Gazi AKADEMİK YAYINCILIK

AIS ARTIFICIAL INTELLIGENCE STUDIES

ARTIFICIAL INTELLIGENCE STUDIES

Classification of Breast Ultrasound Images Based on Regional and Morphological Features

Zeynep Ak*[a], Şerife Gengeç Benli[b]

ABSTRACT

Breast cancer is a highly prevalent and the most lethal cancer type in women, emphasizing the critical importance of early diagnosis and treatment. This study is based on extracting features from breast ultrasound images (BUSI) from a publicly available dataset. In the study, breast cancer types were examined using regional and morphological features obtained from mask images of breast ultrasound images containing lesions.

Regional and morphological features were extracted from BUSI images, and the least absolute shrinkage and selection operator (LASSO) method was used for feature selection. The results demonstrated that the selected features could effectively distinguish between malignant and benign breast lesions with high accuracy. In this study, machine learning methods such as support vector machines (SVM), artificial neural networks (ANN), and naive bayes (NB) were employed to classify benign and malignant lesions. The classification methods were evaluated using various performance criteria. According to the results, in the study conducted with balanced data, the highest classification performance was obtained with the ANN method with an area under the curve (AUC) value of 0.9973 and an accuracy value of 0.9887.

[a,*] [1]Erciyes University Graduate School of Natural and Applied Sciences, Kayseri/Türkiye
ORCID: 0000-0001-8621-9465

[b] Erciyes University, Engineering Faculty, Dept. of Biomedical Engineering , 32985- Kayseri/Türkiye
ORCID: 0000-0002-5527-8574

[*] Corresponding author.
e-mail: zeynepak.bmm@gmail.com

Keywords: Breast ultrasound image, region and morphological features, machine learning

Anahtar Kelimeler: Meme ultrason görüntüsü, bölgesel ve morfolojik özellikler, makine öğrenmesi

Meme Ultrason Görüntülerinin Bölgesel ve Morfolojik Özelliklere Göre Sınıflandırılması

ÖZ

Meme kanserinin kadınlarda oldukça sık görülen ve en ölümcül kanser türü olması, erken tanı ve tedavinin kritik önemini vurgulamaktadır. Bu çalışma, halka açık bir veri kümesinden meme ultrasonu görüntülerinden (BUSI) özelliklerin çıkarılmasına dayanmaktadır. Yapılan çalışmada, lezyon içeren meme ultrason görüntülerine ait maske görüntülerinden elde edilen bölgesel ve morfolojik özellikler kullanılarak meme kanseri türleri incelendi.

BUSI görüntülerinden bölgesel ve morfolojik özellikler çıkarılmış ve özellik seçimi için an az mutlak büzülme ve seçim operatörü (LASSO) yöntemi kullanılmıştır. Sonuçlar, seçilen özelliklerin kötü huylu ve iyi huylu meme lezyonlarını yüksek doğrulukla etkili bir şekilde ayırt edebildiğini gösterdi. Bu çalışmada, iyi huylu ve kötü huylu lezyonların sınıflandırılmasında destek vektör makineleri (SVM), yapay sinir ağları (YSA) ve Naive Bayes (NB) gibi makine öğrenme yöntemleri kullanılmıştır. Sınıflandırma yöntemleri çeşitli performans kriterleri kullanılarak değerlendirilmiştir. Sonuçlara göre dengeli verilerle yapılan çalışmada YSA yöntemi ile 0,9973 eğri altındaki alan (AUC) değeri ve 0,9887 doğruluk değeriyle en yüksek sınıflandırma performansı elde edilmiştir.

