Applied Soft Computing Journal 91 (2020) 106230

Contents lists available at ScienceDirect

Applied Soft Computing Journal

journal homepage: www.elsevier.com/locate/asoc

Fuzzy-rough assisted refinement of image processing procedure for mammographic risk assessment

Yanpeng Qu [a,b,*], Qilin Fu [a], Changjing Shang [b], Ansheng Deng [a], Reyer Zwiggelaar [b], Minu George [b], Qiang Shen [b]

[a] Information Technology College, Dalian Maritime University, Dalian, 116026, China
[b] Department of Computer Science, Institute of Mathematics, Physics and Computer Science, Aberystwyth University, SY23 3DB Aberystwyth, UK

ARTICLE INFO

Article history:
Received 3 June 2019
Received in revised form 27 February 2020
Accepted 8 March 2020
Available online 23 March 2020

Keywords:
Image processing
Feature extraction
Fuzzy-rough sets
Mammographic risk assessment

ABSTRACT

The use of computer aided diagnosis (CAD) systems, which are computer based tools for the automatic analysis of medical images such as mammogram and prostate MRI, can assist in the early detection and diagnosis of developing cancer. In the process of CAD for mammogram, the task of image processing (IP) plays a fundamental role in providing promising diagnostic results, by exploiting high-quality features extracted from the mammographic images. Normally, an IP procedure for mammographic images involves three mechanisms: region of interest (ROI) extraction, image enhancement (IE) and feature extraction (FE). However, an improper utilisation of IE may lead to an inferior composition of the features due to unexpected enhancement of any irrelevant or useless information in ROI. In order to overcome this problem, a fuzzy-rough refined IP (FRIP) framework is presented in this paper to improve the quality of mammographic image features hierarchically. Following the proposed framework, the ROI of each mammographic image is segmented and enhanced locally in the area of the block which is of the highest value of fuzzy positive region (FPR). Here, FPR implies a positive dependency relationship between the block and the decision with regard to the given feature set. The higher a block's FPR value the more certain its underlying image category. To attain a high quality of the image enhancement procedure, the winner block will be further improved by a multi-round strategy to create a pool of IE results. As such, for a mammographic image, after embedding the candidate enhanced blocks into the original ROI, the respectively extracted features from the locally enhanced ROI are compared against each other on the basis of the value of FPR. A given image is therefore represented by a set of features which are supported by the premier FPR among all of the resulting extracted features. The quality of the extracted features by FRIP is compared against that of those directly extracted from the original images, from the globally enhanced images or from the randomly locally enhanced images in performing classification tasks. The experimental results demonstrate that the mammographic risk assessment results based on the features achieved by the proposed framework are much improved over those by the alternatives.

1. Introduction

Breast cancer is the most common cause of death for middle-aged women. According to the estimates by the International Cancer Research Institute of the World Health Organisation, more than one million women in the world die from breast cancer every year [1]. According to a report of the Global Cancer Statistics 2018 released by the International Agency for Research on Cancer (IARC), which is part of the World Health Organization (WHO), as almost 2.09 million new cases were diagnosed, breast cancer caused an estimated 0.63 million deaths in 2018 [2]. Along with the increased levels of breast cancer occurrence, the level of early detection has also been recorded thanks to screening by the use of mammographic imaging and expert opinion. Mammography is a specific type of imaging that uses a low-dose X-ray system to result in a high-shrinkage and high-resolution movie to examine breasts [3]. However, even expert radiologists may sometimes fail to detect a significant proportion of mammographic abnormalities. Also, a large number of detected abnormalities are usually discovered to be benign following further medical investigation.

Existing mammographic computer-aided diagnosis (CAD) systems [4–6] concentrate on the detection and classification of mammographic abnormalities, based on the features extracted from suspicious regions. Such techniques on the digitalised mammogram help the doctors to discover the existence of tumour,

* Corresponding author at: Information Technology College, Dalian Maritime University, Dalian, 116026, China.
E-mail address: yanpengqu@dlmu.edu.cn (Y. Qu).

https://doi.org/10.1016/j.asoc.2020.106230
1568-4946/

especially useful when being applied in early stages of cancer development. Among the components for implementing CAD systems, image processing (IP) plays a vital role in providing a set of high-level features for promising diagnosis result. In general, the process of IP for mammogram involves three mechanisms: the region of interest (ROI) extraction, image enhancement (IE) and feature extraction (FE). Mammogram often contains tags and artefacts that appear as high-intensity markers on a dark background. The noises, vascular and glandular tissues may affect the search for abnormal areas and lead to a poor classification accuracy. Additionally, in mammogram, chest muscles are present in the area of objects with various noises. Such information also adversely affects the outcome of FE. Therefore, FE is normally not appropriate to be applied to a mammographic image as a whole, but it may be suitable for certain specific areas or ROI [7].

An ROI includes samples within a dataset identified for a particular purpose. For medical imaging, ROI is the most commonly used as a particular portion which is of concern during a diagnosis or of interest during research. ROI can be defined as the approximate outlining for the object of interest or as a rectangular region that contains both the object of interest and certain background. To extract ROI, a given image is normally denoised first. In order to avoid any potential negative influence of the microstructure that may appear in certain regions, the breast region is smoothed using a median filter, as done in [8]. Then, fuzzy c-means clustering [9] may be utilised to divide pixel grey values into two separate categories: adipose tissue and dense tissue. In the step of removing the background, techniques such as the so-called modal open image function [10] can be applied. Subsequently, a segmentation method based on area growth is often implemented to prevent large-scale segmentation around the seed with only the suspected mass present in the mammogram is segmented as the ROI [11]. Histogram may be employed to enhance the image and the Sobel detector may be performed to support edge detection for preprocessing and image segmentation as with the work of [12]. In [13], ROI regions are located and extracted on the basis of the maximally inscribed circle and centroid methods. In addition, an algorithm automatically generates a class number that can partition mammogram into the best areas as ROI regions can be found in [14]. Similarly, an algorithm that uses tetrolet filter to reduce the speckle noise and the active contour method based on statistical features to automatically segment breast lesions to obtain an ROI can be found in [15]. Methods for automatic ROI construction have also been proposed to serve as an alternative to the manual process of ROI labelling in CAD, for breast lesions [16,17].

