



A classifier-combined method for grading breast cancer based on Dempster-Shafer evidence theory

Zhou Liu^{1#}, Fuliang Lin^{2,3#}, Junhui Huang², Xia Wu², Jie Wen¹, Meng Wang¹, Ya Ren¹, Xiaoei Wei⁴, Xinyu Song⁴, Jing Qin⁵, Elaine Yuen-Phin Lee⁶, Dan Shao⁷, Yixiang Wang⁸, Xiaoguang Cheng⁹, Zhanli Hu², Dehong Luo¹, Na Zhang²

¹Department of Radiology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Shenzhen, China; ²Paul C. Lauterbur Research Center for Biomedical Imaging, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China; ³School of Information Engineering, Zhengzhou University, Zhengzhou, China; ⁴Institute of Diagnostic and Interventional Radiology, Shanghai Jiao Tong University Affiliated Sixth People's Hospital, Shanghai, China; ⁵Centre for Smart Health, School of Nursing, The Hong Kong Polytechnic University, Hong Kong, China; ⁶Department of Diagnostic Radiology, Clinical School of Medicine, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong, China; ⁷Department of Nuclear Medicine, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China; ⁸Imaging and Interventional Radiology, Faculty of Medicine, The Chinese University of Hong Kong, Shatin Hong Kong SAR, China; ⁹Department of Radiology, Beijing Jishuitan Hospital, Beijing, China

Contributions: (I) Conception and design: Z Liu, X Wu, D Luo, N Zhang; (II) Administrative support: J Qin, EYP Lee, D Shao, Y Wang, X Cheng; (III) Provision of study materials or patients: Z Liu, J Wen, M Wang, Y Ren, D Luo; (IV) Collection and assembly of data: Z Liu, J Wen, M Wang, Y Ren, D Luo; (V) Data analysis and interpretation: Z Liu, F Lin, J Huang, X Wu, X Wei, X Song, Z Hu; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Na Zhang. Paul C. Lauterbur Research Center for Biomedical Imaging, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518110, China. Email: na.zhang@siat.ac.cn; Dehong Luo. Department of Radiology, National Cancer Center/National Clinical Research Center for Cancer/Cancer Hospital & Shenzhen Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, 113 Baohe Avenue, Longgang District, Shenzhen 518110, China. Email: 13926236152@163.com.

Background: Preoperative non-invasive histologic grading of breast cancer is essential. This study aimed to explore the effectiveness of a machine learning classification method based on Dempster-Shafer (D-S) evidence theory for the histologic grading of breast cancer.

Methods: A total of 489 contrast-enhanced magnetic resonance imaging (MRI) slices with breast cancer lesions (including 171 grade I, 140 grade II, and 178 grade III lesions) were used for analysis. All the lesions were segmented by two radiologists in consensus. For each slice, the quantitative pharmacokinetic parameters based on a modified Tofts model and the textural features of the segmented lesion on the image were extracted. Principal component analysis was then used to reduce feature dimensionality and obtain new features from the pharmacokinetic parameters and texture features. The basic confidence assignments of different classifiers were combined using D-S evidence theory based on the accuracy of three classifiers: support vector machine (SVM), Random Forest, and k-nearest neighbor (KNN). The performance of the machine learning techniques was evaluated in terms of accuracy, sensitivity, specificity, and the area under the curve.

Results: The three classifiers showed varying accuracy across different categories. The accuracy of using D-S evidence theory in combination with multiple classifiers reached 92.86%, which was higher than that of using SVM (82.76%), Random Forest (78.85%), or KNN (87.82%) individually. The average area under the curve of using the D-S evidence theory combined with multiple classifiers reached 0.896, which was larger

than that of using SVM (0.829), Random Forest (0.727), or KNN (0.835) individually.

Conclusions: Multiple classifiers can be effectively combined based on D-S evidence theory to improve the prediction of histologic grade in breast cancer.

Keywords: Evidence theory; computer-aided diagnosis (CAD); breast cancer; dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI); machine learning

Submitted Jun 25, 2022. Accepted for publication Jan 04, 2023. Published online Apr 20, 2023.

doi: 10.21037/qims-22-652

View this article at: <https://dx.doi.org/10.21037/qims-22-652>

Introduction

Breast cancer is a serious, life-threatening malignancy in women which accounts for 15.4% of all cancer-related deaths (1,2). Due to the underlying tumor heterogeneity, breast cancers of the same clinical stage can have completely different treatment responses and prognoses. Therefore, in this new era of precision medicine, there is a pressing clinical need for a noninvasive method that enables this tumor heterogeneity to be quantified so that individualized treatment can be tailored, and the treatment response can be promptly evaluated.

Among all the sequences in a routine breast magnetic resonance imaging (MRI) protocol, dynamic contrast-enhanced MRI (DCE-MRI) is the backbone. Not only is it the most sensitive method for detecting breast cancer (3) but, more importantly, it can also reflect the tumor hemodynamic heterogeneity, including micro-vessel density and vascular permeability differences (4).

