Significance of Training Images and Feature Extraction in Lesion Classification

Adél Bajcsi[a], Anca Andreica[b] and Camelia Chira[c]
Babeș–Bolyai University, Cluj-Napoca, Cluj, Romania
{firstname.lastname}@ubbcluj.ro

Keywords: Breast Lesion Classification, Mammogram Analysis, Shape Features, Random Forest, DDSM.

Abstract: Proper treatment of breast cancer is essential to increase survival rates. Mammography is a widely used, non-invasive screening method for breast cancer. A challenging task in mammogram analysis is to distinguish between tumors. In the current study, we address this problem using different feature extraction and classification methods. In the literature, numerous feature extraction methods have been presented for breast lesion classification, such as textural features, shape features, and wavelet features. In the current paper, we propose the use of shape features. In general, benign lesions have a more regular shape than malignant lesions. However, there are exceptions and in our experiments, we highlight the importance of a balanced split of these samples. Decision Tree and Random Forest methods are used for classification due to their simplicity and interpretability. A comparative analysis is conducted to evaluate the effectiveness of the classification methods. The best results were achieved using the Random Forest classifier with 96.12% accuracy using images from the Digital Dataset for Screening Mammography – DDSM.

1 INTRODUCTION

Breast cancer is one of the most common cancer types among women (Chhikara and Parang, 2022). In 2021, 2.26 million women were diagnosed with this cancer type worldwide. Accurate diagnosis and timely treatment of breast cancer are crucial for the successful management of this disease.

Mammography is a widely accepted screening tool for the early detection of breast cancer. In recent years, several studies have been conducted to develop mammogram analysis and classification systems that can help radiologists.

Breast cancer classification based on mammogram analysis consists in (1) detection of possible lesions, and (2)classification of those lesions as malignant or benign. In the current study, we focus on constructing a classification system for the second step, to distinguish the different tumors. The difficulty in mammogram analysis lies in proper preprocessing of the images, as mammograms have complex and diverse characteristics (such as brightness, contrast, and resolution). In these systems, it is crucial to select an appropriate dataset with sufficient variability in tumor shapes and sizes for effective training of the system.

Numerous feature extraction methods have been presented in the literature for the classification of breast lesions, including textural features, shape features, and wavelet features. In the present study, our proposal focuses on using shape features to enhance the classification process. Generally, benign lesions are more circular, whereas malignant lesions are more spiculated. Therefore, we decided to use shape features. Unfortunately, this is not a universal rule, as is evident from our experiments. When training the model, we should pay special attention to including examples of outliers in both the training and test sets. In this paper, we address a subsection to emphasize the importance of a proper split of the dataset.

Computer-aided systems based on mammogram analysis are a frequently researched area in the field of breast cancer detection and classification;hence, several methods have been proposed for the accurate identification of tumors. These methods often utilize machine learning algorithms, such as Support Vector Machines, Gaussian Mixture Models, Decision Tree-based methods, and Artificial Neural Networks (ANNs), to classify breast lesions as malignant or benign based on various features extracted from mammograms. In recent years, several studies have shown

[a] https://orcid.org/0009-0007-9620-8584
[b] https://orcid.org/0000-0003-2363-5757
[c] https://orcid.org/0000-0002-1949-1298

Bajcsi, A., Andreica, A. and Chira, C.
Significance of Training Images and Feature Extraction in Lesion Classification.
DOI: 10.5220/0012308900003636
Paper published under CC license (CC BY-NC-ND 4.0)
In *Proceedings of the 16th International Conference on Agents and Artificial Intelligence (ICAART 2024) - Volume 3*, pages 117-124
ISBN: 978-989-758-680-4; ISSN: 2184-433X
Proceedings Copyright © 2024 by SCITEPRESS – Science and Technology Publications, Lda.

promising results in the classification of breast lesions using ANNs. However, one limitation of these methods is the requirement for a large amount of annotated data for training, which is often challenging to obtain in the medical field. Additionally, for a computer-aided system, it is important to explain the outcome of the classification process to provide transparency and enhance trust in the system. To address these challenges, we decided to focus on Decision Tree-based models, as they offer transparency and interpretability in the classification process. The key contribution of this work is the development and evaluation of a Decision Tree-based model for the classification of breast lesions using shape features extracted from 904 mammograms and achieving a classification accuracy of 96.12%.