The purpose of IE is to improve both the identifiability of target features and the quality of images to meet the analysis needs of specific application scenarios [18]. The commonly used IE methods include the following: (1) Histogram equalisation: the histogram of the original image is converted into an image with a probability density of 1 by the use of integral probability density function, thereby achieving the effect of improving contrast [19] (e.g., through dynamic recursive sub-image histogram equalisation [20]). (2) Median filtering: the intermediate value is taken after sorting out the grey values of the pixels in the centre of a given window, therefore being itself a smoothing filter to eliminate noise and achieve enhancement [21]. (3) Wavelet transform: the relevant wavelet parameters are modified as needed in image decomposition and reconstruction, enabling multi-resolution analysis that helps capture spatial domain localisation characteristics, while focusing on the details of the interested object in the image [22].

From the enhanced ROI features are extracted. Commonly adopted FE methods can be grouped into the following four categories: (1) Statistical method, a typical representative of which is to conduct texture feature analysis through grey level co-occurrence matrix (GLCM) [23], and another is to extract texture representations by manipulating image autocorrelation functions (e.g., to extract the characteristic parameters such as the thickness and directionality of the texture by calculating the energy spectrum function of the image) [24]. (2) Geometry method, which offers means for texture feature analysis based on the theory of texture primitives, including two influential algorithms: the Voronio checkerboard feature method [25] and the structural method [26]. (3) Model-based method, which is based on the structural model of a given image, with the parameters of the model being used as texture features, typically including random field models (e.g., Markov random field [27] and Gibbs random field model [28]). (4) Signal processing method, which works by the use of linear transformation or filters to obtain texture features expressed in the transformed domain while being subject to certain domain constraints (e.g., energy criteria) [29].

For conventional approaches to medical IP, an ROI region is normally cut from the original image for processing in its entirety. In such an ROI, not all information is useful or related to decision-making. A certain zone of the ROI may possess more decisive information to the decision label than the rest. If the process of IP ignores the relative significance of different parts of the ROI, the subsequent operations, i.e., IE and FE may lead to an opposite effect to the quality of the features from the design intention, due to the enhancement and prominence of irrelevant or useless information. This may mislead the composition of the extracted features, thereby adversely impacting upon the following tasks of data analysis, such as feature selection and diagnosis decision making. In order to address this important problem, this paper presents a fuzzy-rough refined IP (FRIP) framework with an aim to improve the quality of mammographic image features hierarchically.

In this work, the ROI of a mammographic image is acquired by the successful means as described in [30]. The popular sliding window (SW) algorithm [31] is adopted to achieve a segmentation, namely a cover of the resulting ROI. The significance of each window block will be evaluated by the value of fuzzy positive regions (FPR) [32,33], with respect to the features extracted using grey level co-occurrence matrix (GLCM) [23] within the given window. Note that the FPR values imply a positive dependency relationship between the objects and the decisions given the extracted feature sets. Thus, the window contains the highest value of FPR is most confirmative to the image label and the most preferable to represent the entire ROI. Such a window (which has the highest FPR value) is further enhanced through the use of a pulse coupled neural network (PCNN) [34]. In particular, to obtain high quality features, the winner window is repeatedly enhanced by PCNN with different thresholds each time when a pool of IE results is created.

In implementation, for a mammographic image, after embedding the candidate enhanced blocks into the original ROI, the respectively extracted features from the locally enhanced ROI are compared against each other on the basis of the values of their FPR. The original image is then represented by the set of features which are obtained from the premier FPR among all extracted feature sets. Furthermore, the mammographic risk assessment results based on the extracted features by FRIP is compared against those based on the features directly extracted from the original image, the globally enhanced image, and the locally enhanced images using randomly chosen windows. The classifiers adopted in this work for verification include J48 [35], JRip [36], PART [37], AdaBoostM1 [38], RandomForest (RF) [39]. The experimental results demonstrate that the mammographic risk assessment outcomes based on the features achieved by the proposed framework are much better than those by the

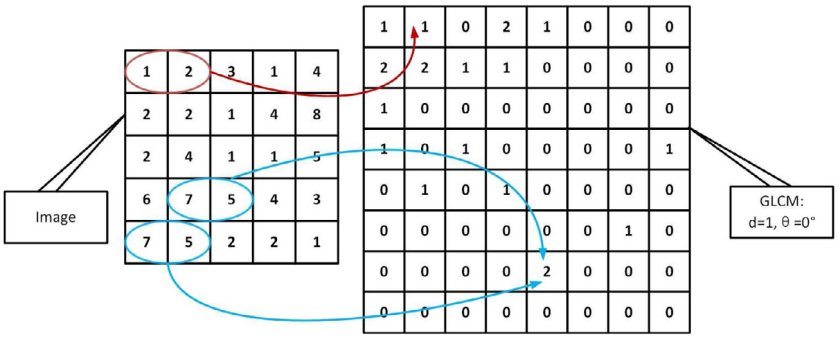


Fig. 1. Example of GLCM computation with distance d =1 and angle $\theta = 0°$, of an 8 grey level image.

alternative features, in terms of both classification accuracy and kappa coefficient [40]. In order to examine the applicability of the proposed approach, the results based on the features obtained by FRIP with different configurations of the sliding window are illustrated also.

The remainder of this paper is structured as follows. The related work on mammogram feature extraction is outlined in Section 2. The improved approach is presented and discussed in Section 4. In Sections 3 and 5, the experimental data and results are described respectively. The paper is concluded in Section 6, with a brief discussion regarding important further work.

2. Background

This section reviews the related approaches to FE, IE and image segmentation that are employed in this work. Moreover, the concept of fuzzy-rough sets which will be used as the underlying computational theory is introduced as well.

2.1. Grey level co-occurrence matrix

The algorithm of grey level co-occurrence matrix (GLCM) refers to a popular method for describing texture by studying the spatial correlation characteristics of grey scales [23,41]. As its name implies, running GLCM is to obtain the co-occurrence matrix by calculating on the values of the grey image. Then the eigenvalues of the matrix are obtained to represent the features of the image. GLCM helps capture comprehensive information of the image grey scales with respect to direction, adjacent interval and variation amplitude. It forms the basis for analysing the local patterns of images and their arrangement rules. As a powerful tool for feature extraction, GLCM is usually used to extract texture features, including contrast, texture entropy, correlation, variance [41]. GLCM can be run for different distances and orientations [42]. Fig. 1 shows the way to generate GLCM features given an image, where the grey level of each pixel ranges from 1 to 8 (see Fig. 1a). The scale of the resulting matrix is 8×8, as shown in Fig. 1b. Each element $G(i, j)$ of this matrix records the number of occurrences that the grey level i is located at a distance d from a grey level j and in the horizontal adjacent direction (i.e., being oriented by an angle $\theta = 0°$). For instance, in Fig. 1a, the ordered tuple of grey levels 1 and 2 only occurs once from two horizontally adjacent pixels. In this case, the value of $G(1, 2)$ in the resulting GLCM matrix is 1. Also, the value of the $G(7, 5)$ of GLCM is 2 indicates that the occurrence of the ordered tuple of grey levels 7 and 5 is observed twice.