Submitted: 19.12.2023
Revised: 24.08.2024
Accepted: 02.10.2024

doi: 10.30855/AIS.2024.07.02.01

To cite this article: Z. Ak, Ş. Gengeç Benli, "Classification of Breast Ultrasound Images Based on Regional and Morphological Features," *Artificial Intelligence Studies,* doi: 10.30855/AIS.2024.07.02.01

1. Introduction (Giriş)

Breast cancer is a very common type of cancer in women, although it occurs as a result of abnormal cell growth in breast tissue [1]. Genetic, hormonal and environmental factors are effective in the development of breast cancer [2]. In 2020, breast cancer was the most common type of cancer affecting women worldwide, with approximately 2.26 million new cases reported [3]. Breast cancer, which accounts for between 2.24 and 2.79 million cases, underlines the significant extent of this challenge [4]. When we look at these statistics, studies towards the diagnosis of the disease are gaining momentum with technological developments and computer-assisted diagnosis (CAD) approaches.

Biopsy is a definitive method to determine whether a breast lesion is malignant or benign [5]. However, less than 30% of breast tumors detected by surgical biopsy are malignant. Therefore, excessive number of biopsies can be reduced with imaging techniques. There are many techniques such as magnetic resonance imaging (MRI), mammography and ultrasound for the diagnosis of the disease [6, 7, 8]. Compared to mammography and MRI, ultrasound is an important method for diagnosing breast cancer. Detection of breast cancer with ultrasound imaging technique is used as one of the most important diagnostic methods in terms of its low radiation exposure, high sensitivity, non-invasiveness and more accessibility. However, diagnoses based on breast ultrasound (BUS) are more expert-dependent than mammography and MRI, leading to high observational variability between clinicians. CAD systems are preferred to increase the usability of ultrasound and to obtain more reliable and accurate diagnostic results [9]. A BUS-based CAD system includes preprocessing, segmentation, feature extraction, and classification stages. Filtering, contrast dilation, and histogram equalization, which are pre-processing steps for accurate segmentation of the breast tumor, are critical for improving images [10].

For the correct classification of mammary tumors, tissue features such as the gray level co-occurrence matrix (GLCM), local binary patterns (LBP), and histogram of the original gradient (HOG) are used quite frequently and give high accuracy about the structure of the tumor [11]. In the study, a deep learning-based approach was proposed for breast lesion classification in the ultrasound technique, and classification performance was obtained with an accuracy of 0.915 [12]. CAD systems have been proposed to improve diagnostic accuracy in breast cancer detection and classification. The preprocessing step removed the speckle noise using speckle-reducing anisotropic diffusion (SRAD), and active contour-based segmentation was used in the study to find the region of interest (ROI). To classify the images as normal, benign, or malignant, texture features were extracted, and k nearest neighbors (kNN) algorithm, decision tree algorithm, and random forest classifier were used. Performance was compared based on classification accuracy. 83% accuracy success was achieved for the kNN algorithm, 85% for the decision tree algorithm, and 88% for the random forest classifier [13].

In the study, which used tissue characteristics by suggesting that speckle tissue properties are more useful in clinical diagnosis, 80.0% accuracy was obtained by classification with logistic regression [14]. In the study, feature selection was made with a recursive feature elimination-based method by extracting multiple different image features of the tumor region, and classification was made with machine learning algorithms using the BUSI data set. With classifiers such as Random Forest, Adaboost, and Gradient Boosting, it achieved 96.7%, 97.4%, and 96.5% accuracy, respectively [15]. In a study based on texture properties, a gray histogram, a GLCM, and LBP were extracted from the images generated by the superpixel. The researchers used K-means and bag-of-words algorithms to extract features from GLCM and LBP. It was then reclassified using a backpropagation neural network (BPNN) for initial classification and the kNN algorithm for postprocessing. BPNN combined features from superpixels, while the kNN algorithm performed the true classification, achieving an accuracy of 86.5% [16]. In another study, 73 features were obtained using five breast ultrasound image features, including gray level histogram, GLCM, HOG, shape, and position. In order to select the best features for the study to give better results, feature selection was made using the Bicluster score. With the top 25 features selected, 98.3% accuracy was achieved using the SVM classifier [17].