The rest of the paper is structured as follows. Existing approaches are presented in Section 2. Section 3 details the approach investigated using shape characteristics, followed by a discussion of the results in Section 4. Finally,Section 5 presents the main conclusions and future directions.

2 RELATED WORK

Breast mass classification is a critical task in the field of breast cancer detection, and various computer-aided systems have been proposed to accurately identify tumors. In the following paragraphs, we present existing approaches from the literature.

An important aspect of mammogram analysis is the extraction of features. In the literature, there are two main categories of features commonly used for breast mass classification: shape-based features, texture-based features, and intensity-based features. The shape-based features (used in (Paramkusham et al., 2021; Gurudas et al., 2022; Singh et al., 2020)) focus on the geometric properties of the lesion, such as its size, shape and contour characteristics. On the other hand, texture-based features (used in (Shanmugam et al., 2020; George Melekoodappattu et al., 2022)), capture the spatial arrangement and patterns of the pixel intensities.

Paramkusham et al. (Paramkusham et al., 2021) applied Beam angle statistics to extract the shape features (1-dimensional signature of the mass). In the experiments, the authors included the K-nearest Neighbor (KNN), Support Vector Machine (SVM), and Artificial Neural Network (ANN) classification methods and reported the best accuracy of 88.8% using 147 contours from the Digital Dataset for Screening Mammography (DDSM) (Heath et al., 2001; Heath et al., 1998) to distinguish benign and malignant lesions.

Gurudas et al. (Gurudas et al., 2022) extracted 18 shape features from images of the Curated Breast Imaging Subset of DDSM (CBIS-DDSM) (Lee et al., 2017) including area, bounding box, convex image, convex hull, length of the minor and major axes, centroid, moments, orientation, aspect ratio, eccentricity, compactness, solidity, extrema, extent and minimum, maximum and mean intensities. The authors compared the performance of SVM and ANN to differentiate lesions and concluded that ANN outperforms SVM, reaching 97. 24% accuracy over 92. 91% achieved by SVM.

Shanmugam et al. (Shanmugam et al., 2020) compared the performance of different texture features combined with statistical features to classify tumors from DDSM (323 benign and 323 malignant images). The authors fed the computed features to an SVM classifier and achieved a precision of 79. 7% using the Gray-Level Cooccurrence Matrix (GLCM), 69% using the Gray-Level Run-Length Matrix (GLRLM), 91. 5% using the Gray-Level Difference Matrix (GLDM) and 97. 5% using the Local Binary Pattern (LBP).

In another experiment, Rani et al. (Rani et al., 2023) segmented the region of interest, followed by the extraction of the texture and shape characteristics. Using images from the DDSM the Adaptive Neuro-Fuzzy Classifier with Linguistic Hedges (ANFC-LH) achived 73% while Principal Component Analysis (PCA) SVM reached 72%.

Kurami et al. (Kumari et al., 2023) proposed a hybrid feature extraction and hybrid feature selection (HFSE) framework with ANN classification to distiguish benign and malign lesion from DDSM. In the feature extraction, the authors included GLCM, Gabor filter, Tamura and LBP features. For feature selection, first, the correlated features are defined by Extremely randomized trees classifier-based feature selector, and then the final selection is performed with the ANOVA F-value test. Kurami et al. (Kumari et al., 2023) reported 94.57% accuracy.

Convolutional Neural Networks (CNNs) have the advantage of fuzing the feature extraction and classification steps. Therefore, they have gained significant attention in the field of breast mass classification. Salama and Aly utilized Convolutional Neural Networks (CNNs) in their work (Salama and Aly, 2021). The authors first applied U-Net to segment the lesion, which was later classified by Inception V3. To distinguish between benign and malignant lesions 98.87% accuracy was achieved. Felconí et al. (Falconí et al., 2020) conducted experiments on popular CNNs such as VGG, ResNet, DenseNet, and Inception with transfer learning. The authors highlighted the importance of fine-tuning, which can increase performance by 20%, resulting in 84.4% accuracy on DDSM. Singh et al. (Singh et al., 2020) presented a CNN for the extraction of shape characteristics and classification of lesions from DDSM in four classes (irregular, lobular, oval, and round) and achieved 80% accuracy.