2.2. Sliding window algorithm

In the context of IP, a sliding window [31] is a rectangular region of fixed width and height that slides across an image. For each of such regions, certain measures are applied to determine if the window contains any information of interest.

Given an image whose size is $m \times n$, the number of the sliding windows of a size $\tilde{m} \times \tilde{n}$ is calculated by

$$num = \left(\frac{m - \tilde{m}}{\lambda} + 1\right) \times \left(\frac{n - \tilde{n}}{\lambda} + 1\right), \tag{1}$$

where, λ is the step size of each sliding. This method slides the entire image in equidistant steps with a fixed-size sliding window, and performs detection within each sliding window. As such, its most significant strength is that the miss detection rate is extremely low because it slides the entire image without missing any position. As shown in Fig. 2, the step size λ is set to 48 pixels, the window size is set to 160×160 pixels, so the given 256×256 ROI is traversed to result in nine areas.

During the procedure of searching for the optimal block, the size of the sliding window should be set appropriately. If the window is too small, the result of the feature calculation is not going to be influential for decision-making; if the window is too large, the small target area will be enlarged, resulting in a decrease in target positioning accuracy. Note that if all windows are mutually disjoint. Thus, they can constitute a segmentation, or otherwise a cover, of the image.

2.3. Pulse coupled neural network model

Pulse coupled neural network (PCNN) is an iterative IE model, which simulates the response of mammalian visual cortex neurons to visual signals. PCNN has proven to be effective in many applications, e.g., for improving the brightness or enhancing the edge of an image, and for making the texture details more prominent [43]. Inspired by this observation, PCNN is adopted herein to process mammogram images to make their texture representation clearer and more recognisable.

The framework of PCNN can be regarded as a single-layer two-dimensional network, in which each neuron corresponds to each pixel of the input image [44]. Fig. 3 shows the PCNN neuron model, which includes a receptive field, a nonlinear modulation field, and a pulse generator. The function of each component is introduced below.

1. For the pixel whose coordinate in an image is (i, j), the input of the receptive field consists of a feedback input F_{ij} and a linear connection input L_{ij}, as defined by

$$F_{ij}[n] = S_{ij}[n] \tag{2}$$

$$L_{ij}[n] = e^{-\alpha_L} L_{ij}[n-1] + V_L \sum W_{ijkl} Y_{kl}[n-1], \tag{3}$$

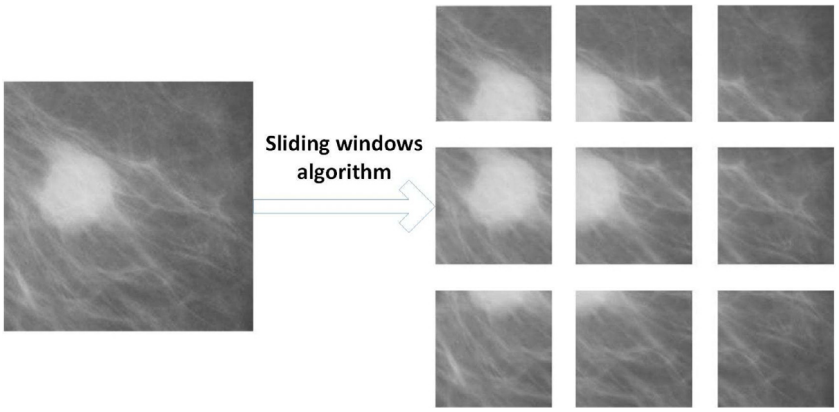


Fig. 2. Sliding window to capture image and local enhancement.

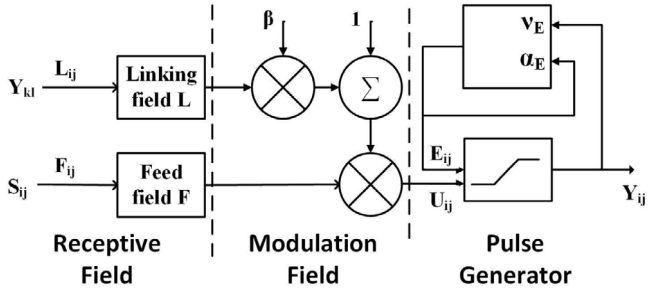


Fig. 3. A PCNN neuron model.

where, as the external input signal, n is the number of pulse ignition iterations; F_{ij} is equal to the value S_{ij} of the grey level of this pixel; L_{ij} is the connection input of the neighbourhood neuron, which is obtained by a weighted summation of Y_{kl} and W_{ijkl}; Y_{kl}, initialised as 0, is the output of the neighbouring neuron (acting as the input of its neighbour); W_{ijkl} is the internal connection matrix, and stores the weighting coefficients for the neighbouring neuron connections; α_L is the attenuation constant of L_{ij}; and V_L is the intrinsic potential of the connection input.

2. As the output of the modulation field, U_{ij} is the internal activity item formed by the signal of two input channels F_{ij} and L_{ij} through the modulation, as defined below:

$$U_{ij}[n] = F_{ij}[n](1 + \beta L_{ij}[n]), \tag{4}$$

where, β is the connection strength coefficient.

3. In the pulse generator, U_{ij} is compared with a dynamic threshold E_{ij} to produce the output pulse Y_{ij}. When E_{ij} exceeds the internal activity item U_{ij}, the pulse generator is turned off and the pulse is stopped. Then, the threshold begins to decrease exponentially. When the dynamic threshold is lower than U_{ij}, the pulse generator is turned on again and the neuron is fired or activated, and a pulse or pulse sequence is produced as the output in the following way [45]:

$$E_{ij}[n] = e^{-\alpha_E} E_{ij}[n-1] + V_E Y_{ij}[n-1], \tag{5}$$

$$Y_{ij}[n] = \begin{cases} 1, & U_{ij}[n] > E_{ij}[n], \\ 0, & otherwise. \end{cases} \tag{6}$$

where, V_E is the amplitude constant; α_E is the time decay constant of the dynamic threshold E_{ij}, determining the number of iterations in a cycle where all pixels are processed; and Y_{ij} is the pulse output function of PCNN [46].