In a study using edge information for breast ultrasound classification, the edge lines of breast ultrasound images were created, and the edge features (maximum curvature sum, maximum curvature and peak sum, maximum curvature sum, and standard deviation) were extracted. Then, morphological features were extracted, and classification was carried out with the SVM algorithm. In the study, in which 192 BUS images were used to evaluate the method, 82.69% accuracy was obtained with edge-based features and 67.31% accuracy was obtained with morphological features [18]. A total of 149 tissue features and 13 morphological features were extracted to examine the effect of different speckle filters in the classification of breast ultrasound images. Classification was performed with SVM using principal component analysis (PCA) to select features from different structures. A total of 100 breast ultrasound images were studied. The results obtained for different filters were recorded as 94.1%, 66.6%, 96.0%, and 68.6% [19]. The aim of this proposed study was to investigate the performance of tumor region and morphological features in differentiating benign and malignant lesions without using tissue features.

In a study conducted by Dörterler et al., clustering accuracy was increased with meta-heuristic algorithms (Differential Evolution Algorithm (DEA) and Harmony Search Algorithm (HSA)) for disease diagnosis and the accuracy of the K-means method was optimized. These meta-heuristic algorithms were tested on the Heart Disease dataset using a hybrid structure with K-means, and an accuracy rate of 85% was obtained with DEA and 66% with HSA [20, 21] In another study, K-means was used with the Death Game Optimization algorithm. By combining the -means algorithm, an effective classification for disease detection has been achieved using lymphography and breast cancer data. The Death Game Optimization algorithm increased the success of the K-means algorithm by 15% [22]. In another study, the effect of optimization methods with the basic Vision Transformer (ViT) model was evaluated. High performance was achieved with 96.6% accuracy and 92.7% F1-score in brain tumor classification divided into four classes with the ViT model [23].

In the presented study, a classification process was carried out with a little number of features from breast ultrasound images in order to analyze the difference only due to the morphological and location characteristics of the lesions. While studies in the literature offer an approach by combining texture and shape features, they focus only on the shape characteristics. In the open access data set used in the study, there are mask images of images containing benign and malignant lesions determined by specialist physicians. Since the morphological features of the lesions will change when the segmentation accuracy decreases, the most accurate classification performance was tried to be obtained by using mask images. It is aimed to contribute to the literature by effectively classifying breast ultrasound images segmented with high accuracy.

2. Materials and Methods (Malzemeler ve Yöntemler)

2.1. Dataset (Veri seti)

In this study for the detection of breast cancer in women with ultrasound images, benign and malignant lesion classification was performed using BUSI (Breast Ultrasound Image), a publicly available data set. The study included 600 female participants (437 benign, 210 malignant, and 133 normal image acquisition) between the ages of 25-75 and a total of 780 images [24]. While mask images with benign and malignant lesions were included in the study, normal images were not included in the study because there was no lesion that could be removed from the mask images. As shown in Figure 1, ultrasound images containing benign and malignant lesions and mask images containing lesion areas determined by specialist physicians were used.

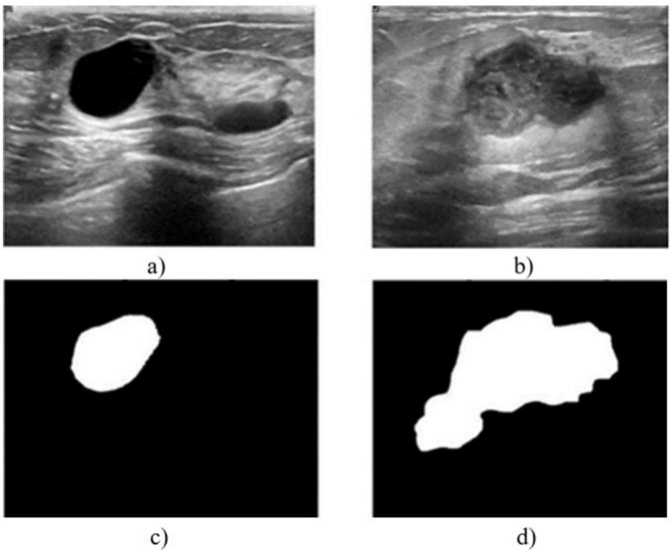

Figure 1: a) Benign US image, b) Malignant US image, c) Mask image (Benign), d) Mask image (Malignant) [24]