Sajid et al. (Sajid et al., 2023) proposed a combination of high- and low-level characteristics. The authors simplified the original VGG model to create the compact VGG (cVGG) in response to a small number of classes. The deep (high-level) features extracted by cVGG were concatenated with low-level features extracted from Histogram of oriented gradients (HOG) and LBP. Using the resulting characteristics, they achieved 75%, 85% and 91. 5% accuracy, respectively, using RF, KNN and Extreme Gradient Boosting (XGBoost) classification methods. The input images were used from CBIS-DDSM. Sajid et al. (Sajid et al., 2023) concluded that complex CNNs are not always robust enough for lesion classification and the positive effect of the combination of features in this classification problem.

In recent years, there has been a growing interest in ensemble approaches, where different machine learning methods are combined. Melekoodappattu et al. (George Melekoodappattu et al., 2022) proposed a combination of CNN and a "traditional" model. The second model was a KNN using as input texture feature extraction (GLRLM) with Maximum Variance Unfolding (MVU) feature selection to distinguish breast lesions. The proposed ensemble model achieved 95.2% accuracy.

In these diagnostics systems, it is important the extraction of features that are used in the classification. These studies highlight the advantages of shape features in mammogram classification systems.

3 PROPOSED APPROACH

The objective of the current study is to differentiate lesions from mammograms. In our approach, a simple computer-aided system is constructed using features that describe the shape of the tumor. Fig. 1 shows the flow diagram of our proposed system. In the following paragraphs, the components of this system are detailed.

3.1 Preprocessing

Mammograms are high-resolution images of breast tissue. An example is presented in Fig. 1a. In order to extract shape features, the image needs to be processed by cropping the area of interest (containing the lesion), as shown in Fig. 1b. The shape characteristics are not influenced by the pixel information; thus, a binary mask of the tumor (shown in Fig. 1c) is generated and used in subsequent steps.

3.2 Feature Extraction

Feature extraction plays a crucial role in classification systems. As mentioned in Section 1, the literature presents several types of characteristics used in mammograms, grouped as texture, shape, and wavelet features. Shape characteristics capture geometric properties related to the boundary of the lesion, such as size (area), perimeter, compactness, irregularity, and asymmetry. In particular, compared to alternative characteristics, shape features are the most robust because they do not depend on intensity, contrast, or resolution. When presented to radiology specialists, the shape features are more straightforward, forming a more understandable decision-making. Last but not least, shape features are computationally less expensive due to their dependence on the contour of the lesion instead of the pixels of the lesion and its surrounding. It is crucial to note that defining the exact boundaries of the lesions is required in order to compute the shape features.

To distinguish between benign and malignant lesions, we decided to use shape features. In general, benign lesions have a more regular (circular) shape, whereas malignant lesions have a more irregular (spiculated) shape. Therefore, we extracted two types of shape features from the lesions: (1) geometrical and (2) contour-based (Li et al., 2017).

Geometrical features are simple features that are used as a baseline, including the perimeter, area, and compactness of the lesion. The contour-based features (Li et al., 2017; Bajcsi and Chira, 2023), on the other hand, are based on the boundary information of the lesion. In the first step, an ellipse is fitted around the lesion by defining its center (C), minor (b) and major (a) axis lengths, and a rotation angle (α). The fitted ellipse to a malignant lesion is presented in Fig. 1d. Next, the distance between C and each point at the lesion boundary is calculated (d_l) as well as the corresponding point distance from the ellipse (d_e), resulting in a chart similar to Fig. 1e. We denote by Δd the difference between d_l and d_e. Finally, the irregularity of the contour is measured by the root mean slope ($R\Delta q$ defined in equation (1)), root mean roughness (Rq defined in equation (2)) and the circularity (defined in equation (3)) of Δd, detailed in (Li et al., 2017). In the experiments carried out, we further investigated the influence of computing the features mentioned above in subsegments (s equal segments