In the clinical setting, the histologic grade for breast cancer is determined by assessing the glandular and tubular differentiation, nuclear pleomorphism, and mitotic activity, which are hallmark characteristics of tumor aggressiveness. In general, high-grade tumors have poorer differentiation and more dismal prognoses than do low-grade tumors (5). The preoperative histological grading of breast cancer is critical for clinical decision-making, including breast cancer surgery and subsequent treatment planning. Therefore, the ability to noninvasively predict the histological grade of breast cancer is critical.

Texture features acquired from breast DCE-MRI images can reflect the microvascular density and spatial permeability heterogeneity within a region of interest (ROI), which is useful for identifying different types of breast lesions (6). Also, pharmacokinetic parameters obtained by fitting the temporal signal intensity profile of T1-weighted images using a modified Tofts model (such as K^{trans}) can

reflect the hemodynamic heterogeneity of tumor tissue (7,8). Therefore, both DCE-MRI-based texture features and pharmacokinetic parameters have potential for the noninvasive preoperative histological grading of breast cancers.

With the widespread application of artificial intelligence in the field of medical image processing seen in recent years, texture feature analysis based on machine learning has achieved groundbreaking results in tumor classification tasks, such as outcome prediction (9), neoadjuvant chemotherapy response prediction (10), and molecular subtype classification (11). However, most of the studies to date have used only spatial or temporal features to achieve the best classification effect (9,10). Pharmacokinetic parameters obtained from DCE-MRI can be used to explore permeability and perfusion changes within a tumor, and thus could potentially be integrated into a conventional texture-based machine learning model to further improve the classification of tumors (11).

Dempster-Shafer (D-S) evidence theory is an inference method that can handle uncertain information (12). Different classifiers have different data processing abilities, and thus, varying levels of fault tolerance and applicability. Using D-S evidence theory, the classification effects of multiple classifiers can be combined to obtain more accurate classification results. In the fields of urban planning and natural image classification, D-S evidence theory was found to substantially improve the classification results of a model based on a single classifier by combining multiple classifiers [support vector machine (SVM) and k-nearest neighbor (KNN)], with a 20% to 30% increase in accuracy (13-15). Therefore, this combination could be expected to improve classification accuracy; however, its potential in the medical field has yet to be explored.

We hypothesized that breast cancers of different pathological grades show spatial and hemodynamic heterogeneity on DCE-MRI, and that this spatial and

Table 1 Clinical characteristics of the enrolled patients.

IBC-NST (n = 33)	Data
Age (years)	50.6±8.3
Histologic subtype	
IBC-NST	19
Mixed type (IBC-NST + DCIS)	14
Size in the largest diameter (≤2 cm/2–5 cm/>5 cm)	22/11/0
TNM stage	
T stage (T1/T2/T3/T4)	13/18/1/1
N stage (N0/N1/N2/N3)	16/5/7/5
M stage (M0/M1)	33/0
Histologic grade	
Grade I/II/III	13/12/8
Proliferation protein (ki-67 index)	
≤20%	18
>20%	15
Metastatic status of ALNs (positive/negative)	18/15
ER (positive/negative)	25/8
PR (positive/negative)	23/10
HER-2 (positive/negative)	17/16
Molecular subtypes	
Luminal A	2
Luminal B	22
HER-2 enriched	9
Triple negative	0
Perineural invasion	
Positive/negative	7/17
Not available	9
Vascular invasion	
Positive/negative	15/9
Not available	9

The data in the table are presented as mean ± standard deviation or frequency. IBC-NST, invasive breast carcinoma of no special type; DCIS, ductal carcinoma in situ; TNM, tumor, node, metastasis; ALN, axillary lymph node; ER, estrogen receptor; PR, progesterone receptor; HER-2, human epidermal growth factor receptor.

hemodynamic heterogeneity can be quantified using texture features and hemodynamic parameters. Therefore, in this study, we investigated and evaluated the performance of a machine learning model based on the combination of pharmacokinetic parameters and texture features derived from DCE-MRI, using D-S evidence theory, in predicting the preoperative histologic grade of breast cancer. We present the following article in accordance with the TRIPOD reporting checklist (available at <https://qims.amegroups.com/article/view/10.21037/qims-22-652/rc>).

Methods

Study participants

This study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and was approved by the ethics committee of Cancer Hospital Chinese Academy of Sciences, Shenzhen Hospital. The requirement for individual consent was waived for this retrospective analysis.

Thirty-three female patients with a histologically confirmed diagnosis of breast cancer who underwent baseline DCE-MRI between December 2019 and July 2020 were retrospectively included. The inclusion criteria were as follows: (I) the histology of each breast lesion was confirmed as invasive breast carcinoma of no special type by biopsy or surgical specimen; (II) each breast cancer lesion had a histologic grade of I–III, and the maximal diameter of the primary breast lesion exceeded 1 cm; (III) conventional MRI and DCE-MRI were performed within the 2 weeks before treatment; and (IV) in one DCE-MRI scan, ten phases had been acquired. The DCE-MRI images were reviewed by two board-certified radiologists (MW and YR). Images of poor quality, including those with overt signal loss, motion artifacts, or geometric distortion, were excluded from the analysis. The clinical characteristics of the enrolled patients are summarized in *Table 1*.

