

# New Features for the Classification of Mammographic Masses

Florian Wagner  
Fraunhofer-Institute for  
Integrated Circuits  
Am Wolfsmantel 33, 91058  
Erlangen, Germany

Thomas Wittenberg  
Fraunhofer-Institute for  
Integrated Circuits  
Am Wolfsmantel 33, 91058  
Erlangen, Germany

## ABSTRACT

Computer-assisted diagnosis (CADx) for the characterization of mammographic masses as benign or malignant has a high potential to help radiologists during the critical process of diagnostic decision making. We have developed a new set of features for the characterization of masses which is especially designed to describe the intensity transition from the center of a mass up to its surrounding tissue. Furthermore, we have investigated the performance of this set with different image quantization (8 bit and 12 bit). The suggested features are based on the idea to characterize the lesion with a predefined number ( $k$ ) of concentric regions defined by the distance to its margin and to the border of its segmentation, respectively. We evaluated the classification performance for different values of  $k$  using the area  $A_z$  under the receiver operating characteristic (ROC) curve. Our dataset contained 750 lesions from a publicly available mammography database. For each  $k$  an optimal feature subset was selected by a genetic algorithm. The  $A_z$  of these subsets ranged from 0.74 to 0.76 on 8 bit images and from 0.76 to 0.77 on 12 bit images.

## Keywords

Breast cancer, CAD, mammography, mass, classification, feature extraction

## 1. INTRODUCTION

Breast cancer is the most common cancer among women worldwide. According to the German Federal Statistical Office 27% of all newly developed cancer diseases in Germany are caused by breast cancer [1]. Several studies have shown that early detection of breast cancer through periodic mammographic screening can reduce the mortality [2].

Today, mammography is the most effective method available for breast cancer screening and indispensable for the diagnostic process. However, mammography has a relatively low specificity for the discrimination between malignant and benign lesions and therefore further image based examination methods such as MRI or US must be used. Nevertheless, today no single one of the available imaging techniques is able to make an exact discrimination between benign and malign lesions. Hence, the biopsy of the lesion in combination with a pathological examination remains the diagnostic gold standard.

However, any biopsy is an invasive procedure, and it has been noticed that only 20% to 30% of all biopsies show a malignant pathology [3]. These numbers show the importance to improve the positive predictive value of mammography in order to

reduce both, the costs of the examinations as well as the patients' discomfort.

## 1.1 Objective

Even after 25 years of research the image-based discrimination of mammographic lesions is a very difficult task and requires knowledge about the appearance of the different lesions in mammograms. The shape of a mass and the characteristics of its margin are important aspects, which can be used for the discrimination of malign and benign masses. Another typical attribute is the presence of spiculations, which are linear structures radiating from the core of the mass into the surrounding tissue. Such spiculations are most often found in connection with malignant masses. A further crucial characteristic is the so-called mammographic density. Differences in the absorption of X-rays lead to varying intensities in the mammogram. High density meaning high gray color intensities in the mammographic image is an indication for a malignant mass. In contrast to that, transparency is an evidence for a benign lesion. The last aspect to mention is the so-called acutance. Benign lesions tend to be sharply outlined with an abrupt transition between lesion and surrounding tissue whereas malignant masses often possess a fuzzy boundary in connection with a smooth intensity transition between the mass and the background.

In order to support radiologists during the challenging task of discrimination and diagnosis of mammographic lesions, we have developed a knowledge-based system for the computer-assisted diagnosis (CADx) of mammographic lesions [4,5]. In contrast to computer-assisted detection (CADe) - systems, which are designed to automatically detect suspicious lesions on a mammogram, our CADx-system focuses on the discrimination of benign and malignant findings. Hence, the system supplies a second opinion to support the decision process of the radiologist. Within this CADx-system, content-based image retrieval (CBIR) methods as well as case-based reasoning (CBR) techniques [6] are applied to retrieve similar lesions with known diagnoses from a database of annotated and histologically validated digital or digitized mammograms of reference lesions. Based on the histological validated ground truth diagnoses of the retrieved lesions, a diagnosis proposal for the unknown query lesion can be generated and displayed to the radiologist. Furthermore, to make the reasoning process more transparent to the operator, the  $n$  most similar retrieved lesions, their corresponding diagnoses and additional annotations are presented to the radiologist in an adequate way together with the generated diagnosis proposal.