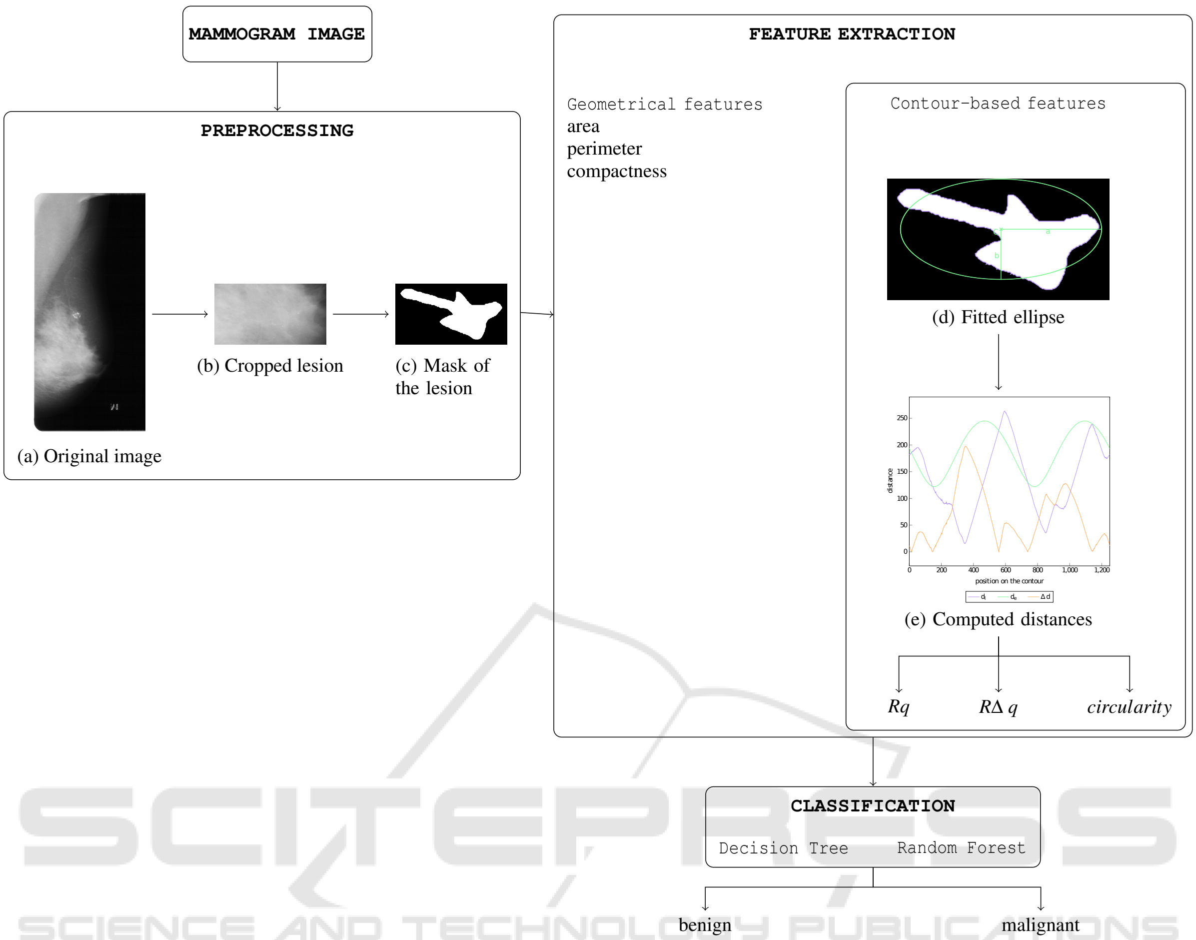

Figure 1: The flow of the system. Illustrations of contour feature extraction are used from (Bajcsi and Chira, 2023).

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS

for $s \in \mathcal{S}, \mathcal{S} = \{1,2,4,6,8,10,12,14,16,18,20\}$).

$$R\Delta q = \sqrt{\mu\left[\left(\sum_i \beta^{(i)}\right)^2\right]}, \text{ where}$$

$$\beta^{(i)} \text{ is the local tilt around point } i \qquad (1)$$

$$\beta^{(i)} = \frac{1}{60\Delta d_i}\big(\Delta d_{i-3} - 9\Delta d_{i-2} + 45\Delta d_{i-1} - 45\Delta d_{i+1} + 9\Delta d_{i+2} - \Delta d_{i+3}\big)$$

$$Rq = \sqrt{\mu(\Delta d^2) - \mu(\Delta d)^2} \qquad (2)$$

$$circularity = \frac{\mu(\Delta d)}{\sigma(\Delta d)} \qquad (3)$$

3.3 Classification

From a lesion mask, three geometrical features and S contour-based features are extracted. Due to the relatively small number of features in the present experiment, all of them are fed to the classification model. To classify the lesion based on the extracted features, two models are used, namely (1) Decision Tree (DT) and (2) Random Forest (RF). These models can generalize from fewer inputs than ANNs because they do not have hidden parameters, which have to be optimized during the training. Furthermore, tree-based models offer interpretability due to their structure, giving them an advantage over ANNs.