MRI protocols

The MRI examinations were performed on a 3T MR scanner (Discovery MR 750w, General Electric Healthcare, Waukesha, WI, USA) using an eight-channel phase-array breast coil. All the participants had undergone DCE-MRI using a three-dimension T1-weighted fast spoiled

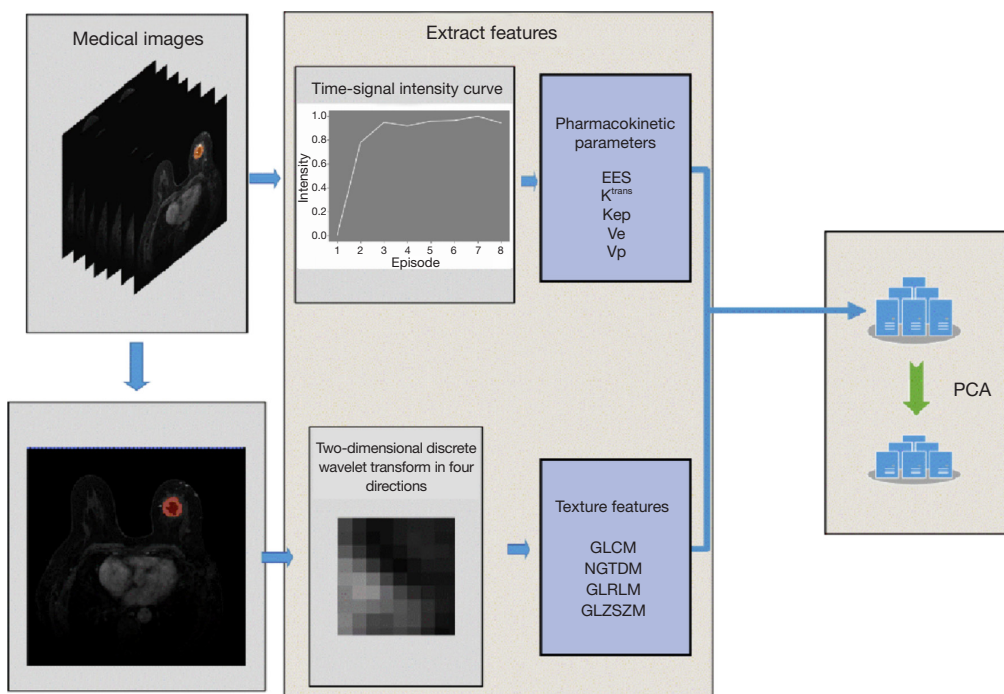


Figure 1 Data processing flow diagram. EES, extravascular extracellular space; K^{trans} , volume transfer constant; K_{ep} , rate constant; V_e , extracellular volume fraction of the imaged tissue; V_p , fractional plasma volume; GLCM, gray-level co-occurrence matrix; NGTDM, neighborhood gray-tone difference matrix; GLRLM, gray-level run-length matrix; GLSZM, gray-level size zone matrix; PCA, principal component analysis.

gradient-echo sequence with the imaging parameters set as follows: repetition time/echo time = 4.5/2.1 ms, field of view = 360 mm × 360 mm, image matrix = 320 × 320, slice thickness = 1.4 mm with no gap, and flip angle = 12°. A total of 10 periods (35–55 s per period), with one pre-contrast and nine post-contrast dynamic periods, were obtained. For dynamic MRI, gadoteric acid meglumine (Gd-DOTA, Dotarem, Guerbet, Roissy CdG Cedex, France) was injected into the antecubital vein at a rate of 2 mL/s (total dose, 0.1 mmol/kg of body weight) using a power injector (Ulrich, Germany) via a 20-gauge needle, which was followed by a 20 mL saline flush.

Image analysis

All the MRI images were transferred to an in-house MATLAB platform developed by X.W. for analysis. A total of 489 slices with breast cancer lesions (171 grade I, 140 grade II, and 178 grade III) were used for analysis. All the lesions were initially segmented manually on the third phase of the DCE-MR images by a radiologist (M.W. or

Y.R., each with 5 years of experience in breast imaging) and then reviewed by a senior radiologist (J.W., with more than 10 years of experience in breast imaging). All three radiologists were blinded to the histology of the lesions. Any discrepancies were resolved through discussion. For each slice, 78 original texture features were extracted using the first-order histogram (FH; 18 features), gray-level co-occurrence matrix (GLCM; 23 features), neighborhood gray-tone difference matrix (NGTDM; 5 features), gray-level run-length matrix (GLRLM; 13 features), gray-level size zone matrix (GLSZM; 13 features), and 6 shape features. The data processing flowchart is shown in *Figure 1*.