With the use of the output of pulse generator, the grey value of each pixel is enhanced according to Eqn. (7).

$$EnhI_{ij} = (\ln(Bri) - \alpha_E(\tilde{n} - 1))Y_{ij}, \tag{7}$$

where $EnhI_{ij}$ represents the grey value of pixel (i, j) in the enhanced image [34]; Bri is the maximum brightness value in the original image; and $\tilde{n}$ records the ignition time of this PCNN neuron. It can also be seen from Eqn. (7) that by adjusting the value of α_E, different IE effects can be obtained. In order to ensure that E_{ij} is attenuated sufficiently slowly to distinguish adjacent grey levels by different ignition timings, the α_E value should be carefully set, which may be done empirically.

2.4. Fuzzy positive region

In a fuzzy-rough set (FRS) [32,33,47], there are two types of approximation: the lower approximation as the fuzzy positive region (FPR) and the upper approximation as the complement of the fuzzy negative region. The former is defined as the set of those objects which can be said with certainty to belong to the concept to be approximated, and the latter is defined as the set of objects which either definitely or possibly belong to the concept to be approximated (given a reference equivalence partition of the universe of discourse). Normally, the concept to be approximated refers to the decision information of the datasets. The difference between the upper and lower approximation is the area known as the fuzzy boundary region, representing the area of uncertainty. When the boundary region is empty, there is no uncertainty regarding the concept which is being approximated and all objects belong to the subset of objects of interest with full certainty.

Definitions for the fuzzy lower and upper approximations can be found in [32,33], where a T-transitive fuzzy similarity relation is used to approximate a fuzzy concept X:

$$\mu_{\underline{R_P}X}(x) = \inf_{y \in \mathbb{U}} I(\mu_{R_P}(x, y), \mu_X(y)), \tag{8}$$

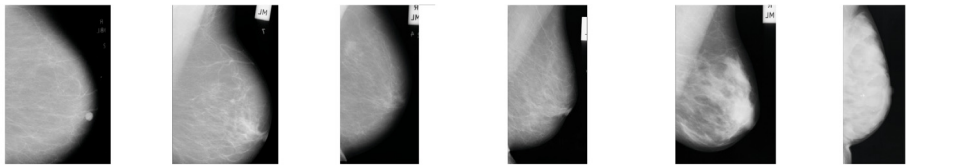
$$\mu_{\overline{R_P}X}(x) = \sup_{y \in \mathbb{U}} T(\mu_{R_P}(x, y), \mu_X(y)), \tag{9}$$

where $\mathbb{U}$ is a nonempty set of finite objects (the universe of discourse); I is a fuzzy implicator; T is a T-norm; and R_P is the fuzzy similarity relation induced by the subset of features P:

$$\mu_{R_P}(x, y) = T_{a \in P}\{\mu_{R_a}(x, y)\}, \tag{10}$$

with $\mu_{R_a}(x, y)$ being the degree to which the objects x and y are regarded to be similar with respect to feature a.

Given the above, the positive region regarding a fuzzy concept X and a set of attributes Q which induces equivalence relations



(a) mdb098 (b) mdb084 (c) mdb304 (d) mdb252 (e) mdb116 (f) mdb318

Fig. 4. Example mammograms, where (a) I, *SCC* 0%, Pattern II, N1; (b) II, *SCC* 0%–10%, Pattern III, P1; (c) II, *SCC* 11%–25%, Pattern III, P1; (d) II, *SCC* 26%–50%, Pattern I, P1; (e) III, *SCC* 51%–75%, Pattern IV, P2; (f) IV, *SCC* >75%, Pattern V, DY.

over $\mathbb{U}$ is defined by

$$\mu_{POS_{R_P}(Q)}(x) = \sup_{X \in \mathbb{U}/Q} \mu_{\underline{R_P}X}(x), \tag{11}$$

The value of Eq. (11) not only indicates the degree of object x belonging to the FPR, but also illustrates the capacity of object x implying the decision. The object with a higher value of FPR is more affirmatively related to the decision regarding the given feature set. In this work, FPR is used as the gauger to select the optimal local information of ROI and henceforth, the best features of each mammographic image in the proposed IP framework.

3. Mammographic dataset used

The experimental data used in this paper is derived from images extracted from the Mammographic Image Analysis Society (MIAS) database [48]. It includes a set of Medio-Lateral-Oblique (MLO) left and right mammogram of 161 woman (322 samples). The spatial resolution of the image is 50 μm×50 μm, which is quantised to 8 bits with a linear optical density in the range 0–3.2. As with the literature, mammographic risk assessment criteria are herein based on BI-RADS [49], Boyd [50], Tabár [51], Wolfe [52] (see Fig. 4 for examples).

In particular, BI-RADS [49] is used to category a mammogram into one of four classes: (*1*) BI-RADS I: Breast density is low; (*2*) BI-RADS II: There exists some fibroglandular tissue; (*3*) BI-RADS III: Breast density is high; (*4*) BI-RADS IV: Breast is extremely dense. In [53], it is reported that associations between BI-RADS I–IV and breast carcinoma (adjusted for weight) in postmenopausal women of which the risks are 1, 1.6, 2.3 and 4.5, respectively

Boyd [50] introduced a quantitative classification of mammographic density. It is based on the proportion of dense breast tissue relative to the overall breast area. The classification is known as Six-Class-Categories (SCC) where the density proportions are: Class1: 0%, Class2: (0%–10%), Class3: (10%–25%), Class4: (25%–50%), Class5: (50%–75%) and Class6: (75%–100%). The increase in the level of breast tissue density is associated with an increase in the risk of developing breast cancer. The increase in the level of breast tissue density has been associated with an increase in the risk of developing breast cancer, specifically the relative risk for SCC 1-6 are 1, 1, 1.9, 2.2, 4.6 and 7.1, respectively [50].

Tabár [51] described breast composition of four building blocks: nodular density, linear density, homogeneous fibrous tissue and radiolucent adipose tissues. These blocks also define mammographic risk classification. In particular, the following patterns are defined, with Patterns I–III corresponding to lower breast cancer risk and Patterns IV–V relating to higher risk: (*1*) Pattern I mammograms are composed of 25, 16, 35 and 24% of the four building blocks, respectively; (*2*) Pattern II has approximate compositions of 2, 14, 2 and 82%; (*3*) Pattern III is quite similar in composition to Pattern II, except that the retroareolar prominent ducts are often associated with periductal fibrosis; (*4*) Pattern IV is dominated by prominent nodular and linear densities, with compositions of 49, 19, 15, and 17%; (*5*) Pattern V is dominated by extensive fibrosis and is composed of 2, 2, 89 and 7% of the building blocks, respectively.

