breast cancer data using deep learning, achieving an accuracy of 93.40%.

As indicated by the accuracy percentages of the methods examined in other studies, these methods, at best, achieved about 97.50% accuracy, and at worst, about 74.24%. In contrast, the method implemented in this study achieved a classification accuracy of 98.20%, demonstrating the effective performance of the quadratic SVM method, which, in addition to being nonlinear, offers simpler computations than the previously mentioned methods.

5. Conclusion

Breast cancer has become a leading cause of mortality among women in developed countries. The most effective way to reduce mortality from breast cancer is through early detection. This requires a precise and reliable diagnostic method that allows physicians to differentiate between benign and malignant breast tumors without resorting directly to surgical biopsy. These predictions aim to categorize patients into a "benign" non-cancerous group or a "malignant" cancerous group. It is essential to mention that cancer cells may spread to other parts of the body. In the case of breast cancer, they are transmitted through lymphatic vessels or blood. In this recent scenario, breast cancer is considered advanced, and surgical intervention, known as biopsy, is usually required. It is strongly advised that breast cancer be contained before it progresses to advanced stages. In this article, considering the importance of diagnosing this cancer, we propose a solution that minimizes the necessary computations. The proposed solution in the data preprocessing phase utilizes a graph-based method to reduce the effective features. Then, using a quadratic support vector machine

(SVM), we proceeded to classify the presented data into benign and malignant groups. The quadratic support vector machine significantly helps reduce processing time and resource usage due to its low computational nature. The accuracy of the proposed method, in comparisons with other methods, indicates the success of the solution. It is suggested that future research should involve comparisons with other existing databases in this field. On the other hand, considering the choice of the SU-CFAM method in the feature selection phase, which is a graph-based approach capable of generalizing to datasets with a very high number of features, it is recommended that the selected datasets in future research should be similar to microarray data.

6. Acknowledgement

This research has received no external funding.

7. References

[1] H. Sung, J. Ferlay, R.L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal and F. Bray, “Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” *CA Cancer J. Clin*, vol. 71, no. 3, 2021, doi: 10.3322/caac.21660.

[2] H. Zerouaoui and A. Idri, “Reviewing machine learning and image processing based decision-making systems for breast cancer imaging,” *Journal of Medical Systems*, vol. 45, no. 8, pp. 1-20, 2021, doi: 10.1007/s10916-020-01689-1.

[3] S. Ekici and H. Jawzal “Breast cancer diagnosis using thermography and convolutional neural networks,” *Medical hypotheses*, vol. 137, p. 109542, 2020, doi: 10.1016/j.mehy.2019.109542.

[4] S. Sadhukhan, N. Upadhyay and P. Chakraborty, “Breast cancer diagnosis using image processing and machine learning,”. in *Emerging Technology in Modelling and Graphics* , vol. 937, pp. 113-127, 2020, doi: 10.1007/978-981-13-7403-6_12.

[5] P. Sahni and N. Mittal “Breast cancer detection using image processing techniques,” in *Advances in interdisciplinary engineering*, pp. 813-823, 2019, doi: 10.1007/978-981-13-6577-5_79.

[6] M. Adel, A. Kotb, O. Farag, M. S. Darweesh and H. Mostafa, "Breast Cancer Diagnosis Using Image Processing and Machine Learning for Elastography Images," *International Conference on Modern Circuits and Systems Technologies (MOCAST)*, Thessaloniki, Greece, 2019, pp. 1-4, doi: 10.1109/MOCAST.2019.8741846.

[7] S. J. S. Gardezi, A. Elazab, B. Lei and T. Wang “Breast cancer detection and diagnosis using mammographic data: Systematic review,” *Journal of medical Internet research*, vol. 21, no. 7, p. e14464, 2019, doi: 10.2196/14464.

[8] I. Varlamis, I. Apostolakis, D. Sifaki-Pistolla, N. Dey, V. Georgoulias and C. Lionis “Application of data mining techniques and data analysis methods to measure cancer morbidity and mortality data in a regional cancer registry: The case of the island of Crete, Greece,” *Computer methods and programs in biomedicine*, vol. 145, pp. 73-83, 2017, doi: 10.1016/j.cmpb.2017.04.011.