The objective of this study is to present and evaluate a set of new features which is designed to characterize the intensity transition from the center of a mammographic mass up to its surrounding tissue. The feature set is able to describe both the density and the acutance of mammographic masses. Furthermore the influence of the image quantization is investigated.

## 1.2 Related Work

Using texture analysis for the discrimination of masses has been a popular approach in research. In the majority of cases features based on gray-level co-occurrence matrices [7], gray level run-length metrics [8] or wavelet decompositions [9] have been used for the characterization of masses [10-13].

Another approach has been the development of features which are characterizing special attributes of mammographic lesions. This is a challenging task and has been widely discussed in the literature. Huo et al. [14-16] have proposed two features that measure the amount of spiculation of a mass based on an analysis of the radial gradient of its contour. Furthermore, Zheng et al. [17] have developed two features that describe the margin characteristics of a mass using the standard deviation and the skew of the gradient strength of the pixels on the contour of a mass. Spiculated margins are an important characteristic of malignant masses. Hence, several research groups have proposed to apply texture analysis on bands of pixels that are close to the margin of a segmented mass. Extending this idea, Sahiner et al. [18] introduced the so-called rubber band straightening transform (RBSST) which transforms a band of pixels surrounding a segmented mass onto the Cartesian plane (the RBSST image). Experiments have shown that texture features extracted from the RBSST image are superior to texture features extracted from the untransformed band of pixels.

Rangayyan et al. [19] have presented features describing acutance and the shape of mammographic masses. To characterize the acutance they used pairwise pixel differences on normals through the contour pixels. On a dataset containing 54 mammograms the best feature subset yielded a classification accuracy of 94.4%. The acutance feature was the best feature with an accuracy of 92.6%. Shi et al. [20] have developed two margin abruptness features that measure the margin sharpness using line detection in RBSST images. ROC-analysis of these two features on a dataset containing 909 lesions yield  $A_z$ -values of  $0.60 \pm 0.02$  and  $0.64 \pm 0.02$ , respectively.

Additionally, Rangayyan et al. [20] have introduced new features based on a polygonal model of the contour of a mass. After transforming the border of the lesion into a  $n$ -sided polygon they computed features describing convexity, concavity, compactness and the probability of spiculation. Using ROC-analysis the classification performance of all features was evaluated on a dataset containing 54 mammograms. The convexity and concavity features lead to an  $A_z$  of 0.75 and 0.76, respectively. The combination of all features yield an  $A_z$  of 0.79.

Varela et al. [11] have divided the mass in different segments (interior region, border, and outer area) and computed different features for each segment. The border segment has been defined as a narrow band along the boundary covering all pixels within 1mm distance on both sides (inside and outside) from the mass boundary. From this segment three types of features were extracted, aimed at characterizing sharpness, micro-lobulation, and texture. They defined the interior segment of a mass as the

area inside its boundary. Three different contrast measures were extracted from this segment. In the outer mass segment two features were calculated which have been developed for the detection of so-called *stellate* distortions [22,23]. These two features are based on the statistical analysis of gradient orientations. Results were obtained using ROC-analysis on a dataset including 981 mammograms. For each segment an optimal subset was selected. The best classification performance was achieved by the border segment ( $A_z = 0.76$ ), followed by the outer ( $A_z = 0.75$ ) and interior segment ( $A_z = 0.69$ ).

te Brake et al. [24] have been proposed an approach which divides a suspicious lesion in an inner and an outer region. From these regions they calculated different contrast measures which are comparable with our measures for  $k=2$ . On two different datasets they showed that these features are able to discriminate between normal and malignant tissue. Interestingly, the results on the publicly available Digital Database for Screening Mammography (DDSM) (which was also used for our experiments) were worse than on the other database which contained cases taken from the Dutch screening program.

Although the task of characterizing mammographic masses has been widely discussed in the literature, only a few of the approaches considered the different gray level distributions in various regions of the lesion. In contrast to these approaches which used only two different regions, we divide the lesion in a various number of concentric regions and perform statistical analysis to them yielding a new and easy computable set of features which is capable of discriminating malignant and benign masses.

The rest of the paper is organized as follows. In Section 2.1 we describe the database of mammograms used in this study. In Section 2.2 the new feature set will be introduced and described. The results of our experiments are presented and discussed in Section 3. Finally a conclusion is given in Section 4.

