BI-RADS assessment categorization for breast cancer using Self-Adaptive Particle Swarm Optimization

Bhavya G1, Dr.Manjunath T N2

**Submitted:** 26/01/2024 **Revised:** 04/03/2024 **Accepted:** 12/03/2024

***Abstract:***Mammography is now the gold standard for screening breast cancer. However, approximately 70% of unnecessary biopsies are benign because of the low positive predictive value of mammography interpreted breast biopsies. Computer-aided diagnosis (CAD) technologies may be utilized to reduce the needlessness of breast biopsies. This device helps physicians decide whether to perform a breast biopsy on a suspicious tumor detected during mammography. The suggested CAD system utilizes clustering theory in conjunction with swarm intelligence to classify mass lesions as benign or malignant based on patient age and three BI-RADS attributes: mass shape, mass density and mass margin.The exploration capacity of natural computing algorithms is mostly defined by how diverse their inputs are. When variety is lost too quickly, poor convergence is inevitable. It is also true that, with nearly every paradigm of natural computing, there are number-sensitive parameters accessible, whose ideal values dictate the quality of the solution. In this research, we apply diversity-based self-adaptive clustering to PSO to improve the quality of the resulting clusters. Many variables, including the inertia constant, the social constant, and the cognitive constant, change depending on the current level of variety.

***Keywords:*** *Mammography, Computer-aided Diagnosis, BI-RADS Assessment, Particle Swarm Optimization, Breast Cancer*

1. Introduction

Breast cancer has a higher mortality rate than other forms of cancer in women but is often curable if detected in its early stages. Therefore, early detection techniques for breast cancer using computer-assisted diagnostics are crucial, particularly in developing nations where women often lack access to expert doctors. Mutations in the genes that regulate cell growth and survival and DNA maintenance account for approximately 10% of the mutations that occur in families. Early onset, familial clustering, and associations with several tumor types are all features of hereditary cancer susceptibility. Both the afflicted and unaffected members of a patient's family may eventually understand whether they have a genetic susceptibility to cancer. Understanding cancer risk may aid in both preventative and curative therapeutic strategies.

The ability of machine learning-based models to nonlinearly map between input and output patterns utilizing a supervised training phase is increasing their attractiveness for use in biomedical applications. Because of the exponential increase in data-based knowledge in scientific investigations, researchers are facing several difficulties when attempting to extract relevant insights from these datasets using conventional database methods.

Therefore, efficient mining techniques are required to extract hidden information from massive data repositories. Data-driven insights may be useful in a wide range of fields, from bioinformatics and engineering to biology, medicine, marketing, CRM, criminology, expert prediction, the web, and mobile

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1 Assistant Professor, Department of Information Science and Engineering, BMS Institute of Technology and Management, Visvesvaraya Technological University, Belagavi-590018,India

ORCID ID: 0000-0002-7990-834X

2 Professor, Department of Information Science and Engineering ,BMS Institute of Technology and Management, Visvesvaraya Technological University, Belagavi-590018,India

ORCID ID:[0000-0002-4733-0598](https://orcid.org/0000-0002-4733-0598)

\* Corresponding Author Email: bhavyasati@bmsit.in

computing. Data mining methods such as classification and clustering are often used in the knowledge extraction process. Grouping data items into meaningful classifications, or clusters, with the goal of maximizing similarity between objects inside a cluster and minimizing similarity across clusters, is one of the most important tasks in data mining. In contrast to traditional methods of data analysis, cluster analysis does not require the user to have any previous knowledge of the groups that are created or the individuals who must belong to it to conduct a thorough assessment of the data. Clustering is a technique in data analysis that helps with classification issues, and it is an unsupervised method. The goal is to cluster cases in such a way that the degree of connection is high within the same cluster but low between different clusters. In this approach, the gathered data and the group membership of each cluster may be described. It is a helpful instrument for finding new things. The data structures and relationships that were hidden from view may have become clear. The findings of cluster analysis might help to develop a more concrete taxonomy. Using clustering, we may discover inherent groupings of components that have certain commonalities. In cluster analysis, observations are divided into smaller groups based on their similarities.