A total of 390 (78 × 5) texture features, including the original features and the features after two-dimensional discrete wavelet transform in four directions [components of the approximation (CA), components of the diagonal detail (CD), components of the horizontal detail (CH), components of the vertical detail (CV)], were finally extracted for each slice. Five pharmacokinetic parameters, including the volume transfer constant of contrast agent leaked into the extravascular extracellular space from the

plasma (K^{trans}), the rate constant of contrast agent reflux to the plasma (K_{ep}), the fractional extravascular extracellular space volume (V_e), the fractional plasma volume (V_p), and the area under the time-intensity curve (AUC), were calculated using the modified Tofts model (7). The time-transformed contrast concentration profile was calculated first using the temporal signal intensity profile (see Eq. [1]), after which the kinetic parameters (K^{trans} and V_e) were calculated by fitting the temporal concentration profile using a modified two-compartment kinetic model (see Eq. [2]).

$$\Delta S(t) \approx \frac{e^{-TR/T_{10}}(1 - \cos \alpha)TRr_1C(t)}{1 - e^{-TR/T_{10}}[1 + \cos \alpha(1 - e^{-TR/T_{10}})(1 - TRr_1C(t))]} \quad [1]$$

where TR is the repetition time (msec), T_{10} is the so-called native relaxation time, $C(t)$ is the tissue contrast agent concentration, r_1 ($\text{mM}^{-1}\text{s}^{-1}$) is the longitudinal contrast agent relaxation coefficient, and α is the flip angle.

$$C_i(t) = \frac{K^{trans} \cdot C_p(0)}{K_{el} - K_{ep}} [e^{-K_{ep}(t-t_0)} - e^{-K_{el}(t-t_0)}] + C_p(0) \cdot V_p \cdot e^{-(t-t_0)K_d} \quad [2]$$

where k_{el} is the excretion rate and represents the loss of contrast agent in the system, $C_p(0)$ denotes the contrast agent concentration in the vascular space at time $t=0$, and t_0 is the time before the injection of the contrast agent.

The data set was first divided into a training set and a test set at a ratio of 7:3. Next, to reduce the dimension, principal component analysis (PCA) was performed for all pharmacokinetic parameters and texture features. Then, the reduced dimension features were used in training sets for the three different types of classifiers (random forest, SVM, and KNN). This paper employed two-fold cross-validation in the training phase.

Study method

In D-S evidence theory, all answers that can be thought of for a question are put into the identification framework. All the answers within the identification framework are mutually exclusive, whereas the answer to the question is unique. Each answer can be regarded as a proposition, and each proposition has a degree of confidence called the basic probability assignment (BPA, also known as the M function). The $m(A)$ reflects the degree of reliability of A .

For the same problem, different evidence sources will yield different BPA values, and the BPA values of different evidence sources can be processed using the orthogonal sum to obtain a new basic probability distribution function (16). For the i^{th} classifier, first, a sample was selected for training,

the best recognition rate $\mathcal{E}_r^{(i)}$ of the classifier was found, and then the probability distribution function was designed according to the distribution of the sample category. The BPA estimation formula is as follows:

$$m_i(j) = \frac{V_i(j,t) \cdot \mathcal{E}_r^{(i)}}{\sum_{t \in U} V_i(j,t) + \alpha} \quad [3]$$

where $V_i(j,t)$ is the number of test samples that actually belong to the t^{th} class when they are judged to be the j^{th} class by the i^{th} classifier, and α is the uncertainty factor.

The D-S evidence theory synthesis rule can be interpreted as follows:

$$M(k) = K^{-1} \times \sum_{\substack{l \cap j \cap h = k \\ l, j, h \in D}} m_1(l) \times m_2(j) \times m_3(h) \quad [4]$$

where $K = \sum_{l \cap j \cap h \neq \emptyset} m_1(l) \times m_2(j) \times m_3(h)$; and $m_1(l)$, $m_2(j)$, and $m_3(h)$ represent the basic probability value that the classification result of the first classifier is pathology grade l , grade j , and grade h , respectively.

The BPAs of the three classifiers, namely m_1 , m_2 , and m_3 , were calculated. Each classifier has three classification levels, and the correct and incorrect prediction results can be represented using a 3×3 matrix. Elements within the matrix represent the BPA for that classifier. The BPA matrices obtained by the three classifiers are denoted by M_1 , M_2 , and M_3 , respectively. M_1 , M_2 , and M_3 are input into Eq. [4] to obtain the synthetic result G . Assuming that the combined BPA value is $G(k)$, the formula for the decision-making process is as follows:

$$j = \arg \max_{k \in D} \{G(k)\} \quad [5]$$

which satisfies $G(j) - \max_{k \in D, k \neq j} G(k) \geq \sigma_1$, where j is the decision-making category. According to the recognition rate of SVM, Random Forest, and KNN, this paper considered the distribution of training samples to construct the BPA and conducted evidence synthesis and decision-making based on the synthesis rules of D-S evidence theory. Receiver operating characteristic curve analysis was performed for each model built. The overall algorithm flowchart is shown in Figure 2.