## 2. MATERIALS AND METHODS

### 2.1 Database

In this study cases from the Digital Database for Screening Mammography (DDSM) have been used, which is publicly available from the University of South Florida [25]. The DDSM contains 2.620 cases, each consisting of two views (CC and MLO) of each breast. Furthermore, each case includes the patient's age, a BI-RADS<sup>TM</sup> compliant breast density rating, a lesion subtlety rating as well as BI-RADS<sup>TM</sup> compliant lesion attributes. The database is partitioned into several volumes which contain normal, benign, or malignant cases.

To minimize the influence of the digitalization process we have included all mammograms of cases from this database which were digitized with a Lumisys laser scanner and contained at least one biopsy proven lesion described as mass. Each case is provided with a ground truth file containing the lesion annotations. For each mammogram a manually drawn ROI is given which contains the mass. This ROI serves as localization and was drawn by an expert. Images containing artifacts (e.g. presence of clips or pencil marks in the area containing the lesion) have been removed from our database. The resulting database consisted of 750 regions (387 malignant, 363 benign) of 423 patients. All images have been digitized with 12 bit depth

and 50 µm pixel size. The distribution of these cases according to BI-RADS™ classification is shown in Table 1.

**Table 1. Distribution of cases in our database according to BI-RADS™ classification.**

BI-RADS™	Benign	Malignant	Overall
1	0	2	2
2	9	0	9
3	187	24	211
4	166	160	326
5	2	200	202

## 2.2 The proposed feature set

A predefined number of concentric regions specified by the distance to the margin of the mass respectively to the border of its segmentation is used to characterize the lesion. To be able to describe the surrounding tissue of a given mass, in a first step the segmentation of our CADx-system is enlarged dependent on the smallest axis of the delineated region. Applying the Chamfer-Distance-Transformation [26] to the grown segmentation we get for each pixel within the segmentation the distance to the border. With respect to the maximum measured distance  $maxValue$  the segmented lesion is separated into  $k$  concentric regions  $R$ . Figure 1 shows an example for  $k = 10$ .

Each region  $R_i$  contains all pixels  $p(x,y)$ , with

$$i(\max Value / k) \leq dist(p(x, y)) < (i + 1)(\max Value / k) \quad (1)$$

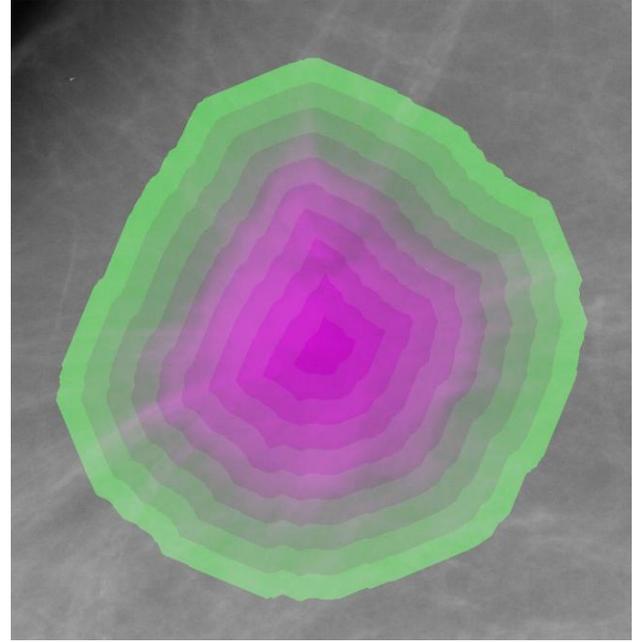
where  $dist(p(x,y))$  denotes the distance between a pixel  $p(x,y)$  and the border of the mass segmentation.

For each of the  $k$  regions  $R_i$  the corresponding gray level histogram  $H_R$  is computed and from this the mean gray value following the equation:

$$Mean_i = \frac{1}{\|H_{R_i}\|} \sum_{j=0}^{N-1} H_{R_i}(j)j, i = 0, \dots, k-1 \quad (2)$$

$\|H_{R_i}\|$  denotes the number of pixels belonging to region  $R_i$ .  $N$  denominates the number of gray levels and  $\|H_{R_i}\|$  denotes the number of pixels with the gray value  $j$  belonging to the region  $R_i$ . The  $k$  mean values are used as features. Additionally we derive another feature which describes the average intensity transition between the regions:

$$IntensityChange = \frac{1}{(k-1)} \sum_{i=1}^{k-1} |Mean_i - Mean_{i-1}| \quad (3)$$



**Fig 1: Example for the division of a mammographic lesion into  $k=10$  concentric regions. Regions containing surrounding tissue are marked green.**