Wolfe [52] used the following four categories to recognise mammogram: (*1*) N1 mainly concludes fatty tissue and a few fibrous tissue stands; (*2*) P1 shows a prominent duct pattern, where a beaded appearance can be found either in the subareolar area or in the upper axillary quadrant; (*3*) P2 indicates server involvement of a prominent duct pattern which may occupy from one-half up to all of the volume of the parenchyma, often with the connective tissue hyperplasia producing coalescence of ducts in some areas; and (*4*) DY features a general increase in density of the parenchyma and there may, or may not, be a minor component of prominent duct. These four groups have an occurrence rate of developing breast cancer of 0.1, 0.4, 1.7, 2.2, respectively.

4. Fuzzy-rough refined image processing

Through an integrated use of the techniques introduced above, this section presents a fuzzy-rough refined IP (FRIP) framework for mammographic risk assessment. Fig. 5 illustrates the flowchart of the proposed scheme, consisting of three key procedures:

1. Image pre-processing and ROI extraction.
2. ROI segmentation and local evaluation.
3. ROI local enhancement and feature fusion.

4.1. Image pre-processing and ROI extraction

The mammographic images are pre-processed to extract the breast tissue region for further analysis. The breast region is denoised by removing the pectoral muscles and other artefacts, such as the orientation tags and the adhesive tapes from the image. Since most of the dense tissues and parenchymal patterns are located within the breast fibroglandular disk area, it is expected that the features will only be extracted from such regions. Segmentation of the fibroglandular disk region can help to extract the tissue characteristics of the right region. Take the BI-RADS criterion for instance, in most cases the areas outside the fibroglandular region contain fatty tissue regardless of their BI-RADS density class as most of the dense tissue patterns develop in the fibroglandular region. So extracting similar features outside the fibroglandular region does not provide discriminative features for tissue density classification. An appropriate ROI is extracted from each mammographic image for the fibroglandular disk area, of a 256 × 256 pixel size, as shown in Fig. 6. For extracting the fibroglandular disk region, the longest perpendicular distance from the breast boundary is considered which is usually the distance starting from the nipple area and then, a parallel distance line from the breast boundary is considered. The intersection point is regarded as the central point of the ROI region. With this central point, an ROI of a 256 × 256 pixel size is extracted as the sample fibroglandular disk region [30].

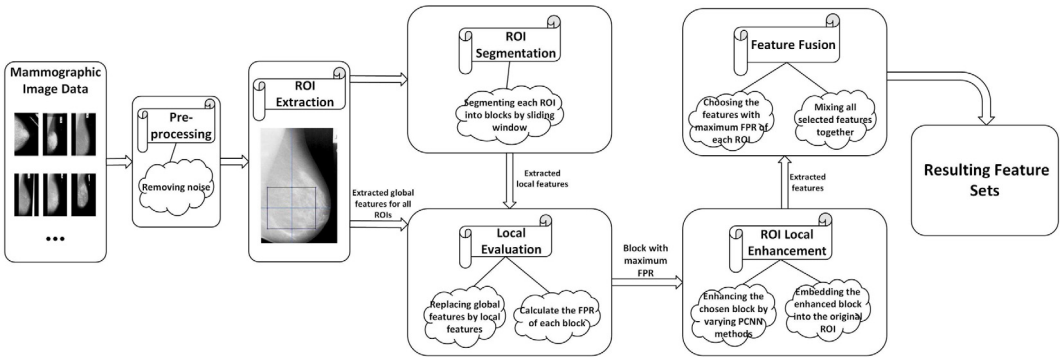


Fig. 5. Architecture of proposed framework.

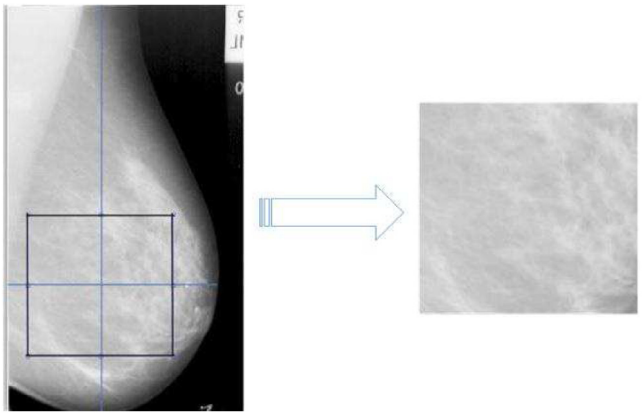


Fig. 6. Extraction of fibroglandular disk area from mammogram.

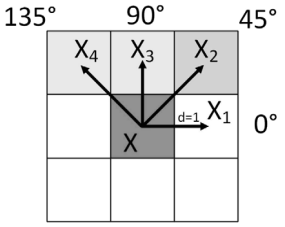


Fig. 7. Pixels X_1, X_2, X_3 and X_4 are located at a distance $d = 1$ from the central pixel X following the angles: $\theta = 0°$, $\theta = 45°$, $\theta = 90°$ and $\theta = 135°$.

4.2. ROI segmentation and local evaluation

In the proposed framework, for capturing local information of the initially extracted ROI, the sliding window algorithm is employed to generate a segmentation or a cover of the ROI region. For the purpose of making the experiment more comprehensive and credible, the sizes of the sliding window are assigned to be 160×160, 176×176, 192×192, 208×208, 224×224, respectively. Note that if the step size λ of each sliding is a relatively large number, the differences between the blocks may be missed. On the contrary, if λ is rather small, the process of going throughout the entire image with the sliding window will be time-consuming. Given the sizes of the sliding window used in this paper, the step size λ of each sliding is set to 16 moderately. In so doing, with the sliding windows whose sizes are 160×160, 176×176, 192×192, 208×208 and 224×224, a mammographic image will be segmented into 49, 36, 25, 16 and 9 blocks, respectively. In order to evaluate the significance of each block in an ROI, from the original ROI regions and their window blocks, the features are extracted by the GLCM algorithm. As shown in Fig. 7, the GLCM results are generated in four directions: 0°, 45°, 90°, 135°, and the distance is set to equalling to 1 pixel.

Given the element $G(x, y)$ of the resulting GLCM, the summations of the ith row and jth column are defined as shown in Eqn. (12) and Eqn. (13), respectively.

$$G_x(i) = \sum_{j=1} G(i, j). \tag{12}$$

$$G_y(j) = \sum_{i=1} G(i, j). \tag{13}$$

Let μ_x, μ_y, θ_x, θ_y be the respective means and standard deviations of G_x and G_y. The types of feature extracted in this work include contrast, entropy, correlation, inverse difference moment (IDM) and angular second moment (ASM), which are introduced as follows [54].