PSO is a metaheuristic because it searches extremely wide areas of possible optimizations while making few to no assumptions about the situation at hand. Furthermore, PSO does not depend on the gradient of the problem being improved, unlike conventional optimization techniques such as gradient descent and quasi-Newton methods. A perfect solution may never be discovered with a metaheuristic such as PSO. To achieve a global solution with a higher probability in this work, an adaptive form of PSO, which is based on the diversity of the population in the current stage is proposed.

This work has been divided into several sections. The related work has been presented in section 2, while the detailed mammogram mass data is discussed in section 3. The proposed algorithm is discussed in section 4, while experimental results with numerical optimization and over-the-mammographic data are given in section 5. Finally, the conclusion is given.

1. Related Work

The National Breast Screening Service infrequently makes use of diagnostic mammography equipment located in trusts that do not provide screening services. The research in [1] examined whether women who go to existing screening locations also visit symptomatic units in hospitals. Several mutations in certain genes are linked to an increased risk of breast cancer as well as other cancers in carriers [2]. The use of next-generation sequencing (NGS) has made it possible to examine all genes linked to a family history of breast cancer at once. Clinical management recommendations for individualized preventative and therapeutic care of at-risk persons have been compiled by major multinational working groups in most circumstances due to sufficient research. The work proposed in [3] proposes a system for the detection of breast cancer on mammograms, with the primary objective of making the system's implementation efficient, simple, and effective while minimizing the need for difficult processes or expensive equipment. The proposed solution employs digital image processing methods for this purpose, namely, for enhancing and segmenting the areas of interest. After the mammograms have been analyzed, a smaller collection of characteristics is taken from them and classified using machine learning methods (for example, closest neighbors, decision trees, ANNs, and naive Bayes) in accordance with labels approved by expert doctors.

In [4], the authors introduced a unique YOLOv3 framework for an ideal mass detector that can perform both mass detection and classification in full-field digital mammograms using a single deep convolutional neural network (DCNN). The connected detection head was swapped out for a decoupled head to resolve the remaining model-wide conflict between classification and regression. CAD-based breast cancer detection method published in [5]. To classify tissues as normal or pathological, images are preprocessed before the classification phase, and machine learning is used for the actual classification. To acquire the most useful identifying features, CAD implements the use of the Gabor wavelet to perform preprocessing transformations such as binarization, thresholding, smoothing, and the main operation. Principal component analysis (PCA), t-distributed stochastic neighbor embedding (TSNE), and models of analysis of variance were used for feature selection and dimensionality reduction, and k-nearest neighbors (k-NN) was utilized for classification. Classifying worrisome regions in digital mammograms requires an effective classifier. To improve the classification performance, the authors of [6] explored a stacked ensemble technique that reduces the overall classification error by merging the results of many classifiers and minimizing their variance. Accuracy issues in mammographic image categorization were addressed by proposing a region-specific Multi Attribute white mass estimate method [7]. White mass, density, and binding were used to determine microcalcifications.

Work published in [8] employed AdaBoost with RBFSVM and Hybrid KNN-RBFSVM as the basis estimator, with adaptively modified kernel parameters of RFBSVM (and C). Using the optimized parameters of KNN to adaptively change the SVM kernel parameters, the KNN-RBFSVM classifier was developed for breast mammography classification. The goal of the studies in [9] was to determine the optimal adaptive threshold for mass segmentation and detection in mammography images, which will help radiologists make more accurate diagnoses. The model is a one-layer Legendre network that was developed and trained via the block-based normalized sign-sign least mean square (BBNSSLMS) technique. The study in [10] described a strategy based on the power spectral fractal dimension for the purpose of finding, categorizing, and predicting normal, benign, and malignant regions in a mammogram. This method of area categorization deviates from the norm by making use of the values of fractal dimensions found in the radial directions outward from a central point. Compared with benign and normal lesions, malignant lesions have a substantially greater fractal dimension.