DTs are nature-inspired models and can be represented as a series of `if-then-else` structures, where every branch either defines the output class or contains another `if` statement. RFs make their decisions by constructing multiple DTs and summarizing their results.

The disadvantage of these models is their tendency to overfit, failing to generalize the features given in the training set. To address this problem, we propose the use of ensemble models (RF). On the other hand, pre-pruning is also applied.

4 EXPERIMENTS AND RESULTS

The scope of our experiment is to distinguish the character of lesions based on shape features using the system presented in Section 3. In the following subsection, the selected dataset and the achieved results are presented.

4.1 Data Processing

In the current experiments, images from the Digital Mammography Screening Dataset (Heath et al., 1998; Heath et al., 2001) are used from side view (mediolateral oblique MLO). The dataset contains for each mammogram a corresponding mask of the lesion, thus allowing the extraction of shape features. The DDSM dataset is made up of 1,440 instances of cancerous cells that are classified as benign (712 samples) or malignant (728 samples). In order to have a balanced dataset, we randomly selected 712 images from the malignant class, resulting in a dataset containing 1424 images.

After conducting preliminary experiments, presented in (Bajcsi and Chira, 2023), we decided to further analyze the lesions recorded in the dataset. The analisys revealed that the border (and the calculated features) of benign and malignant lesions had small differences. Fig. 2 shows the irregularity of the lesions grouped by class. The irregularity is measured by the average distance ($\overline{\Delta d}$) between the fit ellipse and the boundary of the tumor. Training input can greatly influence the construction of the classification model. If the model is trained using images from the first part and tested for the second part, then there will be a massive difference in the training and testing accuracy. In our previous study (Bajcsi and Chira, 2023), we reached 100% training accuracy when only 64.99% test accuracy was achieved using the same features and classifiers. Therefore, the difference can be explained by the inadequate split of the images. To overcome this issue, we decided to extract subsets from the dataset as shown in Fig. 2 where the minimum difference in the irregularity is fixed. The lesions corresponding to the specified condition are selected and randomly split (75% train, 25% test). In the creation of the new splits, we took special care to always have the same images in the same set. This approach is equivalent to a stratified sampling, where in addition to the type of the lesion, its contour is considered. This approach resulted in 6 new split files. The distribution of samples in each partition is presented in Table 1, which includes the number of images used for training and testing. In the following, we present the results using these subsets.

4.2 Results

The scope of the present study is to verify the importance of a proper train test split and to evaluate the performance of shape features in the binary classification of lesions. To train the models, the previously mentioned splits are used. Furthermore, in the training process, k-fold cross-validation ($k = 5$) was applied to avoid overfitting and increase the reliability of the results. In the current subsection, the results achieved in the experiments conducted are presented.

As mentioned in Section 3.2, there are three types of features extracted from the contour of the lesion. First, we want to select the best contour features. Therefore, we built separate models for each computed feature and for different numbers of segments, obtaining 3 features $\times |\mathcal{S}| = 33$ RF models. In addition, we built models using the combination (concatenation) of contour features, obtaining $|\mathcal{S}| = 11$ models corresponding to each number of $\mathcal{S}$. Due to the fact that with a lower $\overline{\Delta d}$ value the difference in the result is more emphasized, we present the results obtained on the largest split ($\overline{\Delta d} > 0$) containing a total of 904 mammograms. The performance of contour-based features is evaluated and shown in Fig. 3. We can conclude that from the proposed contour-based features (root mean roughness, root mean slope, and circularity), the models that use root mean roughness (Rq) outperform the models trained using the other two features, independent of the number of segments. Moreover, Fig. 3 shows that the combination of the extracted contour features slightly improves the classification accuracy. Based on the data presented in Fig. 3, we can conclude that the Random Forest algorithm is the most effective when using a combination of contour-based features and $s = 10$ segments. The best performing model reached an accuracy of 96.12%.