The images used in the study are the c and d images given in Figure 1 [24] In many studies, it has been observed that the tissue characteristics and gray-level characteristics of the tumor have been examined. The aim of this study is to make inferences about the location and morphological of benign and malignant tumors.

2.2. Feature Extraction (Özellik Çıkarımı)

The study was carried out in the MATLAB (Matrix Laboratory) environment, which is a multi-paradigm numerical computing software and fourth-generation programming language. The MATLAB program, which is frequently preferred in image processing and classification studies, has been actively used in the execution of the study with different toolbox options. The regionprops architecture, which provides information about region and shape properties in image processing tools, can extract 27 different properties of the image [25]. These properties include information such as the area of objects, their central coordinates, surface area, circumference, axis ratios, and convexity. This information provides information about the shape, size, position, and other properties of objects. A total of 23 features were extracted and tested in classification models.

2.3. Data Balancing (Veri Dengeleme)

When the data sets used for classification are not evenly distributed, the performance of the classification algorithms is not at the desired level. In order to solve this problem and increase performance, the adaptive synthetic sampling approach (ADASYN) was applied to generate synthetic data and create a balanced data distribution. ADASYN is a machine learning algorithm used specifically in the context of imbalanced datasets [26]. Unbalanced data sets are data sets in which the distribution of classes is significantly different. ADASYN aims to reduce this imbalance, but compared to traditional oversampling that replicates existing minority class samples, ADASYN relies on local density distribution when generating synthetic samples for the minority class [27]. Classification was carried out in two ways to see the performance of ADAYSN [28]. Data balancing was performed with ADAYSN in various studies conducted for breast cancer detection. In the study conducted using mammogram images, higher accuracy was achieved by data balancing (ADAYSN) and feature selection (Lasso) [29]. In another study, data balancing was performed with ADASYN using mammogram images and high accuracy was achieved by using a small number of features [30].

2.4 Feature Selection (Özellik Seçimi)

Different feature selection algorithms were examined to select the best features out of 23 features extracted from the images. The Least Absolute Shrinkage and Selection Operator (Lasso) was used for feature selection. Lasso is used to model the relationship between dependent and independent variables. This technique is used to reduce excessive prediction biases during regression analysis as well as to determine which independent variables are important in the model [31].

2.5 Classification (Sınıflandırma)

Classification methods is presented in 3 different ways as Support Vector Machines (SVM), Artificial Neural Networks (ANN) and Naïve Bayes (NB). Average results were given by 10-fold cross-validation for each classification method.

SVM is an important classification and regression algorithm in the field of machine learning. The main goal of SVM is to set a decision boundary to best separate data points. This decision boundary can be a line, a plane, or a hyperplane that separates data points from two different classes. The standout feature of SVM is its focus on creating the maximum margin surrounding this decision boundary [32].

ANN is an important machine learning method in the field of deep learning, inspired by biological nervous systems. Artificial Neural Networks are basically created by combining many artificial neurons in layers. By processing the incoming data, these neurons calculate the outputs that are transmitted to the next layer. ANN learns through the training process. In this process, the error between the inputs given to the network and the expected outputs is corrected, and the weights and parameters inside the network are updated. This feedback loop allows the network to adapt to the data and learn a specific task [33].

Naive Bayes (NB) is an effective algorithm used in machine learning and statistical classification problems. This algorithm is based on Bayes' Theorem and is used to model data sets and make predictions in classification tasks [34].