To characterize the contrast of the lesion in comparison to its surrounding tissue, hence to describe the acutance of the lesion we are using two features which have also been derived from the  $k$  mean values and two features which are adapted from Varela et al.[11]:

$$MaxRatio_1 = \frac{\sum_{i=0}^{Q_1-1} Mean_{k-(i+1)}}{\sum_{i=0}^{Q_1-1} Mean_i} \quad (4)$$

$$MaxRatio_2 = \frac{\sum_{i=0}^{Q_2-1} Mean_{k-(i+1)}}{\sum_{i=0}^{Q_2-1} Mean_i} \quad (5)$$

$$ContrastMeasure_1 = \sum_{i=0}^{Q_1-1} \frac{(Mean_i - Mean_{k-(i+1)})^2}{\sigma_i^2 + \sigma_{k-(i+1)}^2} \quad (6)$$

$$ContrastMeasure_2 = \sum_{i=0}^{Q_2-1} \frac{(Mean_i - Mean_{k-(i+1)})^2}{\sigma_i^2 + \sigma_{k-(i+1)}^2} \quad (7)$$

$\sigma^2$  denotes the variance of the corresponding region and  $Q_1$  and  $Q_2$  are defined as follows:

$$Q_1 = \begin{cases} 1 & , k < 10 \\ \lfloor 0.1k \rfloor & , else \end{cases} \quad (8)$$

$$Q_2 = \begin{cases} Q_1 + 1 & , \lfloor 0.2k \rfloor \leq Q_1 \\ \lfloor 0.2k \rfloor & , else \end{cases} \quad (9)$$

To compute  $MaxRatio2$  and  $ContrastMeasure2$  the value of  $k$  has to be larger than 3. Furthermore, we compute three features from the gray level histograms which are characterizing the

difference of the gray level distribution inside and outside the lesion and the difference between the gray level distributions of neighboring regions. These features are defined as:

$$HistoDiff_1 = \frac{1}{Q_1} \sum_{i=0}^{Q_1-1} \sum_{j=0}^{N-1} \left| \frac{H_{R_i}(j)}{\|H_{R_i}\|} - \frac{H_{R_{k-(i+1)}}(j)}{\|H_{R_{k-(i+1)}}\|} \right| \quad (10)$$

$$HistoDiff_2 = \frac{1}{Q_2} \sum_{i=0}^{Q_2-1} \sum_{j=0}^{N-1} \left| \frac{H_{R_i}(j)}{\|H_{R_i}\|} - \frac{H_{R_{k-(i+1)}}(j)}{\|H_{R_{k-(i+1)}}\|} \right| \quad (11)$$

$$HistoDiff_{Global} = \sum_{i=1}^{k-1} \sum_{j=0}^{N-1} \left| \frac{H_{R_i}(j)}{\|H_{R_i}\|} - \frac{H_{R_{i-1}}(j)}{\|H_{R_{i-1}}\|} \right| \quad (12)$$

In total we are calculating  $k + 8$  features which characterize the intensity distribution inside and outside the lesion and also the intensity transition into the surrounding tissue.

### 2.3 Classification and feature selection

To evaluate the classification performance of each single feature we are using a Support Vector Machine (SVM). A SVM is a class of related supervised learners used for classification as well as regression. The main idea of SVMs is to map feature vectors which cannot be separated by a hyperplane in feature space into a higher dimensional space where an optimal separating hyperplane can be found. The implementation of the mapping is realized using so-called kernel functions. To be more exact, two parallel hyperplanes are constructed on each side of the optimal hyperplane that separates the data. The optimal separating hyperplane is then defined as the hyperplane that maximizes the distance between the two parallel hyperplanes. This approach is based on the assumption that a larger margin between these parallel hyperplanes results in a better generalization error of the classifier.

For each examined value of  $k$  we tried to find an optimal feature subset from the  $k + 8$  features. Hence, we have used a genetic algorithm (GA) for this feature selection task. GAs are known for their ability to efficiently search large search spaces with little or no a priori information on their structure. A GA is defined by a problem representation, genetic operators, and an objective function. The basic idea of a GA is to evolve a population of candidate solutions (individuals) towards a problem solution using genetic operators. The evolution starts with an initial population of individuals and evolves in generations. For each generation, the fitness of every individual is determined using the objective function. Individuals are then selected from the current population based on their fitness. They are modified using genetic operators (e.g. mutation and crossover) and form a new population for the next iteration of the algorithm.