Fig. 4 compares the performance of different machine learning algorithms using the most effective features found in previous experiments (combination of contour-based features computed from 10 segments) and our baseline geometric features. First, it can be observed that with a higher threshold at $\overline{\Delta d}$ we can increase the accuracy of the classification. The results then indicate that contour-based features outperform baseline geometric features. We can also remark that Random Forest consistently provided the best classification accuracy across all scenarios tested.

The presented method has the drawback that it requires a precise lesion mask for extracting shape features, thus restricting the experiments to datasets with this information. This could potentially limit its applicability in real-world scenarios where such detailed

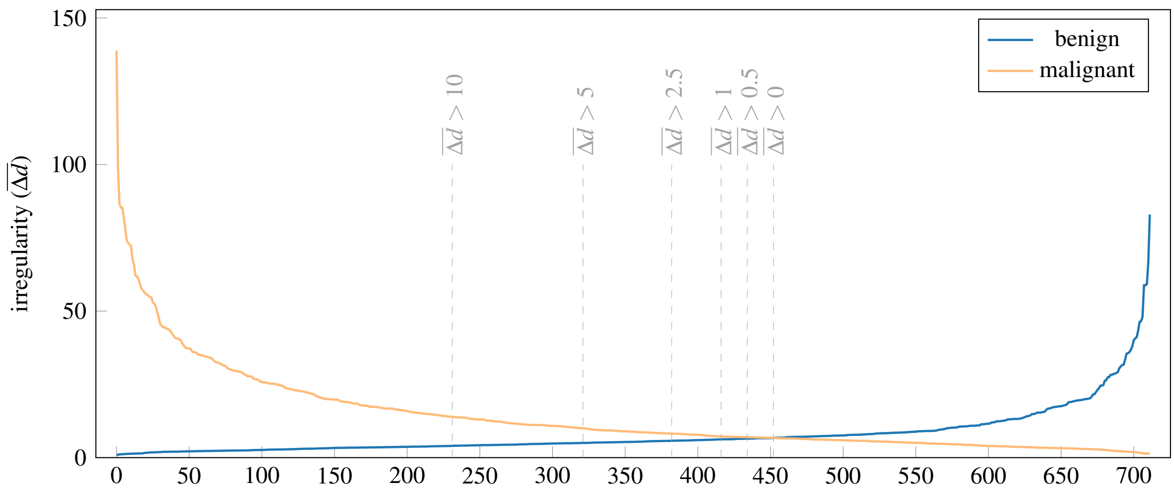

Figure 2: Irregularity ($\overline{\Delta d}$) between the ellipse and the lesion boundary for benign and malignant samples. Distances from the benign class are sorted increasingly, whereas distances from the malignant class are sorted decreasingly, based on the presumption that benign lesions are usually round, whereas malignant lesions tend to be irregular in shape.

Table 1: Number of selected samples in total after splitting the dataset to have at least the specified difference between the average distances.

$\overline{\Delta d}$		samples per class	total samples	train samples	test samples
10	<	231	462	346	116
5	<	321	642	479	163
2.5	<	382	764	569	195
1	<	416	832	620	212
0.5	<	434	868	647	221
0	<	452	904	674	230

SCITEPRESS
SCIENCE AND TECHNOLOGY PUBLICATIONS

data may not always be available.

4.3 Discussion

In the previous section (Section 4.2), the performance of the proposed approach is presented. In the current section, the results obtained are compared with other methods presented in the literature. Table 2 collects and compares existing approaches from the literature.

Compared to our previous model (Bajcsi and Chira, 2023), where we achieved 64.99% due to overfitting, the current approach shows a significant improvement. Hence, the importance of proper train test split arises.

Li et al. (Li et al., 2017) reported 99.66% accuracy using root mean slope features as input to the SVM classifier. To train the reported model, the authors used 323 images from the DDSM (from a total of 14,440). The presented method, using an RF classifier, outperformed the presented method in (Li et al., 2017), by reaching 100% test accuracy using 642 images to train and test the model.

Paramkusham et al. (Paramkusham et al., 2021) proposed another applied Beam angle statistics method to extract the shape features. The authors reported the best accuracy of 88.8% using 147 contours from the DDSM. In our experiment, the best result achieved is 96.12%.

In the research presented by Falconí et al. (Falconí et al., 2020), different CNNs were compared using transfer learning on images from CBIS-DDSM (an updated version of DDSM). The authors reported the best performance by VGG-16 achieving 64.4% using the original mammograms (1696). The result of VGG-16 was increased by employing fine tuning (generating 60 000 images with augmentation) reaching 84.4%.