Within the scope of the study, firstly, mask images of the lesion areas of ultrasound images in the BUSI data set, determined by specialist physicians, were selected. Various region and shape features were extracted from these images using Matlab software. By feature selection, the best features were determined, synthetic data was created for data balance, and Lasso regression was applied to select the best features. In order to test the performance of various machine learning models, classification models were run 10 times with 10-fold cross validation and average results are presented. To examine the effect of data balancing on classification performance, the findings were evaluated with both unbalanced data (original data) and balanced data. The lambda value, which is an important parameter in Lasso regression, was taken as 0.005 and the machine learning models used were improved with hyperparameters. The method presented in the study is explained in Figure 2.

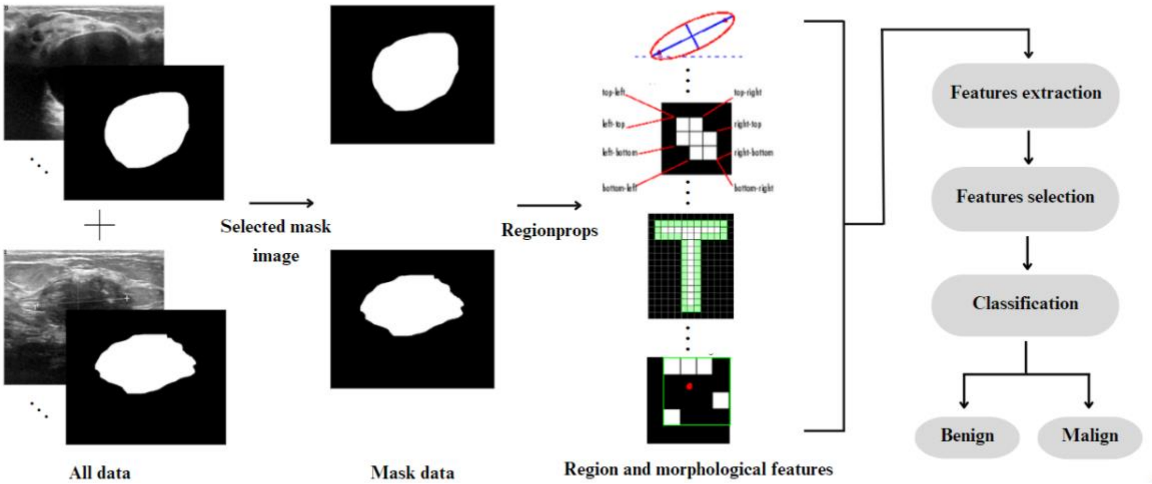

Figure 2: Feature extraction and classification diagram of mask images

3. Results (Sonuçlar)

In the presented study, mask images of BUS images were used. Various region and shape features were extracted and feature selection was made with Lasso. The classification results using the features obtained before and after data balancing are presented comparatively in Table 1. The success of region and shape features in detecting benign and malignant BUS images has been tested with different machine learning methods such as SVM, ANN and NB. The performance criteria of machine learning methods presented using region and shape features are given in Table 1 as Area Under the Curve (AUC), Accuracy (Acc), Sensitivity (Sens), Specificity (Spec), Positive Predictive Value (PPV), Negative Predictive Value (NPV), F-Score (F1) is presented. Classification methods are presented as results by obtaining average performance using a 10-fold cross-validation approach. The study was conducted on an i5 Acer Nitro 5 computer with Windows 11 operating system and Matlab2023a version was used. Bayesian optimization was used and the model was improved with hyperparameters by making less oversampling with regularization parameters.