The role of computerized mammography in screening for and diagnosing breast cancer was investigated in [11]. The system is based on the successful multiple-instance learning (MIL) paradigm for healthcare decision-making. The co-training strategy was used in breast mammography classification [12] to make use of unlabeled data. The suggested work incorporates the co-training active learning (COAL) framework, which concurrently uses two-view pictures to increase the classification performance, in addition to leveraging the unlabeled data and importing the annotation efficiency. In [13], the authors presented a morphologically based technique for the automated identification of micro calcifications in digital mammograms. To accommodate the wide range of micro calcification features, several structural components are used. A continuous suspicion map (SM) was generated from the collected morphological maps using the structural similarity index (SSIM). Masses must be manually segmented from native mammograms, which is a laborious and error prone process. Radiologists need help with automated and exact breast mass delineation; thus, a CAD system that works together well is essential. A two-stage multiscale pipeline was reported in [14] to accomplish this goal, allowing for precise mass delineations from high-resolution complete mammograms. There was some discussion of the prior works on PSO in [15].

1. Mammographic Mass Data

The proposed method can be used to distinguish between benign and malignant mammographic mass lesions based on BI-RADS characteristics and patient age. The inclusion criteria included the patient's age, three characteristics from the BI-RADS scale, and the underlying reality of the evaluation. In a double-review procedure, doctors award a BI-RADS score to each patient, with higher scores indicating a greater likelihood of cancer. The scores range from 1 (absolutely benign) to 5. It is assumed that all instances with a BI-RADS score of 5 or above are malignant and that all others are benign. Both the sensitivity and specificity will be determined. These may be utilized to compare the accuracy of a CAD system to that of radiologists. The sample data are given in Table 1.

“Attribute Information:

1. BI-RADS assessment: 1 to 5 (ordinal)

2. Age: Patient age in years (integer)

3. Shape: mass shape:

Round = 1

Oval = 2

Lobular = 3

Irregular = 4 (nominal)

4. Margin: mass margin:

Circumscribed =1

Micro lobulated =2

Obscured =3

ill-defined =4

speculated =5 (nominal)

5. Density: mass density

High =1

Iso =2

Low =3

fat-containing =4 (ordinal)”

**Table 1.** Data Sample

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ***BI-RADS Assessment*** | ***Age*** | ***Shape*** | ***Margin*** | ***Density*** | ***Severity*** |  |
| 5 | 67 | [3](http://en.wikipedia.org/w/index.php?title=Glomerellaceae&action=edit&redlink=1) | 5 | [3](http://en.wikipedia.org/wiki/Sordariomycetes) | Malignant |  |
| 5 | 58 | 4 | 5 | [3](http://en.wikipedia.org/wiki/Leotiomycetes) | [Malignant](http://en.wikipedia.org/wiki/Leotiomycetidae) |  |
| 4 | 28 | [1](http://en.wikipedia.org/w/index.php?title=Glomerellaceae&action=edit&redlink=1) | 1 | [3](http://en.wikipedia.org/wiki/Sordariomycetes) | begign |  |

1. Proposed Solutions: Diversity based self-adaptivness in PSO

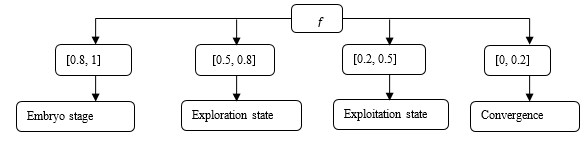
PSO is a metaheuristic that works by defining inspirational development in the population to derive progress as it happens with respect to the number of social species. Conceptually, the PSO has been derived from keeping reference to birds and fish schooling and, to some extent, carrying human social bias. There are three populations whose data are updated with time. It is assumed that each solution has a velocity that is updated to change the solution. The influencing factors that ruled the changes in the velocity were leader inspiration and self-motivation from their past performance. The possibilities of leader criteria can be local or global. A locally based leader can take more time to converge, while a global leader can increase the speed of convergence. The first population can be considered the solution population, which is an encoded form of desired variable values. The second population is considered the velocity population, which provides the upgrading facility to the solution population. The third population is the self-best population, which carried the best values of solution members driven by an individual in the past. This work has used the global variant of PSO. Considering the D-dimensional space of the solution, a ‘jth’ member Mj from the solution population can be represented as Mj =[mj1, mj2,…… mjD]. The sizes of all the tree populations are the same; hence, the velocity population member is also a D-dimensional vector Vj = [vj1, vi2, . . . vjD]. In a similar manner, the previous best solution member can be represented by Bi = [bi1, bi2, . . . biD]. Based on the linear influencing behavior, the velocity upgrading can be described through Eq. 1.