In a study conducted by Salama and Aly (Salama and Aly, 2021), the performance of CNN was investigated. The authors cropped the region of the lesion, applied data augmentation to achieve a more robust model, and reported 98.87% accuracy on DDSM (from 564 images, 1804 were generated with augmentation). Our model is behind the model reported by

Table 2: Comparison of the current approach to existing solutions from the literature.

Approach	Images	Model	Accuracy
(Rani et al., 2023)	518	ANFC-LH	73%
(Falconí et al., 2020)	60 000	VGG-16	84.4%
(Paramkusham et al., 2021)	147	SVM	88.8%
(Kumari et al., 2023)	428	ANN	94.57%
current	904	RF	96.12%
(Salama and Aly, 2021)	1804	U-Net + Inception V3	98.87%
(Li et al., 2017)	323	SVM	99.66%

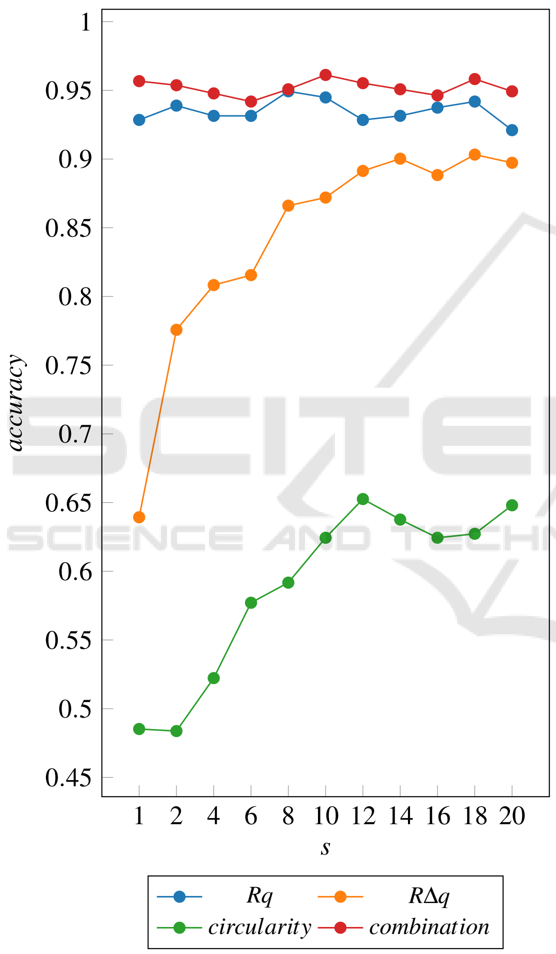

Figure 3: Results achieved with the different contour features computed for the different number of segments (s), using RF classifier.

Salama and Aly (Salama and Aly, 2021) by 2.75%. This difference can be decreased by further investigation of the model's parameters.

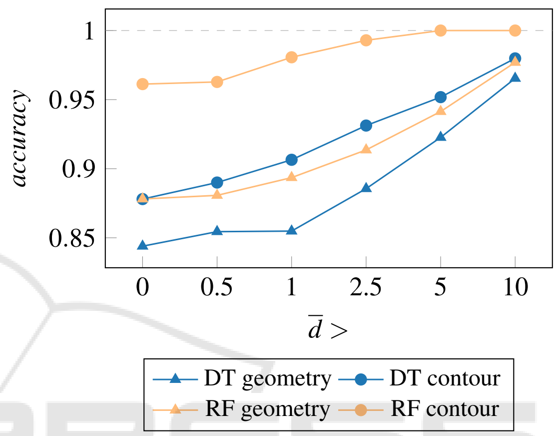

Figure 4: Comparison of the geometry- and contour-based features using different classification model.

5 CONCLUSIONS AND FUTURE WORK

Mammogram analysis plays a key role in the early detection of breast cancer. The earlier a lesion is explored, the higher the chances of recovery. In the presented paper, an early detection system is proposed using shape features to distinguish between benign and malignant tumors. To classify the extracted features Random Forest and Decision Tree methods are used. The best performance achieved was 96.12% accuracy. The results of the experiments conducted have shown that a proper split between train and test is crucial to achieve accurate classification of the lesions. The reported results are comparable with other state-of-the-art approaches.