### 3. RESULTS AND DISCUSSION

We have evaluated and compared the classification performance of the introduced features for  $k = 2n$  ( $n = 2, 3, \dots, 15$ ) using 10-fold cross validation and receiver operating characteristic (ROC) curve analysis. The area under the ROC curve  $A_z$  was used as a performance metric and was computed for each single feature. The classification was done using a SVM classifier.

Additionally, for each value of  $k$  denoting the number of concentric regions a preferable optimal feature subset has been selected using a GA. The  $A_z$  of this subset is regarded as the  $A_z$  of the appropriate extractor. Table 2 outlines the classification performance of the introduced features for  $k = 6, 12, 18, 24, 30$ . Details on SVMs and GAs can be found in Section 2.1.

Table 2 shows that most of the introduced features are able to discriminate malignant and benign masses. The mean features range from 0.57 to 0.69 with an average of 0.63. It is apparent that the classification performance is increasing with greater distance to the border of the segmentation. This means that the intensity and hence the density of regions in the center of a mass are more diagnostically conclusive than the intensity of regions containing outer parts of a mass or surrounding tissue. This observation can be seen for every  $k$ , independent of the image quantization and is displayed in Figure 2 for 12 bit images.

Another interesting observation is that for every investigated value of  $k$  the mean features calculated on 12 bit images lead to a better classification performance than the corresponding features calculated on 8 bit images. The average  $A_z$  is 0.64 and 0.62, respectively. Hence, the loss of information has a negative influence on the classification performance of these features. In contrast to that it is interesting to note that in average the features  $MaxRatio_1$  and  $MaxRatio_2$  perform better on 8 bit images. All other features seem to be independent of the image quantization.

Looking at the overall performance of the feature extractors we obtain  $A_z$  values ranging from 0.74 ( $k = 16$ ) to 0.76 ( $k = 26$ ) for 8 bit images and  $A_z$  values ranging from 0.76 ( $k = 16$ ) to 0.77 ( $k = 10$ ) for 12 bit images. This emphasizes the ability of each single feature extractor to discriminate benign and malignant masses. On 8 bit images we get over all  $k$  an average performance of  $0.75 \pm 0.01$  and on 12 bit images an average performance of  $0.77 \pm 0.00$ . Hence the loss of information also has a negative influence on the overall performance. The small standard regression shows that for the investigated range of  $k$  the choice of this parameter does not have a big influence on the discrimination ability. Figure 2 gives an overview of the classification performance for every investigated value of  $k$ .

### 4. CONCLUSION

In this work a set of new features designed to characterize the intensity transition from the center of a mammographic mass up to its surrounding tissue has been presented and evaluated on a subset of the DDSM reference data set. To characterize the lesion the proposed features use a predefined number ( $k$ ) of concentric regions defined by the distance to the mass margin respectively to the border of its segmentation. For each region the mean value was used as a feature. Additionally, features describing the *acutance* were derived from the mean values respectively from the corresponding gray level histograms. We have evaluated the classification performance for  $k = 2n$  ( $n = 2, 3, \dots, 15$ ). For each  $k$  an optimal feature subset was selected by a genetic algorithm. The  $A_z$  values of these subsets ranged from 0.74 ( $k = 16$ ) to 0.76 ( $k = 26$ ) on 8 bit images and from 0.76 ( $k = 16$ ) to 0.77 ( $k = 10$ ) on 12 bit images. The results show that the value of  $k$  has only a slight influence on the overall performance and that the classification performance of the mean features is increasing with greater distance to the border of the segmentation. Furthermore we have shown the eligibility of

these fast to compute features to help to discriminate benign and malignant masses. Although the classification performance is not sufficient to use these features alone for mammographic mass classification it is conceivable that a combination with other features might improve the overall classification performance of different CADx-approaches.