Tablo 1. Classification with region and morphological features

Type of study		AUC	Acc	Sens	Spec	PPV	NPV	F1
Classification	SVM	0.9922	0.9658	0.9328	0.9816	0.9608	0.9681	0.9466
	ANN	0.9924	0.9743	0.9500	0.9846	0.9675	0.9762	0.9586
	NB	0.9784	0.9374	0.9042	0.9533	0.9033	0.9539	0.9037
Classification with Adaysn	SVM	0.9930	0.9703	0.9777	0.9631	0.9628	0.9779	0.9702
	ANN	0.9973	0.9887	0.9955	0.9821	0.9819	0.9955	0.9887
	NB	0.9759	0.9339	0.9362	0.9315	0.9298	0.9382	0.9329

4. Discussion and Suggestions (Tartışma ve Öneriler)

In the study, an approach based on the region and figural features obtained from the mask images of the BUS images, consisting of 437 benign and 210 malignant images, was presented. When the findings of the study were examined, it was observed that the extracted features and classification methods achieved high accuracy in the classification of breast tumors. It has been observed that classification by data balancing increases classification accuracy in SVM and ANN methods. The best classification performance was obtained at 0.9973 with the ANN classifier using balanced data.

When studies in the literature are examined, various medical imaging applications using different methods and algorithms are seen. For example, a study using speckle texture features achieved 80.0% accuracy with logistic regression [14]. In another study, various image features were extracted from the tumor area and classification was made with machine learning algorithms, and high accuracy rates were obtained [15]. The use of edge information in classifying breast ultrasound images was also examined and different accuracy rates were observed compared to morphological features [18]. Improving clustering accuracy in disease diagnosis through the use of meta-heuristic algorithms has also been investigated [20,21]. Finally, high performance in brain tumor classification was achieved with the basic Vision Transformer model. These studies shed light on future research by evaluating the effectiveness of various methods in the field of medical imaging.

When we look at the other studies examined, it is seen that the study resulted in better accuracy than the literature. It is predicted that the classifications to be made by adding different textures and morphological features to the study can increase the level of success. However, since the size of the features that will emerge when different features are added will increase, classifying the features with the best distinguishing structure with different feature selection techniques will carry the success to higher levels.

Data availability statement

The performed breast ultrasound dataset, generated by Al-Dhabyani et al., can be downloaded from https://doi.org/10.1016/j.dib.2019.104863 (accessed on 10 October 2021).

Conflict of interest

The authors declare no conflict of interest.

References (Kaynaklar)

[1] Feng, Y., Spezia, M., Huang, S., Yuan, C., Zeng, Z., Zhang, L., ... and Ren, G. Breast cancer development and progression: Risk factors, cancer stem cells, signaling pathways, genomics, and molecular pathogenesis. Genes & diseases, 5(2), 77-106, 2018, doi: 10.1016/j.gendis.2018.05.001.

[2] American Cancer Society, https://www.cancer.org/cancer/types/breast-cancer/about.html.

[3] J. Ferlay, M. Ervik, F. Lam, M. Colombet, L. Mery, M. Piñeros, A. Znaor, I. Soerjomataram, F. Bray, Global Cancer Obser-Vatory: Cancer Today; International Agency for Research on Cancer: Lyon, France, 2020, https://gco.iarc.fr/today.

[4] Xu, H., & Xu, B. Breast cancer: Epidemiology, risk factors and screening. Chinese Journal of Cancer Research, 35(6), 565, 2023, doi: 10.21147/j.issn.1000-9604.2023.06.02.

[5] B. Liu, H. D. Cheng, J. Huang, J. Tian, X. Tang and J. Liu, Fully automatic and segmentation-robust classification of breast tumors based on local texture analysis of ultrasound images. Pattern Recognition, 43(1), 280-298, 2010, doi:10.1016/j.patcog.2009.06.002.

[6] A. W. C. Liew and Hong Yan, "An adaptive spatial fuzzy clustering algorithm for 3-D MR image segmentation," in IEEE Transactions on Medical Imaging, vol. 22, no. 9, pp. 1063-1075, Sept. 2003, doi: 10.1109/TMI.2003.816956.