In future experiments, we will investigate the parameters of the classification models used. We will also compare the performance of the texture features with the proposed contour-based features on these new splits. We will also investigate the effect of the combination of texture- and contour-based features on the performance of the system. We will consider the

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS

use of other explainable models (e.g. interpretable ANNs) or representation learning. To increase the amount of input image, augmentation will be taken into consideration.

REFERENCES

Bajcsi, A. and Chira, C. (2023). Textural and shape features for lesion classification in mammogram analysis. In *Hybrid Artificial Intelligent Systems*, pages 755–767. Springer Nature Switzerland.

Chhikara, B. S. and Parang, K. (2022). Global cancer statistics 2022: the trends projection analysis. *Chemical Biology Letters*, 10(1):451.

Falconí, L. G., Pérez, M. H., Aguila, W. G., and Conci, A. (2020). Transfer learning and fine tuning in breast mammogram abnormalities classification on CBIS-DDSM database. *Advances in Science, Technology and Engineering Systems Journal*, 5(2):154–165.

George Melekoodappattu, J., Sahaya Dhas, A., Kumar K., B., and Adarsh, K. S. (2022). Malignancy detection on mammograms by integrating modified convolutional neural network classifier and texture features. *International Journal of Imaging Systems and Technology*, 32(2):564–574.

Gurudas, V. R., Shaila, S. G., and Vadivel, A. (2022). Breast cancer detection and classification from mammogram images using multi-model shape features. *SN Computer Science*, 3(5):404.

Heath, M., Bowyer, K., Kopans, D., Kegelmeyer, P., Moore, R., Chang, K., and Munishkumaran, S. (1998). *Current Status of the Digital Database for Screening Mammography*, pages 457–460. Springer Netherlands, Dordrecht.

Heath, M., Bowyer, K., Kopans, D., Moore, R., and Kegelmeyer, P. (2001). The digital database for screening mammography. In Yaffe, M., editor, *Proceedings of the Fifth International Workshop on Digital Mammography*, pages 212–218. Medical Physics Publishing.

Kumari, D., Yannam, P. K. R., Gohel, I. N., Naidu, M. V. S. S., Arora, Y., Rajita, B., Panda, S., and Christopher, J. (2023). Computational model for breast cancer diagnosis using hfse framework. *Biomedical Signal Processing and Control*, 86:105121.

Lee, R. S., Gimenez, F., Hoogi, A., Miyake, K. K., Gorovoy, M., and Rubin, D. L. (2017). A curated mammography data set for use in computer-aided detection and diagnosis research. *Sci Data*, 4:170–177.

Li, H., Meng, X., Wang, T., Tang, Y., and Yin, Y. (2017). Breast masses in mammography classification with local contour features. *BioMedical Engineering OnLine*, 16(1):44.

Paramkusham, S., Thotempuddi, J., and Rayudu, M. S. (2021). Breast masses classification using contour shape descriptors based on beam angle statistics. In *2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA)*, pages 1340–1345.

Rani, J., Singh, J., and Virmani, J. (2023). Mammographic mass classification using dl based roi segmentation and ml based classification. In *2023 International Conference on Device Intelligence, Computing and Communication Technologies, (DICCT)*, pages 302–306.

Sajid, U., Khan, R. A., Shah, S. M., and Arif, S. (2023). Breast cancer classification using deep learned features boosted with handcrafted features. *Biomedical Signal Processing and Control*, 86:105353.

Salama, W. M. and Aly, M. H. (2021). Deep learning in mammography images segmentation and classification: Automated cnn approach. *Alexandria Engineering Journal*, 60(5):4701–4709.

Shanmugam, S., Shanmugam, A. K., Mayilswamy, B., and Muthusamy, E. (2020). Analyses of statistical feature fusion techniques in breast cancer detection. *International Journal of Applied Science and Engineering*, 17:311–317.

Singh, V. K., Rashwan, H. A., Romani, S., Akram, F., Pandey, N., Sarker, M. M. K., Saleh, A., Arenas, M., Arquez, M., Puig, D., and Torrents-Barrena, J. (2020). Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network. *Expert Systems with Applications*, 139:112855.

SCITEPRESS SCIENCE AND TECHNOLOGY PUBLICATIONS