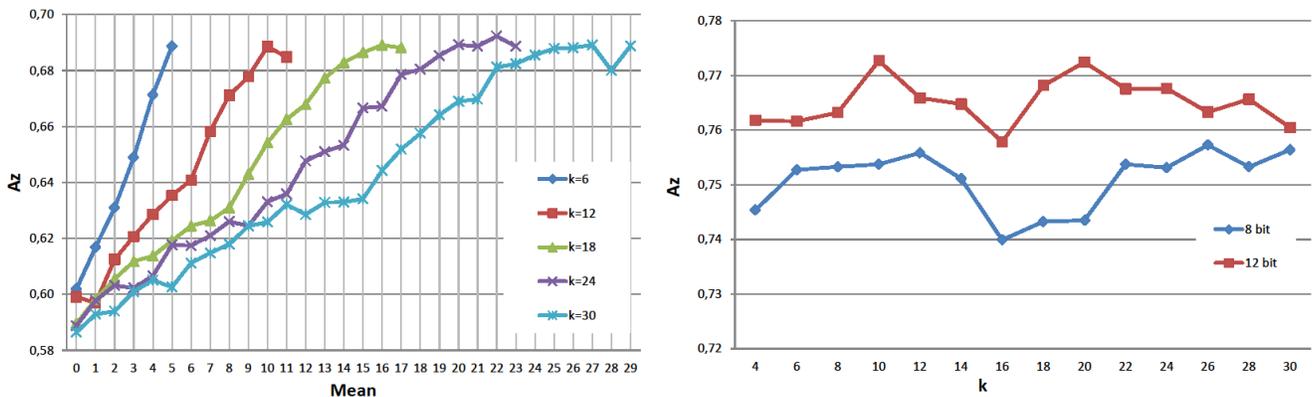
It is noticeable that a direct comparison with the results of other workgroups is hard to perform because the used databases are too different in composition. Especially the very good results of some workgroups reported in section 1.2 were determined on very small databases with less than 100 cases and might be worse on larger databases. Additionally, in [24] it was noted that differences in screening practice can have an influence on the results. They reported that their features performed better on database containing cases taken from the Dutch screening program in comparison to the DDSM. Hence, it would be desirable if more authors would use public available databases

such as the DDSM for their experiments to make an objective comparison possible.

Looking at the number of cases Shi et al. [20] and Varela et al. [11] used databases which are comparable to our database for their experiments. Compared to their results most of our feature sets performed slightly better (e.g. for  $k = 10$ ). te Brake et al. [24] used similar features and also the DDSM for the evaluation of their experiments. Unfortunately they employed FROC analysis to evaluate the performance of their features. Hence, a direct comparison is not possible.

## 5. CONFLICT OF INTEREST STATEMENT

We certify that there is no conflict of interest with any financial and personal relationships with other people or organizations regarding the material discussed in the manuscript.



**Fig 2: Left: Increase of the classification performance of the Mean features from the outer to the inner region for  $k = 6, 12, 18, 24, 30$  calculated on 12 bit images. Right: Overall performance of the feature sets for different  $k$  on 8 and 12 bit images.**

**Table 2: Classification performance ( $A_z$ ) of the introduced features for  $k = 6, 12, 18, 24, 30$  calculated on 8 and 12 bit images. Bold values represent those features which belong to the subsets selected by a genetic algorithm. The  $A_z$  values of these subsets can be seen in the last row of this table.**