- Contrast (Con):

$$Con = \sum_i \sum_j (i-j)^2 * G(i, j). \tag{14}$$

- Entropy (Ent):

$$Ent = -\sum_i \sum_j G(i, j) * log(G(i, j)). \tag{15}$$

- Correlation (Cor):

$$Cor = \left[\sum_i \sum_j (i * j) * G(i, j) - \mu_x * \mu_y \right] / \sigma_x * \sigma_y. \tag{16}$$

- Inverse Difference Moment (IDM):

$$IDM = \sum_i \sum_j \frac{1}{1 + (i-j)^2} G(i, j). \tag{17}$$

- Angular Second Moment (ASM):

$$ASM = \sum_i \sum_j G(i, j)^2. \tag{18}$$

With the use of the above features, the block that has the highest value of FPR will be chosen to receive further enhancement. That is, with regard to the proposed framework, the ROI of each mammographic image is segmented and enhanced locally in the area of the highest FPR value. To ensure high quality of the image enhancement procedure, the winner block which enjoys the highest FPR value will be further improved by a multi-round strategy to create a pool of image enhancement results. Therefore, for a

mammographic image, after embedding the candidate enhanced blocks into the original ROI, the respectively extracted features from the locally enhanced ROI will be compared against each other in terms of the FPR value. As a result, the original image will be represented by the set of features which are entailed by the premier FPR among all extracted features.

Following the example in Fig. 2, the FPR value of each block is calculated and marked in Fig. 8. It can be seen that block 4 has the highest FPR = 0.6848. As such, it will be chosen as the winner to receive further enhancement. Moreover, blocks 1, 2, 5, 7 and 8, which are the neighbouring blocks of block 4, achieve higher FPR values than blocks 3, 6, 9. This observation helps verify the rationality of the proposed method to choose the winner block.

4.3. ROI local enhancement and feature fusion

As indicated previously, the value of the time decay constant α_E should be determined carefully in order to ensure that the dynamic threshold E_{ij} can be attenuated sufficiently slowly to distinguish adjacent grey levels in different ignition timings. Thus, the IE process on the chosen block is implemented with a multi-round strategy by PCNN which is devised to have a different value of α_E in each round, to create a pool of IE results. In so doing, the number of the sets of the candidate features to represent each mammographic image is the same as that of the distinct values of α_E for the PCNN algorithm. As shown in Fig. 9, all of the candidate enhanced windows will be embedded into the original ROI to produce the features by GLCM for use in further evaluation.

In the proposed framework, the appraisal of the candidate extracted features from the locally improved ROI is conducted with the use of FPR as well. Note that the features extracted from the entire original ROI, which have been achieved in the process of evaluating the window blocks, are employed to represent the rest of the image. In addition, the winner set of features associated with each mammographic image can be achieved from the different rounds of IE process. Ultimately, the fusion of these sets of features will be utilised as the ultimate FE results for the mammographic image dataset. Overall, the proposed FRIP framework is outlined in Alg. 1.

Algorithm 1 Fuzzy-rough Refined Image Processing Framework

Input:
Data: Mammographic Image Dataset
α_E: time decay constant;
Max: number of rounds for updating α_E.
Output: X_{best}: resulting feature sets produced by FRIP.

Image pre-processing and ROI extraction on *Data*
Extracting feature dataset X_0 from all ROIs
Initialisation: $\tau = 0$, $num = |Data|$, $i = 1$
// *ROI Segmentation and Local Evaluation*
while $i \leq Max$ **do**
  for $j = 1;\ j \leq num;\ j++$ **do**
    Segmenting the jth ROI into n blocks by sliding window
    for $k = 1;\ k \leq n;\ k++$ **do**
      x_{jk} = features extracted from kth window of the jth ROI
      X_{jk} = updated X_0 where features of jth ROI are replaced by x_{jk}
      μ_{jk} = value of FPR of x_{jk} in X_{jk}
      if $\mu_{jk} > \tau$ **then**
        $\tau = \mu_{jk}$
        $k_b = k$
      end if
    end for
    // *ROI Local Enhancement*
    Enhancing k_bth window block of jth ROI by PCNN
    Embedding enhanced block into original ROI
  end for
  Extracting feature dataset X_i from all locally enhanced ROIs
  $i++$
end while
// *Feature Fusion*
Initialise: $j = 1$
while $j \leq num$ **do**
  for $i = 0;\ i \leq Max;\ i++$ **do**
    x_{ij} = features of jth sample in X_i
    $\tilde{X}_{ij}$ = updated X_0 where features of jth sample are replaced by x_{ij}
    $\tilde{\mu}_{ij}$ = value of FPR of x_{ij} in X_i
    if $\tilde{\mu}_{ij} > \tau$ **then**
      $\tau = \tilde{\mu}_{ij}$
      $i_b = i$
    end if
  end for
  $j = j++$
  X_{best} = feature dataset where jth sample is equal to $x_{i_b j}$
end while

As can be readily inferred from Alg. 1, the time complexity of the proposed FRIP framework is $O(\max\{Max, num, n\}^3)$.

In the following experiments, the time decay constant α_E in PCNN is initialised as 0.01 and will be updated 100 times with the step size 0.01. In so doing, as the IE process is repeatedly executed, the value of α_E increases from 0.01 to 1. Moreover, as the number of pulse ignition iterations n is set to 10,000, the associated dynamic threshold E_{ij}, which is used to provoke the effect of image enhancement by PCNN, will be updated 10^6 rounds in total.

In addition, the comparative study will apply two variants of the FRIP framework: the globally enhanced ROI (GE) strategy and the randomly locally enhanced ROI (RLE) strategy. Specifically, the GE strategy refers to the means of enhancing the ROI entirely in the IE process within the proposed framework. When FRIP randomly chooses the window block for further enhancement, it will be referred to as the RLE strategy below. Running these two counterparts will help demonstrate the validity and superiority of the FRIP method.

5. Experimental results

In this section, the mammographic risk assessment results based on the features extracted by the proposed FRIP framework are compared against those based on the original ROI strategy, the GE strategy and the RLE strategy. The performance criteria used are: confusion matrix and the statistical tests on both classification accuracy and the kappa coefficient [40]. Moreover, the performance of the FRIP framework with different sizes of the sliding window is also discussed.

The configuration of the implemented FRIP is given in Table 1, where σ_a is the standard deviation of feature a in feature fusion.