Results

Clinical characteristics of the participants

A total of 489 slices with breast cancer lesions (171 grade I, 140 grade II, and 178 grade III) were used for analysis. Figure 3 shows representative MRI images of the first,

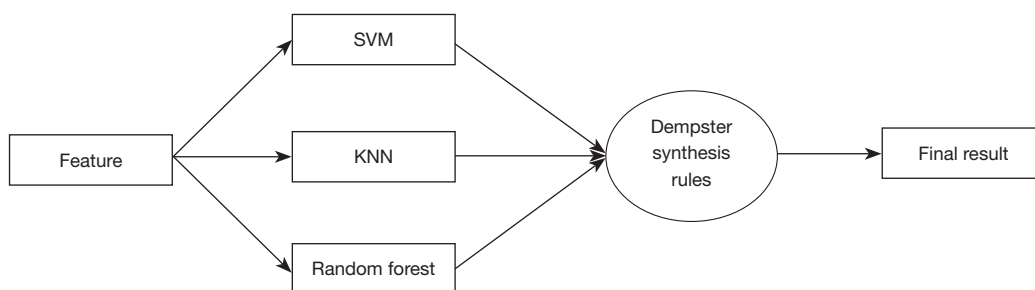


Figure 2 The overall algorithm flow diagram. SVM, support vector machine; KNN, k-nearest neighbor.

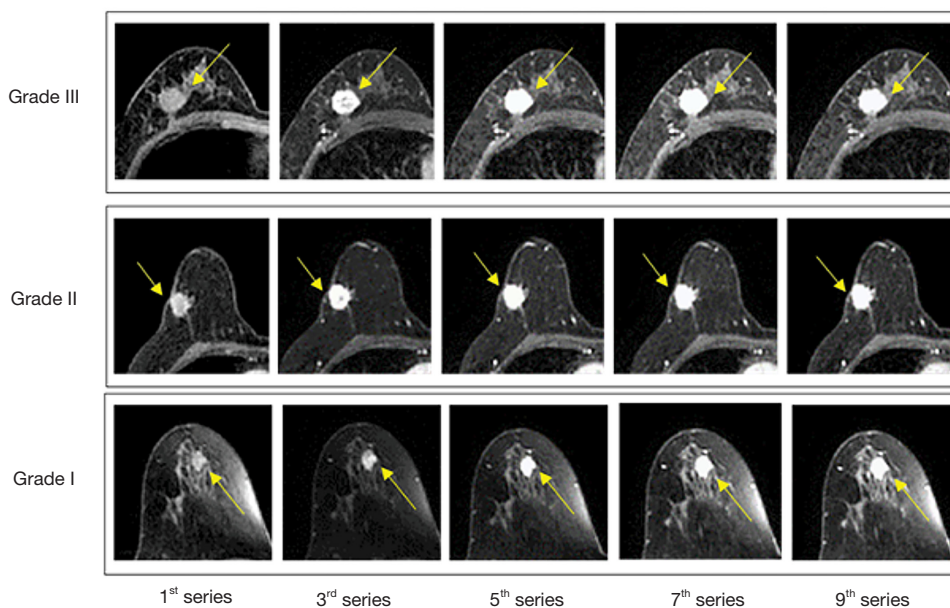


Figure 3 The 1st, 3rd, 5th, 7th, and 9th phases after contrast for grade I–III breast cancer.

third, fifth, seventh, and ninth phases acquired in the corresponding acquisition times for grade I–III breast cancer.

Comparison between the single classifier results and the classification results based on D-S evidence theory

After reducing the dimension of 390 texture features extracted from slices, 113 new features were selected. The data were divided into training and testing sets at a ratio of 7:3 and then sent to three different classifiers (SVM, random forest, and KNN) for training and testing. The performance of the machine learning techniques was evaluated in terms of accuracy, sensitivity, and specificity

(Table 2).

From the case classification results using a single classifier (Table 2), KNN had the highest accuracy among the three classifiers. However, in terms of specificity and sensitivity in classifying breast cancers of different pathological grades, the performance of KNN was not always the best. For example, in the recognition of pathology grade II, the sensitivity of the SVM classifier was 1.0, which was much higher than that of KNN or Random Forest; however, in the classification of pathology grade I, the sensitivity of the SVM classifier was lower than that of KNN and Random Forest. These observations show that the different classifiers have different advantages in classifying slices at different pathological levels. Table 3 shows the results of the

Table 2 Histologic grading evaluation index results of single classifiers (support vector machine, Random Forest, and K-nearest neighbor)

Histologic grade	SVM		Random Forest		KNN	
	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity
Grade I	0.716	0.990	0.761	0.836	0.894	0.967
Grade II	1.000	0.629	0.816	0.738	0.791	0.810
Grade III	0.934	0.798	0.804	0.774	0.936	0.830
Accuracy	0.828		0.789		0.878	

SVM, support vector machine; KNN, K-nearest neighbor.