[7] A. T. Stavros, D. Thickman, C. L. Rapp, M. A. Dennis, S. H. Parker, and G. A. Sisney, Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. Radiology, 196(1), 123-134, 1995, doi: 10.1148/radiology.196.1.7784555.

[8] E. Warner, D. B. Plewes, K. A. Hill, P. A. Causer, J. T. Zubovits, R. A. Jong, ... and Narod, S. A. Surveillance of BRCA1 and BRCA2 mutation carriers with magnetic resonance imaging, ultrasound, mammography, and clinical breast examination. Jama, 292(11), 1317-1325, 2004, doi: 10.1001/jama.292.11.1317.

[9] K. Drukker, M. L. Giger, C. J. Vyborny and E. B Mendelson, Computerized detection and classification of cancer on breast ultrasound. Academic radiology 11.5 (2004): 526-535. https://doi.org/10.1016/S1076-6332(03)00723-2.

[10] A. Madabhushi and D. N. Metaxas, "Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions," in IEEE Transactions on Medical Imaging, vol. 22, no. 2, pp. 155-169, Feb. 2003, doi: 10.1109/TMI.2002.808364.

[11] A. H. Farhan, and M. Y. Kamil, Texture Analysis of Breast Cancer via LBP, HOG, and GLCM techniques. IOP conference series: materials science and engineering. Vol. 928. No. 7. IOP Publishing, 2020. doi: 10.1088/1757-899X/928/7/072098.

[12] M. Byra, ''Breast mass classification with transfer learning based on scaling of deep representations'', Biomedical Signal Processing and Control, Volume 69, August 2021, 102828, doi: https://doi.org/10.1016/j.bspc.2021.102828.

[13] S. Pavithra, R. Vanithamani, J. Justin, ''Computer aided breast cancer detection using ultrasound images'', Materialstoday: Proceedings Volume 33, Part 7, 2020, Pages 4802-4807, doi: https://doi.org/10.1016/j.matpr.2020.08.381.

[14] C. M. Lo, R. Chang, C. Huang and W. Moon, Computer-aided diagnosis of breast tumors using textures from intensity transformed sonographic images. 1st Global Conference on Biomedical Engineering and 9th Asian-Pacific Conference on Medical and Biological Engineering, Tainan, Taiwan. (pp. 124–127), 2015. doi: 10.1007/978-3-319-12262-5_35.

[15] A. K. Mishra, P. Roy, S. Bandyopadhyay and S. K. Das, Breast ultrasound tumour classification: A Machine Learning—Radiomics based approach. Expert Systems, 38(7), e12713, 2021. https://doi.org/10.1111/exsy.12713.

[16] Q. Huang, Y. Huang, Y. Luo, F. Yuan and X. Li, Segmentation of breast ultrasound image with semantic classification of superpixels. Medical Image Analysis, 61, 101657, 2020. doi: 10.1016/j.media.2020.101657.

[17] Q. Huang, F. Yang, L. Liu & X. Li, Automatic segmentation of breast lesions for interaction in ultrasonic computer-aided diagnosis. Information Sciences, 314, 293-310, 2015. https://doi.org/10.1016/j.ins.2014.08.021.

[18] Y. Liu, L. Ren, X. Cao, Y. Tong, Breast tumors recognition based on edge feature extraction using support vector machine. Biomed Signal Process Control 58:101825 2020. https://doi.org/10.1016/j.bspc.2019.101825.

[19] Kriti, J. Virmani, and R. Agarwal. Effect of despeckle filtering on classification of breast tumors using ultrasound images. Biocybernetics and biomedical engineering 39.2: 536-560. 2019. https://doi.org/10.1016/j.bbe.2019.02.004.

[20] Dörterler, S. Hybridization of k-means and meta-heuristics algorithms for heart disease diagnosis. New Trends In Engineering And Applied Natural Sciences, 55, 2022. doi:10.30855/gmbd.0705N01.