The classification accuracies using the features achieved by FRIP, the original ROI, GE and RLE are compared here. For completeness, the classification methods used in this comparative study are J48 [35], JRip [36], PART [37], AdaboostM1 [38], RF [39].

5.1. Performance evaluation

The task of mammographic risk assessment is carried out in terms of confusion matrix and the t-tests on both classification accuracy and the kappa coefficient.

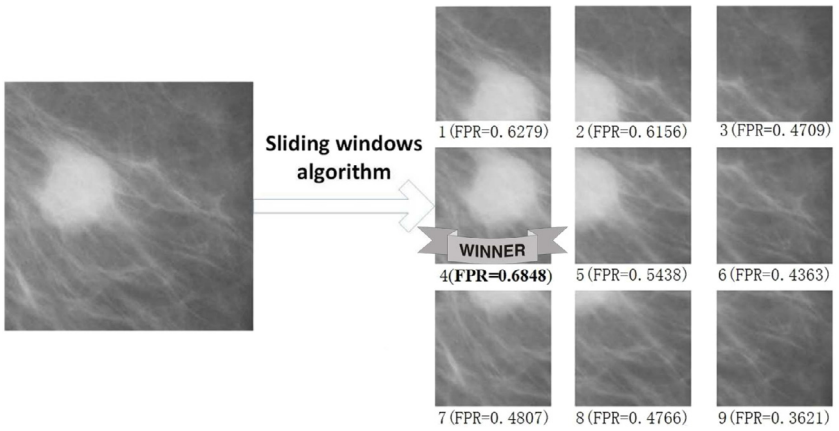


Fig. 8. ROI segmentation and local evaluation.

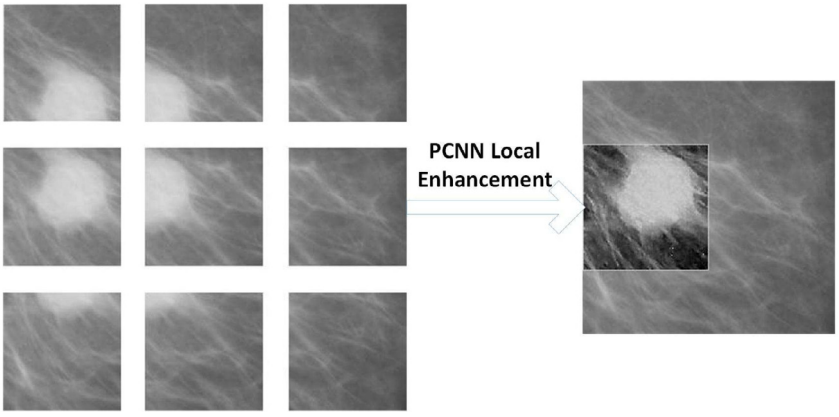


Fig. 9. ROI local enhancement.

Table 1
Configuration of implemented FRIP framework.

Method	Parameters
Sliding window	1. sliding window size is set to 160 × 160, 176 × 176, 192 × 192, 208 × 208, 224 × 224; 2. step size $\lambda = 16$.
PCNN	1. attenuation constant $\alpha_L = 0.06931$; 2. intrinsic potential $V_L = 1.00$; 3. connection strength coefficient $\beta = 0.2$; 4. amplitude constant $V_E = 200$; 5. time decay constant $\alpha_E = 0.01$; 6. number of rounds for updating $\alpha_E = 100$; 7. number of pulse ignition iterations $n = 10000$.
GLCM	1. distance $d = 1$; 2. direction angle $\theta = 0°$, $\theta = 45°$, $\theta = 90°$, $\theta = 135°$.
Feature fusion	1. Gödel T-norm (min(x,y)); 2. Łukasiewicz fuzzy implicator (min(1 − x + y, 1)); 3. fuzzy similarity relation $\mu_{R_a}(x, y) =$ $\max\left(\min\left(\frac{(a(y)-(a(x)-\sigma_a))}{(a(x)-(a(x)-\sigma_a))}, \frac{((a(x)+\sigma_a)-a(y))}{((a(x)+\sigma_a)-a(x))}\right), 0\right)$.

Table 2
Confusion matrices and classification accuracies using the BI-RADS criterion.

Original (Accuracy = 67.19%)					RLE (Accuracy = 90.48%)				
	I	II	III	IV		I	II	III	IV
I	37	18	4	0	I	48	6	0	5
II	13	54	19	0	II	7	76	0	3
III	0	11	120	12	III	0	0	142	1
IV	0	2	20	12	IV	5	2	0	27
GE (Accuracy = 90.09%)					FRIP (Accuracy = 94.04%)				
	I	II	III	IV		I	II	III	IV
I	45	2	11	1	I	52	2	1	4
II	5	81	0	0	II	2	84	0	0
III	4	2	135	2	III	1	0	140	2
IV	0	0	4	30	IV	5	1	0	28

5.1.1. Confusion matrix

Confusion matrix offers a standard means to support evaluation of classification accuracy, expressed in a square matrix form with regard to the number of the class categories. Classification accuracy is herein defined as the ratio of the total number of samples correctly classified to the number of all samples. Within a confusion matrix, each column represents the prediction category, and the total number of each column represents the number of data predicted for that category. Each row represents the true category to which the data belongs, and the total number of data in each row represents the number of data instances in that category.

The confusion matrices as shown in Table 2 in response to the use of the BI-RADS criterion are based on the features achieved by the original, RLE, GE and FRIP, respectively. In particular, the RF method, with forests of 100 trees each, is herein applied to implement the classification tasks. The size of the sliding window is set to 160 × 160. For Tables 3–5, the identical experimental configuration is employed for consistency.

Note that the FRIP framework seeks to reduce class confusion, such as that between class II and class III in Table 2. This is of practical significance because these two classes constitute the

Table 3
Confusion matrices and classification accuracies using the Boyd criterion.

Original (Accuracy = 50.12%)							RLE (Accuracy = 79.69%)						
	I	II	III	IV	V	VI		I	II	III	IV	V	VI
I	0	5	0	1	0	0	I	1	3	0	2	0	0
II	1	37	10	8	2	2	II	1	53	0	4	2	0
III	0	12	19	11	3	1	III	0	0	38	6	2	0
IV	0	5	8	37	21	4	IV	0	0	0	63	10	2
V	0	1	1	16	59	14	V	0	0	1	18	69	3
VI	0	0	0	11	23	10	VI	0	0	0	0	6	38
GE (Accuracy = 74.64%)							FRIP (Accuracy = 84.25%)						
	I	II	III	IV	V	VI		I	II	III	IV	V	VI
I	0	5	1	0	0	0	I	1	2	0	0	3	0
II	3	45	3	3	4	2	II	0	49	3	2	6	0
III	2	3	38	2	1	0	III	0	2	40	1	3	0
IV	0	1	0	62	12	0	IV	0	2	0	62	11	0
V	0	0	0	6	77	8	V	1	2	1	5	82	0
VI	0	1	0	1	18	24	VI	0	0	0	0	8	36

Table 4
Confusion matrices and classification accuracies using the Tabár criterion.