Table 3 Grading evaluation index results of D-S evidence theory combined with multiple classifiers

Evaluation indicator	Grade I	Grade II	Grade III
Specificity	0.896	0.976	0.814
Sensitivity	0.972	0.796	0.833
Accuracy	0.929		

D-S, Dempster-Shafer.

classification evaluation index of the D-S evidence theory combined with multiple classifiers. The accuracy rate of the D-S evidence theory was 0.929; its sensitivity for the different grades was 0.896 for grade I, 0.976 for grade II, and 0.814 for grade III; and its specificity for the different grades was 0.972 for grade I, 0.796 for grade II, and 0.833 for grade III.

According to the experimental results, the receiver operating characteristic curve of each level was drawn, and the average AUC value was calculated (*Figure 4*). The average AUC value of the D-S-based method reached 0.896, which was higher than that of the methods using a single classifier (SVM: 0.829, Random Forest: 0.727, and KNN: 0.835).

Discussion

In this study, we investigated the performance of a combination of multiple classifiers based on D-S evidence theory in classifying the histologic grade of breast cancer and compared it with the performances of three classical single classifiers. Our proposed method outperformed the single classical classifiers, which indicates that D-S evidence theory can effectively combine the effects of multiple classifiers to obtain higher classification accuracy.

The three classifiers included in this study have different recognition situations in different categories and have

unique advantages for different histologic grades. Using D-S evidence theory combined with multiple classifiers eliminated the blind spots of the single classifiers for certain pathology levels by combining the recognition advantages of each classifier for those pathology levels, to achieve a balanced recognition of all pathology levels (17,18). According to the training results and the sample size, the BPA values of the three classifiers were calculated and the evidence was synthesized, and from this, the decision was made. Since the KNN method mainly relies on a limited number of neighboring samples around to determine the class to which it belongs, rather than the method of discriminating class domains, it is more advantageous for classifying sample sets with more intersection or overlap of class domains. Moreover, KNN involves no assumptions on the data, is highly accurate, and is insensitive to outliers (19). Therefore, these factors might explain our observation that KNN had the best overall performance among the three classifiers.

Regarding SVM classifiers, K secondary classification SVM classifiers need to be trained for the sample data of the K categories during training. The sample data of the i^{th} SVM subcategory is marked as a positive category, while the sample data not belonging to category I are marked as a negative category. Therefore, the negative category sample data appears for each level of the pathological section, and the data of the negative category samples are much larger