|  |  |
| --- | --- |
| V (n+1) id = χ [w Vnid + Cc Rc (Bnid − Mnid) + Cg Rg (Bngd − Mnid)] | (1) |
| M (n+1) id = M nid + V (n+1) id | (2) |

There are three parameters, w, Cc and Cg, that functionally evolve the characteristics of the velocity. The inertia weight, or 'w,' parameter, is crucial to the convergence properties. Cognitive and social constants (Cc and Cg) are terms for two such quantities. A parameter denoted as a constriction factor is used, with values always smaller than 1, to regulate the divergence behavior of velocity change. In the range [0 1], Rc and Rg are uniformly distributed random numbers.

The applied values of inertia weight w in Eq.1. strongly determine the outcome; hence, proper attention should be given to these values. A low value of ‘w’ supports better exploration in the local region but is the cause of a small change in velocity because of a small change in the solution member. A high value supports exploration in a larger region and helps individuals avoid local optima; however, a high value does not properly explore local regions and can cause convergence in the suboptimal solution. In short, a large inertia weight helps people explore globally, while a small weight supports local exploration. Hence, fine tuning is required to converge the solution optimally. A dynamic approach that carries the larger value of ‘w’ in the beginning, while with iteration, there is a corresponding decrement in the inertia weight. The logic behind this approach is that initially, a random solution exists, and there are likely very high numbers of solutions that are far from the optimal solution; hence, a larger change is needed, while with iteration, it is assumed that the solution is getting better and moving toward the optimal value and hence needs a small inertia weight. Initially, a high value such as 1.2 can be considered, which decreases linearly with iteration up to a value of 0.1 and remains constant after that.

The approach Dynamic PSO seems fine, but the problem is that there is no guarantee that the solution will get better as ‘w’ is changing. This means that an approach that can estimate the quality of population progress with time is needed, and the inertia weight can be determined accordingly. One possibility is to estimate the diversity of the population at present, which could help to determine where the population is present. High diversity indicates the exploration stage, while low diversity indicates convergence. In this work, a diversity-based self-adaptation of inertia weight was proposed.

* 1.  Dynamic PSO

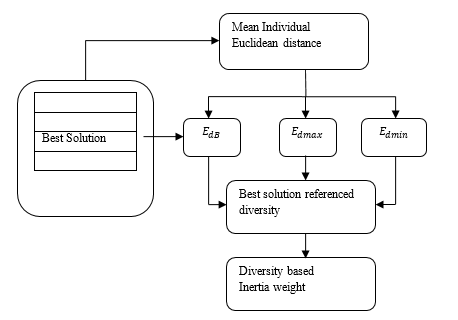
|  |  |
| --- | --- |
|  | (3) |

There is a linear decrease in the ‘w’ values, as given by Eq. 3, where after up to 75% of the iterations, the value of ‘w’ decreases.

where and are the maximum and minimum allowed weight values, respectively; is the maximum allowed number of iterations; and is the counter of the iterations.