[9] M. Tan, B. Zheng, J. K. Leader and D. Gur, "Association Between Changes in Mammographic Image Features and Risk for Near-Term Breast Cancer Development," in *IEEE Transactions on Medical Imaging*, vol. 35, no. 7, pp. 1719-1728, July 2016, doi: 10.1109/TMI.2016.2527619.

[10] K. Yan *et al.,* "Comprehensive autoencoder for prostate recognition on MR images," *IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, Prague, Czech Republic, 2016, pp. 1190-1194, doi: 10.1109/ISBI.2016.7493479.

[11] M. Zhao, A. Wu, J. Song, X. Sun and N. Dong, “Automatic screening of cervical cells using block image processing,” *Biomedical engineering online*, vol. 15, no. 1, pp. 1-20, 2016, doi: 10.1186/s12938-016-0131-z.

[12] K. Sivakami and N. Saraswathi, “Mining big data: breast cancer prediction using DT-SVM hybrid model,” *International Journal of Scientific Engineering and Applied Science (IJSEAS)*, vol. 1, no. 5, pp. 418-429, 2015.

[13] S. Bakhshandeh, R. Azmi and M. Teshnehlab, “Symmetric uncertainty class-feature association map for feature selection in microarray dataset,”. *Int. J. Mach. Learn. and Cyber*., vol. 11, pp. 15–32, 2020, doi: 10.1007/s13042-019-00932-7.

[14] A. Shmilovici, “Support Vector Machines,” in *Data Mining and Knowledge Discovery Handbook*, Springer, Boston, MA, 2005, pp. 257-276, doi: 10.1007/0-387-25465-X_12.

[15] https://archive.ics.uci.edu/dataset/17/breast+cancer+wisconsin+diagnostic.

[16] Q. H. Nguyen et al., "Breast Cancer Prediction using Feature Selection and Ensemble Voting," *International Conference on System Science and Engineering (ICSSE)*, Dong Hoi, Vietnam, 2019, pp. 250-254, doi: 10.1109/ICSSE.2019.8823106.

[17] A. H. Osman and H. M. A. Aljahdali, "An Effective of Ensemble Boosting Learning Method for Breast Cancer Virtual Screening Using Neural Network Model," in *IEEE Access*, vol. 8, pp. 39165-39174, 2020, doi: 10.1109/ACCESS.2020.2976149.

[18] D. Dumitru, "Prediction of recurrent events in breast cancer using the Naive Bayesian classification," *Annals of the University of Craiova*, vol. 36, no. 2, 2009.

[19] D. Kaushik and K. Kaur, "Application of Data Mining for high accuracy prediction of breast tissue biopsy results," *Third International Conference on Digital Information Processing, Data Mining, and Wireless Communications (DIPDMWC)*, Moscow, Russia, 2016, pp. 40-45, doi: 10.1109/DIPDMWC.2016.7529361.

[20] A. Mert, N. Kılıç, E. Bilgili and A. Akan, "Breast cancer detection with reduced feature set," *Computational and Mathematical Methods in Medicine*, 2015, doi: 10.1155/2015/265138.

[21] W. K. Moon *et al.,* "Classification of breast tumors using elastographic and B-mode features: comparison of automatic selection of representative slice and physician-selected slice of images," in *Ultrasound in medicineand biology*, vol. 39, no. 7, pp. 1147-1157, 2013, doi: 10.1016/j.ultrasmedbio.2013.01.017.