[21] Dörterler, S. Dumlu, H. Özdemir, D. and Temurtaş, H. Melezlenmiş K-means ve Diferansiyel Gelişim Algoritmaları ile Kalp Hastalığının Teşhisi. In International Conference on Engineering and Applied Natural Sciences (Vol. 1844), 2022.

[22] Dörterler, S. Kanser Hastalığı Teşhisinde Ölüm Oyunu Optimizasyon Algoritmasının Etkisi. Mühendislik Alanında Uluslararası Araştırmalar VIII, 15, 2023.

[23] Şahin, E. Özdemir, D. and Temurtaş, H. Multi-objective optimization of ViT architecture for efficient brain tumor classification. Biomedical Signal Processing and Control, 91, 105938, 2024. https://doi.org/10.1016/j.bspc.2023.105938.

[24] W. Al-Dhabyani, M. Gomaa, H. Khaled and A. Fahmy, Dataset of breast ultrasound images, Data in Brief, 2020, 28, 104863, https://doi.org/10.1016/j.dib.2019.104863.

[25] https://www.mathworks.com/help/images/ref/regionprops.html.

[26] Pristyanto, Y. Nugraha, A. F. Pratama, I., Dahlan, A. and Wirasakti, L. A. Dual approach to handling imbalanced class in datasets using oversampling and ensemble learning techniques. In 2021 15th International Conference on Ubiquitous Information Management and Communication (IMCOM) (pp. 1-7). IEEE, 2021. doi: 10.1109/IMCOM51814.2021.9377420.

[27] Wang, A. X. Chukova, S. S. and Nguyen, B. P. Synthetic minority oversampling using edited displacement-based k-nearest neighbors. Applied Soft Computing, 148, 110895, 2023. doi:10.1016/j.asoc.2023.110895.

[28] H. He, Y. Bai, E. A. Garcia and S. Li, ADASYN: Adaptive synthetic sampling approach for imbalanced learning, Proc. IEEE Int. Joint Conf. Neural Netw. IEEE World Congr. Comput. Intell., pp. 1322-1328, Jun. 2008. doi: 10.1109/IJCNN.2008.4633969.

[29] Fusco, R. Piccirillo, A. Sansone, M. Granata, V. Rubulotta, M. R. Petrosino, T., ... and Petrillo, A. Radiomics and artificial intelligence analysis with textural metrics extracted by contrast-enhanced mammography in the breast lesions classification. Diagnostics, 11(5), 815, 2021. DOI: 10.3390/diagnostics11050815.

[30] Khan, T. M. Xu, S. Khan, Z. G. and Uzair Chishti, M. Implementing multilabeling, ADASYN, and ReliefF techniques for classification of breast cancer diagnostic through machine learning: Efficient computer-aided diagnostic system. Journal of Healthcare Engineering, 2021(1), 5577636, 2021. https://doi.org/10.1155/2021/5577636.

[31] R. Tibshirani, Regression Shrinkage and Selection Via the Lasso, J. R. Stat. Soc. Ser. B. 58 (1996) 267–288. https://doi.org/https://doi.org/10.1111/j.2517-6161.1996.tb02080.x.

[32] M. Lloyd-Williams, Case studies in the data mining approach to health information analysis, Knowledge Discovery and Data Mining 1998/434) IEEEColloquium on, pp. 1/1-1/4. doi: 10.1049/ic:19980641.

[33] J. J. Hopfield, "Artificial neural networks," in IEEE Circuits and Devices Magazine, vol. 4, no. 5, pp. 3-10, Sept. 1988, doi: 10.1109/101.8118.

[34] L. Jiang, H. Zhang and Z. Cai, "A Novel Bayes Model: Hidden Naive Bayes," in IEEE Transactions on Knowledge and Data Engineering, vol. 21, no. 10, pp. 1361-1371, Oct. 2009, doi: 10.1109/TKDE.2008.234.

This is an open access article under the CC-BY license