Feature	$k = 6$		$k = 12$		$k = 18$		$k = 24$		$k = 30$	
	8 bit	12 bit								
$Mean_0$	0.58	0.60	0.57	0.60	0.58	0.59	0.57	0.59	0.57	<b>0.59</b>
$Mean_1$	0.60	<b>0.62</b>	0.58	0.60	0.59	0.60	0.58	0.60	0.58	<b>0.59</b>
$Mean_2$	<b>0.61</b>	<b>0.63</b>	0.60	<b>0.61</b>	0.59	0.61	0.57	0.60	0.57	0.59
$Mean_3$	0.63	0.65	0.60	0.62	0.59	0.61	0.59	0.60	0.58	0.60
$Mean_4$	<b>0.65</b>	<b>0.67</b>	0.61	<b>0.63</b>	0.60	0.61	0.59	<b>0.61</b>	<b>0.59</b>	0.61
$Mean_5$	0.66	0.69	<b>0.61</b>	<b>0.64</b>	0.60	0.62	0.59	0.62	0.60	0.60
$Mean_6$			0.63	0.64	0.61	0.62	<b>0.60</b>	0.62	0.59	0.61
$Mean_7$			<b>0.64</b>	<b>0.66</b>	<b>0.61</b>	<b>0.63</b>	0.60	0.62	0.60	0.61
$Mean_8$			0.65	0.67	0.62	0.63	0.61	0.63	0.61	0.62
$Mean_9$			0.65	0.68	0.62	0.64	0.61	<b>0.62</b>	0.60	<b>0.62</b>
$Mean_{10}$			0.66	0.69	0.64	<b>0.65</b>	<b>0.62</b>	0.63	0.61	0.63
$Mean_{11}$			0.66	0.68	0.64	0.66	0.62	0.64	0.61	0.63
$Mean_{12}$					<b>0.65</b>	0.67	0.63	0.65	0.61	0.63
$Mean_{13}$					<b>0.65</b>	0.68	<b>0.64</b>	0.65	<b>0.62</b>	<b>0.63</b>
$Mean_{14}$					0.66	0.68	0.64	0.65	0.62	<b>0.63</b>
$Mean_{15}$					0.66	0.69	0.64	<b>0.67</b>	0.62	0.63
$Mean_{16}$					0.66	0.69	0.65	0.67	0.63	0.64
$Mean_{17}$					0.66	0.69	0.65	0.68	<b>0.64</b>	<b>0.65</b>
$Mean_{18}$							0.66	0.68	0.64	0.66
$Mean_{19}$							0.66	0.69	0.65	0.66
$Mean_{20}$							0.66	0.69	0.65	0.67
$Mean_{21}$							0.66	0.69	<b>0.65</b>	0.67
$Mean_{22}$							0.66	0.69	0.66	0.68
$Mean_{23}$							0.67	0.69	0.66	0.68
$Mean_{24}$									0.66	<b>0.69</b>
$Mean_{25}$									0.66	<b>0.69</b>
$Mean_{26}$									0.67	0.69
$Mean_{27}$									0.66	0.69
$Mean_{28}$									0.66	0.68
$Mean_{29}$									0.66	<b>0.69</b>
<i>IntensityChange</i>	<b>0.65</b>	0.66	<b>0.65</b>	<b>0.65</b>	0.65	<b>0.64</b>	<b>0.65</b>	<b>0.64</b>	<b>0.64</b>	<b>0.65</b>
$MaxRatio_1$	0.62	0.61	0.62	0.61	0.62	0.61	0.62	0.62	0.61	0.61
$MaxRatio_2$	<b>0.61</b>	<b>0.61</b>	0.62	0.61	<b>0.62</b>	0.61	0.63	0.62	0.62	<b>0.61</b>
$HistoDiff_1$	0.67	0.67	0.67	0.67	0.66	0.66	0.66	0.66	0.66	0.65
$HistoDiff_2$	0.66	0.66	0.67	0.67	0.66	0.66	0.66	0.65	0.65	0.64
$HistoDiff_{Global}$	<b>0.65</b>	0.63	0.60	0.51	0.57	0.57	<b>0.53</b>	0.59	<b>0.54</b>	<b>0.59</b>
$ContrastMeasure_1$	<b>0.65</b>	<b>0.66</b>	0.65	0.65	0.65	0.65	<b>0.64</b>	0.65	<b>0.60</b>	<b>0.63</b>
$ContrastMeasure_2$	<b>0.65</b>	0.65	<b>0.65</b>	<b>0.65</b>	<b>0.65</b>	<b>0.65</b>	<b>0.64</b>	<b>0.65</b>	<b>0.64</b>	<b>0.63</b>
Opt. $A_z$	0.75	0.76	0.76	0.77	0.74	0.77	0.75	0.77	0.76	0.76