* 1. Self-Adaptive PSO

Diversity measures the differences between two entities, and the estimation of the diversity of one solution from others can be estimated by Euclidean distances. For each member, the mean differences are estimated. To determine the diversity with reference to the best solution, the members with the maximum and minimum diversity were used. The available diversity with the best solution is used to estimate the value of the inertia weight. The whole process is defined in Fig. 1. The estimations of individual diversity, best referenced diversity and corresponding values of inertia weights were evaluated via Eq. 4, Eq. 5, and Eq. 6, respectively. The complete workflow of the proposed solution is shown in Fig. 3



**Fig. 1.** Development of self-adaptive inertia weight

(4)

(5)

(6)

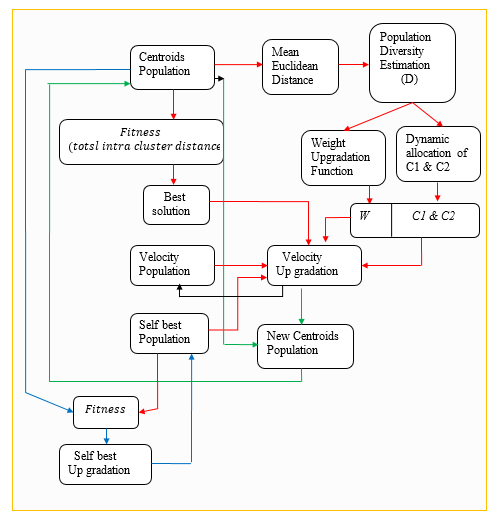
where and are constant parameters and should be between 1 and 3. Empirically, it was estimated that 1.5 and 2.5 were good choices. Depending upon the value of , four categories of population states can be broadly defined, as shown in Fig. 2.

Cognitive and social constants were also allowed to change through a random variation, as given by Eq. 7

**Fig. 2.** Different states of the population with time

(7)

|  |  |
| --- | --- |
|  | (8) |



**Fig. 3.** Working flow of the proposed solution (different connection colors indicate the sequence of flow: red connections as the first flow, green colors as the second flow and blue colors as the third flow)

1. Experimental results
   1. Numerical optimization

To understand the benefit of the proposed self-adaptive PSO algorithm, a numeric benchmark problem called Ackley’s function, as given by Eq. 8, is considered. The problem has dimensions of 10 and multimodal characteristics, where finding the optimal solution is a challenging task. The population size was 50, and the allowed number of iterations was 100. The obtained performances in 10 independent trials are shown in Table 2. It is as seen that dynamic PSO has inferior performance compared to self-adaptive PSO. The inertia weight variations over different iterations are as shown in Fig. 4, and such inertia weights are not possible to define.



**Fig. 4.** Inertia weight variation as a function of diversity

|  |  |  |
| --- | --- | --- |
| **Table 2:** Performance value of Ackley’s function for 10 independent trials | | |
| **Trail No.** | **Dynamic weighted PSO(DPSO)** | **Self-Adaptive PSO(SAPSO)** |
| **1** | **16.9830** | **3.8913** |
| **2** | **24.2531** | **6.1206** |
| **3** | **12.6554** | **7.5474** |
| **4** | **9.6900** | **4.2692** |
| **5** | **20.9369** | **8.1691** |
| **6** | **21.3479** | **2.7714** |
| **7** | **11.7232** | **8.7878** |
| **8** | **19.2179** | **3.3341** |
| **9** | **24.8900** | **4.3881** |
| **10** | **16.8536** | **5.8217** |
| **Mean**  **(Std.Dev)** | **17.8551**  **( 5.2393 )** | **5.5101**  **(2.1105)** |

5.2. Experimental results with Mammographic Mass Data

**Table 4.** Centroid Value Parameters

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | ***BI-RADS Assessment*** | ***Age*** | ***Shape*** | ***Margin*** | ***Density*** |  |
| Benign | 4.3147 | 45.1067 | 2.6344 | 2.6786 | 2.6862 |  |
| Malignant | 4.5407 | 66.0266 | 2.6434 | 3.0938 | 2.6596 |  |

To obtain the “benign” and “malignant” categories, more than 120 datasets considered for the experiment. The adaptive form of PSO applied to develop the clusters, and the process iterated over 100 iterations. The total intra cluster distance considered the fitness

value parameter, which is minimized to make individual clusters compact and enable large separations between clusters. More than 76 sets of data considered for training purposes, and more than 76 were considered for testing purposes. The initial population was selected as a random sample from the dataset to define the centroids of the clusters. Figure 5 illustrates the obtained convergence characteristics, in which it is possible to see evidence of rapid convergence. The obtained performances in terms of sensitivity and specificity are as shown in Table 3. The obtained centroid value parameters are as shown in Table 4.