Appendix

Table 1. WDBC Dataset

Total number of samples: 569
Number of features: 32 features including ID, LAS (benign or malignant), and 30 input features.
These 30 mentioned features consist of three values: mean, standard error, and worst or largest (mean of the three largest values) calculated for each of the ten features below (which are computed for each cell nucleus): • Radius (mean distances from center to points on the perimeter) • Texture (standard deviation of gray-scale values) • Perimeter • Area • Smoothness (local variation in radius) • Compactness • Concavity (severity of concave portions of the contour) • Concave points (number of concave portions of the contour) • Symmetry • Fractal dimension (fractality)

Table 2. The numbers of true and fals detection (Quadratic SVM, without and with using SU-CFAM)

Method	**Detection**	**Malignant Samples**	**Benign Samples**	**Percentage of Malignant Data**	**Percentage of Benign Data**
Without SU-CFAM	Correct	356	202	62.7	35.6
	Incorrect	1	9	0.2	1.6
With SU-CFAM	Correct	357	206	62.9	36.4
	Incorrect	0	5	0	0.8

Table3. The accuracy of the detection (Quadratic SVM, without and with using SU-CFAM)

Method	**Correct Classification Percentage**	**Incorrect Classification Percentage**
Without SU-CFAM	98.2	1.8
With SU-CFAM	99.1	0.9

Table4. Comparison of different kernels of SVM without and with using SU-CFAM on three parameters (Accuracy, Precision and Recall)

SVM with Different Kernels		**Accuracy**	**Precision**	**Recall**
Quadratic	**Without SU-CFAM**	**98.2**	**97.6**	**96.5**
	With SU-CFAM	**99.1**	**98.1**	**97.3**
Gaussian	**Without SU-CFAM**	96.0	93.2	94.9
	With SU-CFAM	97.1	93.4	95.1
RBF	**Without SU-CFAM**	97.6	94.5	96.4
	With SU-CFAM	96.8	95.2	96.3
Tanh	**Without SU-CFAM**	98.1	**97.8**	95.1
	With SU-CFAM	99.90	96.8	97.1

Table5. Comparison of different kernels of SVM with using PCA and SU-CFAM on three parameters (Accuracy, Precision and Recall)

SVM with Different Kernels		**Accuracy**	**Precision**	**Recall**
Quadratic	**PCA**	**99.3**	98.0	**97.3**
	SU-CFAM	**99.5**	**99.2**	**98.5**
Gaussian	**PCA**	96.3	94.5	96.5
	SU-CFAM	97.5	95.4	95.7
RBF	**PCA**	98.3	94.2	96.4
	SU-CFAM	97.4	96.9	97.2
Tanh	**PCA**	98.4	97.2	94.5
	SU-CFAM	98.1	**98.1**	96.9

Table6. The numbers of true and false detection (Quadratic SVM, SOM)

Method	**Detection**	**Malignant Samples**	**Benign Samples**	**Percentage of Malignant Data**	**Percentage of Benign Data**
Our Method	Correct	357	206	62.9	36.4
	Incorrect	0	5	0	0.8
SOM	Correct	356	129	62.7	22.7
	Incorrect	1	82	0.2	14.4

Table7. The Accuracy (Quadratic SVM, SOM)

Method	**Correct Classification Percentage**	**Incorrect Classification Percentage**
Our Method	99.1	0.9
SOM	85.4	14.6

Table 8. Accuracy of the other methods in the breast cancer detection field

Method	**Year**	**Description**	**Accuracy Percentage**
[17]	2020	Ensemble Learning	97.00
]18]	2020	Bayes Algorithm	74.24
[19]	2016	Perceptron NN	83.50
[20]	2015	RBF-NN	88.00
[20]	2015	KNN	91.09
[20]	2015	ANN	97.53
[21]	2013	Deep Learning	93.40
Our Method		Quadratic SVM-Without SU-CFAM	**98.20**
		Quadratic SVM-With SU-CFAM	**99.10**

Declaration of Competing Interest: Authors do not have conflict of interest. The content of the paper is approved by the authors.

Publisher's Note: All content expressed in this article is solely that of the authors, and does not necessarily reflect the views of their affiliated organizations or the publisher, editors, and reviewers. Any content or product that may be reviewed and evaluated in this article is not guaranteed or endorsed by the publisher.

Author Contributions: All authors reviewed the manuscript.

Open Access: Journal of Southern Communication Engineering is an open access journal. All papers are immediately available to read and reuse upon publication.