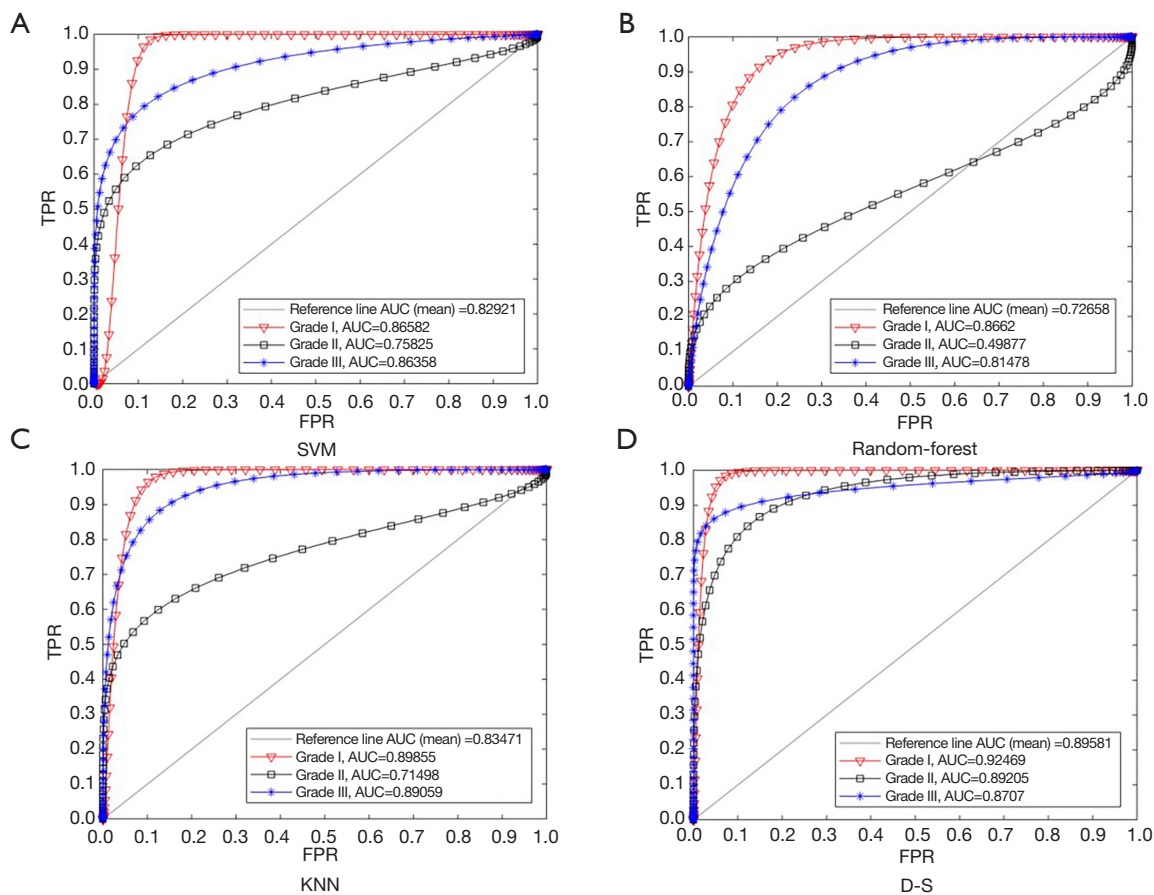


Figure 4 Receiver operating characteristic curves for each pathological grade. Receiver operating characteristic curves of (A) SVM, (B) Random Forest, (C) KNN, and (D) multiple classifiers based on D-S evidence theory. SVM, support vector machine; KNN, k-nearest neighbor; AUC, area under the curve; TPR, true positive rate; D-S, Dempster-Shafer.

than those of the positive category samples. Consequently, class imbalance appears, and this situation tends to become more serious with the increase of training data. For example, in histologic grade II recognition, the specificity of the SVM classifier reaches 1 but its sensitivity is only 0.6286.

As previously reported, using D-S evidence theory combined with multiple classifiers can achieve a balanced recognition of classification tasks. Consistently, we observed that the accuracy and average AUC of the fusion strategy proposed in this paper were higher than those of single classifiers for breast histologic grade classification, indicating that our fusion strategy can integrate the advantages of different classifiers to improve the classification performance. Therefore, the results of this study demonstrate that based on D-S evidence theory, the knowledge of multiple classifiers can be fused at the same time using the Dempster synthesis rules, and the prediction

results of multiple predictors (SVM, KNN, and Random Forest) can be fused to improve the classification accuracy.

This study had a number of limitations. First, the sample size in this study was relatively limited. However, our study mainly focused on verifying the proof of concept of using D-S evidence theory to merge and take advantage of multiple classical classifiers. To augment our dataset, two-dimensional slices were used instead of three-dimensional volume to extract texture features and pharmacokinetic parameters. Second, the PCA method was applied in this study to solve the “curse of dimensionality” problem. However, PCA is a linear reduction method, which might have resulted in the loss of some valuable information. In the future, the appropriateness of PCA should be evaluated and more feature selection methods should be explored. Third, this paper applied the D-S formula to the SVM, Random Forest, and KNN classifiers without analyzing the

independence between the different classifier-based models; this aspect should be further investigated in the future. Fourth, this paper did not apply bootstrapping techniques to obtain confidence intervals of the classification performance. Finally, we arbitrarily chose KNN, SVM, and Random Forest, as three of the most commonly used classifiers. Many other commonly used classifiers, such as naïve Bayes, logistic regression, and decision tree classifiers, were not explored in our study. However, the primary aim of this work was to explore the potential benefit of D-S evidence theory in taking advantage of multiple classifiers, not to explore the performance of specific classifiers themselves.

Conclusions

The results of this study suggest that the effective combination of multiple classifiers under D-S evidence theory can improve histological grade prediction in breast cancer.

Acknowledgments

Funding: The study was partially supported by the Key Laboratory for Magnetic Resonance and Multimodality Imaging of Guangdong Province (grant No. 2020B1212060051); the Key Technology and Equipment R&D Program of Major Science and Technology Infrastructure of Shenzhen (grant No. 202100102 and 202100104); the Guangdong Innovation Platform of Translational Research for Cerebrovascular Diseases, the Shenzhen Basic Research Program (grant No. KCXFZ202002011010360); Shenzhen Clinical Research Center for Cancer (grant No. [2021] 287); and the Shenzhen High-level Hospital Construction Fund.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <https://qims.amegroups.com/article/view/10.21037/qims-22-652/rc>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://qims.amegroups.com/article/view/10.21037/qims-22-652/coif>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all

aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013) and was approved by the ethics committee of Cancer Hospital Chinese Academy of Medical Sciences, Shenzhen Hospital. The requirement for individual consent for this retrospective analysis was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Momenimovahed Z, Salehiniya H. Epidemiological characteristics of and risk factors for breast cancer in the world. *Breast Cancer* (Dove Med Press) 2019;11:151-64.
2. Chen H, Wu K, Wang M, Wang F, Zhang M, Zhang P. A standard mastectomy should not be the only recommended breast surgical treatment for non-metastatic inflammatory breast cancer: A large population-based study in the Surveillance, Epidemiology, and End Results database 18. *Breast* 2017;35:48-54.
3. Abramson RG, Li X, Hoyt TL, Su PF, Arlinghaus LR, Wilson KJ, Abramson VG, Chakravarthy AB, Yankeelov TE. Early assessment of breast cancer response to neoadjuvant chemotherapy by semi-quantitative analysis of high-temporal resolution DCE-MRI: preliminary results. *Magn Reson Imaging* 2013;31:1457-64.
4. Holli K, Lääperi AL, Harrison L, Luukkaala T, Toivonen T, Ryymin P, Dastidar P, Soimakallio S, Eskola H. Characterization of breast cancer types by texture analysis of magnetic resonance images. *Acad Radiol* 2010;17:135-41.
5. Yankeelov TE, Lepage M, Chakravarthy A, Broome EE, Niermann KJ, Kelley MC, Meszoely I, Mayer IA, Herman CR, McManus K, Price RR, Gore JC. Integration of quantitative DCE-MRI and ADC mapping to monitor treatment response in human breast cancer: initial results. *Magn Reson Imaging* 2007;25:1-13.
6. Turkki R, Byckhov D, Lundin M, Isola J, Nordling S,

- Kovanen PE, Verrill C, von Smitten K, Joensuu H, Lundin J, Linder N. Breast cancer outcome prediction with tumour tissue images and machine learning. *Breast Cancer Res Treat* 2019;177:41-52.
7. Tahmassebi A, Wengert GJ, Helbich TH, Bago-Horvath Z, Alaei S, Bartsch R, Dubsky P, Baltzer P, Clauser P, Kapetas P, Morris EA, Meyer-Baese A, Pinker K. Impact of Machine Learning With Multiparametric Magnetic Resonance Imaging of the Breast for Early Prediction of Response to Neoadjuvant Chemotherapy and Survival Outcomes in Breast Cancer Patients. *Invest Radiol* 2019;54:110-7.
 8. Wu T, Sultan LR, Tian J, Cary TW, Sehgal CM. Machine learning for diagnostic ultrasound of triple-negative breast cancer. *Breast Cancer Res Treat* 2019;173:365-73.
 9. Chitalia RD, Kontos D. Role of texture analysis in breast MRI as a cancer biomarker: A review. *J Magn Reson Imaging* 2019;49:927-38.
 10. Henderson S, Purdie C, Michie C, Evans A, Lerski R, Johnston M, Vinnicombe S, Thompson AM. Interim heterogeneity changes measured using entropy texture features on T2-weighted MRI at 3.0 T are associated with pathological response to neoadjuvant chemotherapy in primary breast cancer. *Eur Radiol* 2017;27:4602-11.
 11. Machireddy A, Thibault G, Tudorica A, Afzal A, Mishal M, Kemmer K, Naik A, Troxell M, Goranson E, Oh K, Roy N, Jafarian N, Holtorf M, Huang W, Song X. Early Prediction of Breast Cancer Therapy Response using Multiresolution Fractal Analysis of DCE-MRI Parametric Maps. *Tomography* 2019;5:90-8.
 12. Chen C, Wang JZ, Chang HY, Li J. Lane Detection of Multi-visual-features Fusion Based on D-S Theory. *Chin Contr Conf*. 2011:3047-52.
 13. Tang YC, Wu DD, Liu ZJ. A new approach for generation of generalized basic probability assignment in the evidence theory. *Pattern Anal Appl* 2021;24:1007-23.
 14. Liu ZG, Huang LQ, Zhou K, Denoeux T. Combination of Transferable Classification With Multisource Domain Adaptation Based on Evidential Reasoning. *IEEE Trans Neural Netw Learn Syst* 2021;32:2015-29.
 15. Zhou R, Fang WP, Wu JS. A risk assessment model of a sewer pipeline in an underground utility tunnel based on a Bayesian network. *Tunnelling and Underground Space Technology* 2020;103:103473.
 16. Kumar V, Bhatele M. Proceedings of all India seminar on biomedical engineering 2012 (AISOBE 2012). Springer New Delhi 2012.
 17. Singh R, Vatsa M, Noore A, Singh SK. DS theory based fingerprint classifier fusion with update rule to minimize training time. *IEICE Electronics Express* 2006;3:429-35.
 18. Kisku DR, Tistarelli M, Sing JK, Gupta P. Face Recognition by Fusion of Local and Global Matching Scores using DS Theory: An Evaluation with Uni-classifier and Multi-classifier Paradigm. 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Miami, FL, USA, 2009, pp. 60-65.
 19. Yan WF, Wu GX, Li CZ, Zhou L. Evidence Theory of One-dimensional Compression KNN Classification Method. *Advanced Materials Research* 2011;143-144:1337-41.

Cite this article as: Liu Z, Lin F, Huang J, Wu X, Wen J, Wang M, Ren Y, Wei X, Song X, Qin J, Lee EYP, Shao D, Wang Y, Cheng X, Hu Z, Luo D, Zhang N. A classifier-combined method for grading breast cancer based on Dempster-Shafer evidence theory. *Quant Imaging Med Surg* 2023;13(5):3288-3297. doi: 10.21037/qims-22-652