**Fig. 5**. Convergence characteristics over MM data

The sensitivity and specificity of the obtained results were greater than 90% for both the training and test datasets; in particular, the differences were very small for the test dataset. These outcomes ensure the usefulness of the achieved results in practical aspects. The obtained centroids for both the benign and malignant categories provide a very convenient approach for BI-RADS assessment.

**Table 3** Performance on MMdata by SAPSO

|  |  |  |
| --- | --- | --- |
|  | **Sensitivity** | **Specificity** |
| Training data | 93% | 95% |
| Test Data | 91.2% | 92.6% |

1. Conclusion

The process of analyzing the mammogram was defined through BI-RADS assessment by assigning numeric values to the lesion mass characteristics. The numeric values physically interpreted by the radiologist further contributed to the decision to belong to the benign/malignant category. The proposed method provides a very convenient approach that is based on clustering to determine the final outcomes and has shown a very impressive success rate. With the available data, the radiologist can recheck his/her decision to minimize the final decision error. The diversity-based adaptive value of the inertia weight ensures the convergence of the PSO algorithm to optimality and speed. The proposed method can be integrated into the clinical environment as a support mechanism for radiologists in BI-RADS assessment to improve decision quality and reduce time consumption. Even though a good number of data samples have been considered, considering the dynamic condition of the clinical environment to make the system more robust, it is advisable to continue evolving the solution with new data samples. The performances of other metaheuristic methods, such as differential evolution and other swarm intelligence methods, can be explored further.

Acknowledgements

We thank our institution BMSIT&M and our Management for their abundant support in conducting this research. We thank our collegues, family members for their direct and indirect support.

Author contributions

**Bhavya G:** Conceptualization, Methodology, Software, Field study, Investigation, writing Reviewing and Editing.

**Manjunath T N:** Validation and reviweing.

Conflicts of interest

The authors declare no conflicts of interest.