## 6. REFERENCES

- [1] Statistisches Bundesamt, "Gesundheitsberichterstattung des Bundes: Gesundheit in Deutschland", (2006). <http://www.gbe-bund.de/>.
- [2] Heywang-Köbrunner, S. H. and Schreer, I., [Bildgebende Mammadiagnostik: Untersuchungstechnik, Befundmuster, Differenzialdiagnose und Interventionen ], Georg Thieme Verlag, 2., überarbeitete und erweiterte Auflage ed. (2003).
- [3] Kopans, D., "The positive predictive value of mammography.", *Am. J. Roentgenol* 158, 521-526 (1992).
- [4] Elter, M. and Haßmeyer, E., "A knowledge-based approach to the cadx of mammographic masses", in [Medical Imaging: Computer-Aided Diagnosis ], 6915, 69150L:1-8, SPIE (2008).
- [5] Elter, M. and Held, C., "Semiautomatic segmentation for the computer aided diagnosis of clustered micro calcifications", in [Medical Imaging: Computer-Aided Diagnosis ], 6915, 691524:1-8, SPIE (2008).
- [6] Lehmann, T., Güld, M., Thies, C., Fischer, B., Spitzer, K., Keysers, D., Ney, H., Kohnen, M., Schubert, H., and Wein, B., "Content-based image retrieval in medical applications", *Methods. Inf. Med.* 43(4), 354-61 (2004).
- [7] Haralick, R., Shanmugam, K., and Dinstein, I., "Textural features for image classification", *IEEE Trans. on Systems, Man, and Cybernetics* 3(6), 610-621 (1973).
- [8] Galloway, M., "Texture analysis using gray level run lengths", *Computer Graphics and Image Processing* 4, 172-179 (1975).
- [9] Laine, A. and Fan, J., "Texture classification by wavelet packet signatures", *IEEE Trans. Pattern Anal. Mach. Intell.* 15, 1186-1191 (1993).
- [10] Mudigonda, N., Rangayyan, R., and Desautels, J., "Gradient and texture analysis for the classification of mammographic masses", *IEEE Trans. Med. Imaging* 19, 1032-1043 (2000).
- [11] Varela, C., Timp, S., and Karssemeijer, N., "Use of border information in the classification of mammographic masses", *Phys Med Biol* 51, 425-441 (2006).
- [12] Gupta, S. and Markey, M. K., "Correspondence in texture features between two mammographic views", *Med Phys* 32, 1598-1606 (2005).
- [13] Mavroforakis, M. E., Georgiou, H. V., Dimitropoulos, N., Cavouras, D., and Theodoridis, S., "Mammographic masses characterization based on localized texture and dataset fractal analysis using linear, neural and support vector machine classifiers", *Artif Intell Med* 37, 145-162 (2006).
- [14] Huo, Z., Giger, M. L., Vyborny, C. J., Bick, U., Lu, P., Wolverton, D. E., and Schmidt, R. A., "Analysis of spiculation in the computerized classification of mammographic masses", *Med Phys* 22, 1569-1579 (1995).
- [15] Huo, Z., Giger, M. L., Vyborny, C. J., Wolverton, D. E., Schmidt, R. A., and Doi, K., "Automated computerized classification of malignant and benign masses on digitized mammograms", *Acad Radiol* 5, 155-168 (1998).
- [16] Huo, Z., Giger, M. L., Vyborny, C. J., Wolverton, D. E., and Metz, C. E., "Computerized classification of benign and malignant masses on digitized mammograms: a study of robustness", *Acad Radiol* 7, 1077-1084 (2000).
- [17] Zheng, B., Lu, A., Hardesty, L. A., Sumkin, J. H., Hakim, C. M., Ganott, M. A., and Gur, D., "A method to improve visual similarity of breast masses for an interactive computer-aided diagnosis environment", *Med Phys* 33, 111-117 (2006).
- [18] Sahiner, B., Chan, H., Petrick, N., Helvie, M., and Goodsitt, M., "Computerized characterization of masses on mammograms: the rubber band straightening transform and texture analysis", *Med. Phys.* 25(4), 516-526 (1998).
- [19] Rangayyan, R. M., El-Faramawy, N. M., Desautels, J. E., and Alim, O. A., "Measures of acutance and shape for classification of breast tumors", *IEEE Trans Med Imaging* 16, 799-810 (1997).
- [20] Shi, J., Sahiner, B., Chan, H.-P., Ge, J., Hadjiiski, L., Helvie, M. A., Nees, A., Wu, Y.-T., Wei, J., Zhou, C., Zhang, Y., and Cui, J., "Characterization of mammographic masses based on level set segmentation with new image features and patient information", *Med Phys* 35, 280-290 (2008).
- [21] Rangayyan, R. M., Mudigonda, N. R., and Desautels, J. E., "Boundary modelling and shape analysis methods for classification of mammographic masses", *Med Biol Eng Comput* 38, 487-496 (2000).
- [22] Karssemeijer, N. and te Brake, G., "Detection of stellate distortions in mammograms", *IEEE Trans. Med. Imaging* 15(5), 611-619 (1996).
- [23] te Brake, G. and Karssemeijer, N., "Single and multiscale detection of masses in digital mammograms", *IEEE Trans. Med. Imaging* 18, 628-639 (1999).
- [24] Guido M te Brake, Nico Karssemeijer, and Jan H C L Hendriks. "An automatic method to discriminate malignant masses from normal tissue in digital mammograms". *Phys Med Biol*, 45:2843-2857, (2000).
- [25] Heath, M., Bowyer, K., Kopans, D., Moore, R., and Kegelmeyer, W., "The digital database for screening mammography", in [Proc's 5th Int. Workshop on Digital Mammography ], 212-218 (2001).
- [26] Butt, M. A. and Maragos, P., "Optimum design of chamfer distance transforms", *IEEE Trans Image Process* 7, 1477-1484 (1998).