**References**

1. Jones, E., Zammit, C. & Rubin, G. Should we be using nonscreening symptomatic units' mammogram machines for screening? Women's attitudes and factors likely to affect whether they attend. Breast Cancer Res 13 (Suppl 1), P39 (2011). <https://doi.org/10.1186/bcr2991>.
2. Apessos, A., Nasioulas, G. (2021). Breast Cancer Genetics. In: Rezai, M., Kocdor, M.A., Canturk, N.Z. (eds) Breast Cancer Essentials. Springer, Cham. <https://doi.org/10.1007/978-3-030-73147-2_3>.
3. Carrera, E.V., Sandoval, B., Carrasco, C. (2021). Identification of Breast Cancer Through Digital Processing of Mammograms. In: Guarda, T., Portela, F., Santos, M.F. (eds) Advanced Research in Technologies, Information, Innovation and Sustainability. ARTIIS 2021. Communications in Computer and Information Science, vol 1485. Springer, Cham. <https://doi.org/10.1007/978-3-030-90241-4_30>.
4. Quy, H.D., Son, N.N., Anh, H.P.H. (2023). DeYOLOv3: An Optimal Mass Detector for Advanced Breast Cancer Diagnostics. In: Huang, YP., Wang, WJ., Quoc, H.A., Le, HG., Quach, HN. (eds) Computational Intelligence Methods for Green Technology and Sustainable Development. GTSD 2022. Lecture Notes in Networks and Systems, vol 567. Springer, Cham. <https://doi.org/10.1007/978-3-031-19694-2_29>.
5. Chachalo, B. et al. (2019). Automated Identification of Breast Cancer Using Digitized Mammogram Images. In: Nyström, I., Hernández Heredia, Y., Milián Núñez, V. (eds) Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. CIARP 2019. Lecture Notes in Computer Science, vol 11896. Springer, Cham. <https://doi.org/10.1007/978-3-030-33904-3_32>.
6. Abubacker, N.F., Hashem, I.A.T. & Hui, L.K. Mammographic Classification Using Stacked Ensemble Learning with Bagging and Boosting Techniques. J. Med. Biol. Eng. 40, 908–916 (2020). <https://doi.org/10.1007/s40846-020-00567-y>.
7. Padmavathy, T.V., Vimalkumar, M.N. & Sivakumar, N. Region-specific multiattribute white mass estimation-based mammogram classification. Pers Ubiquit Comput 22, 1093–1098 (2018). <https://doi.org/10.1007/s00779-018-1135-4>.
8. Bhosle, U., Deshmukh, J. Mammogram classification using AdaBoost with RBFSVM and Hybrid KNN–RBFSVM as base estimator by adaptively adjusting γ and C value. Int. j. inf. tecnol. 11, 719–726 (2019). <https://doi.org/10.1007/s41870-018-0241-x>.
9. Sarangi, S., Rath, N.P. & Sahoo, H. Mammogram mass segmentation and detection using Legendre neural network-based optimal threshold. Med Biol Eng Comput 59, 947–955 (2021). <https://doi.org/10.1007/s11517-021-02348-4>.
10. Renjini, A., Swapna, M.S., Raj, V. et al. Power Spectral Fractal Dimension and Wavelet Features for Mammogram Analysis: A Machine Learning Approach. Pattern Recognit. Image Anal. 32, 419–428 (2022). <https://doi.org/10.1134/S105466182202016X>.
11. G. Quellec, M. Lamard, M. Cozic, G. Coatrieux and G. Cazuguel, "Multiple-Instance Learning for Anomaly Detection in Digital Mammography," in IEEE Transactions on Medical Imaging, vol. 35, no. 7, pp. 1604-1614, July 2016, doi: 10.1109/TMI.2016.2521442.
12. Z. Yang, W. Wu, J. Zhang, Y. Zhao and L. Gu, "Deep Co-Training Active Learning for Mammographic Images Classification," 2020 Chinese Automation Congress (CAC), Shanghai, China, 2020, pp. 1059-1062, doi: 10.1109/CAC51589.2020.9327642.
13. Touil, K. Kalti, P. -H. Conze, B. Solaiman and M. A. Mahjoub, "Morphological-based microcalcification detection using adaptive thresholding and structural similarity indices," 2020 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Sousse, Tunisia, 2020, pp. 1-6, doi: 10.1109/ATSIP49331.2020.9231731.
14. Y. Yan, P . -H. Conze, G. Quellec, M. Lamard, B. Cochener and G. Coatrieux, "Two-Stage Multi-Scale Mass Segmentation From Full Mammograms," 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), Nice, France, 2021, pp. 1628-1631, doi: 10.1109/ISBI48211.2021.9433946.
15. Bonyadi, M. R.; Michalewicz, Z. (2017). "Particle swarm optimization for single objective continuous space problems: a review".Evolutionary Computation-25 (1): 1-54.  [doi](https://en.wikipedia.org/wiki/Doi_(identifier)):[10.1162/EVCO\_r\_00180](https://doi.org/10.1162%2FEVCO_r_00180). [PMID](https://en.wikipedia.org/wiki/PMID_(identifier)) [26953883](https://pubmed.ncbi.nlm.nih.gov/26953883).[S2CID](https://en.wikipedia.org/wiki/S2CID_(identifier)) [8783143](https://api.semanticscholar.org/CorpusID:8783143).
16. Kennedy, J.; Eberhart, R. (1995). "Particle Swarm Optimization". Proceedings of IEEE International Conference on Neural Networks. Vol. IV. pp. 1942–1948. [doi](https://en.wikipedia.org/wiki/Doi_(identifier)):[10.1109/ICNN.1995.488968](https://doi.org/10.1109%2FICNN.1995.488968).
17. Kennedy, J.; Eberhart, R.C. (2001). Swarm Intelligence. Morgan Kaufmann.[ISBN](https://en.wikipedia.org/wiki/ISBN_(identifier)) [978-1-55860-595-4](https://en.wikipedia.org/wiki/Special:BookSources/978-1-55860-595-4).