Original (Accuracy = 57.04%)						RLE (Accuracy = 78.16%)					
	I	II	III	IV	V		I	II	III	IV	V
I	91	8	2	15	3	I	110	1	5	1	2
II	9	28	14	1	1	II	4	29	15	2	3
III	16	16	6	2	0	III	4	17	14	3	2
IV	25	0	1	50	6	IV	0	1	2	77	2
V	7	0	0	13	8	V	1	2	2	1	22
GE (Accuracy = 79.88%)						FRIP (Accuracy = 86.21%)					
	I	II	III	IV	V		I	II	III	IV	V
I	109	2	1	7	0	I	111	1	1	4	2
II	1	32	13	7	0	II	10	36	2	5	0
III	5	11	21	3	0	III	0	2	38	0	0
IV	3	4	2	73	0	IV	6	2	0	72	2
V	1	1	0	8	18	V	5	0	1	3	19

Table 5
Confusion matrices and classification accuracies using the Wolfe criterion.

Original (Accuracy = 57.91%)					RLE (Accuracy = 89.64%)				
	I	II	III	IV		I	II	III	IV
I	41	17	4	0	I	48	3	2	9
II	15	51	24	2	II	3	86	1	2
III	1	17	71	19	III	2	1	105	0
IV	0	2	26	32	IV	2	5	0	53
GE (Accuracy = 88.60%)					FRIP (Accuracy = 90.34%)				
	I	II	III	IV		I	II	III	IV
I	49	7	4	2	I	53	2	1	6
II	8	81	3	0	II	2	86	1	3
III	2	4	100	2	III	3	0	98	7
IV	2	0	5	53	IV	2	1	2	55

majority of BI-RADS; it is therefore more useful, though more difficult, to identify class II and III separately. As shown in Table 2, the original dataset performs poorly to distinguish between these two classes. For example, by RF, 19 class II members are incorrectly classified as class III, and 11 class III members are incorrectly classified as class II. Although by RLE and GE, such classification results are improved, the performance of FRIP still outperforms all of them. Few elements in classes III and II are incorrectly classified into one another. Considering the results shown in Tables 3–5 jointly, it can be seen that the FRIP framework improves the ability to distinguish classes significantly. Indeed, these experimental comparisons have shown that the FRIP framework can function better at the level of individual risk types. This is of great practical significance. Moreover, the results also demonstrate that the use of FPR to evaluate the window block is more promising than a random choice.

5.1.2. Statistical tests on classification accuracy

The paired t-test with a significance level of 0.05 is employed to provide a statistical analysis of the resulting classification accuracy rates by the 5 classification methods introduced previously. This is done in order to ensure that results are not discovered by coincidence. Note that In these evaluations, the size of the sliding window is set to be 160 × 160 pixels again.

The results of t-tests are annotated with three tags: better (v), equivalent ( ) or worse (*) in Table 6, signifying that the relative performance achievable by the features using the FRIP framework in comparison to the features of the original dataset, GE and RLE approaches for all of the 4 mammographic risk assessment criteria. Such statistical significance results are further summarised in the rightmost column of this stable, showing the count of the number of statistically better, equivalent and worse results per approach in comparison to the proposed. For example, by using the BI-RADS criterion, the last entry "(0/0/5)" in the RLE row indicates that the set of features generated by RLE performs worse than FRIP for 0 classifiers, equivalently to FRIP for 0 classifiers, and worse than FRIP for 5 classifiers.

From the results shown in Table 6, it can be concluded that except for the occasional inferior results as compared to the use of RLE by AdaboostM1 and JRip for the Wolfe criterion, the proposed FRIP framework statistically outperforms its counterparts in most cases. This from one aspect demonstrates that FRIP improves the quality of the features in an effective and outstanding way for mammographic risk assessment.

5.1.3. Statistical tests on kappa coefficient

To further compare with the existing work, the kappa coefficient [40] is employed to evaluate the experimental results also. The kappa coefficient is generally regarded to be a more robust measure than simple percentage agreement calculations on accuracy since it summarises the level of any agreement between observers after agreements by chance are removed. It tests how well observers agree with themselves (repeatability) and with each other (reproducibility). A high value of the kappa coefficient represents better performance.

Table 7 presents the comparative results on the quality of the features produced by the FRIP framework and by again, the

Table 6
T-test on classification accuracy.

Dataset		AdaboostM1	JRip	PART	J48	RF	Summary
BI-RADS	FRIP	91.83	90.09	91.52	91.64	94.04	(v/ /*)
	RLE	88.37 *	85.72 *	87.69 *	87.63 *	90.48 *	(0/0/5)
	GE	88.63 *	87.02 *	87.48 *	87.82 *	90.09 *	(0/0/5)
	Original	61.12 *	65.90 *	59.38 *	60.86 *	67.19 *	(0/0/5)
Boyd	FRIP	83.63	81.71	82.70	83.48	84.25	(v/ /*)
	RLE	76.27 *	75.13 *	74.96 *	75.27 *	79.69 *	(0/0/5)
	GE	72.16 *	68.09 *	71.45 *	73.31 *	74.64 *	(0/0/5)
	Original	48.13 *	41.30 *	46.87 *	46.20 *	50.12 *	(0/0/5)
Tabár	FRIP	86.80	80.39	85.17	85.57	86.21	(v/ /*)
	RLE	76.24 *	71.71 *	74.54 *	74.79 *	78.16 *	(0/0/5)
	GE	77.39 *	75.29 *	76.40 *	77.39 *	79.88 *	(0/0/5)
	Original	52.91 *	49.66 *	51.39 *	52.91 *	57.04 *	(0/0/5)
Wolfe	FRIP	88.59	86.49	88.91	89.16	90.34	(v/ /*)
	RLE	89.49 v	87.99 v	88.28 *	88.71 *	89.64 *	(2/0/3)
	GE	85.71 *	82.08 *	84.53 *	84.91 *	88.60 *	(0/0/5)
	Original	57.29 *	58.26 *	55.55 *	57.47 *	57.91 *	(0/0/5)

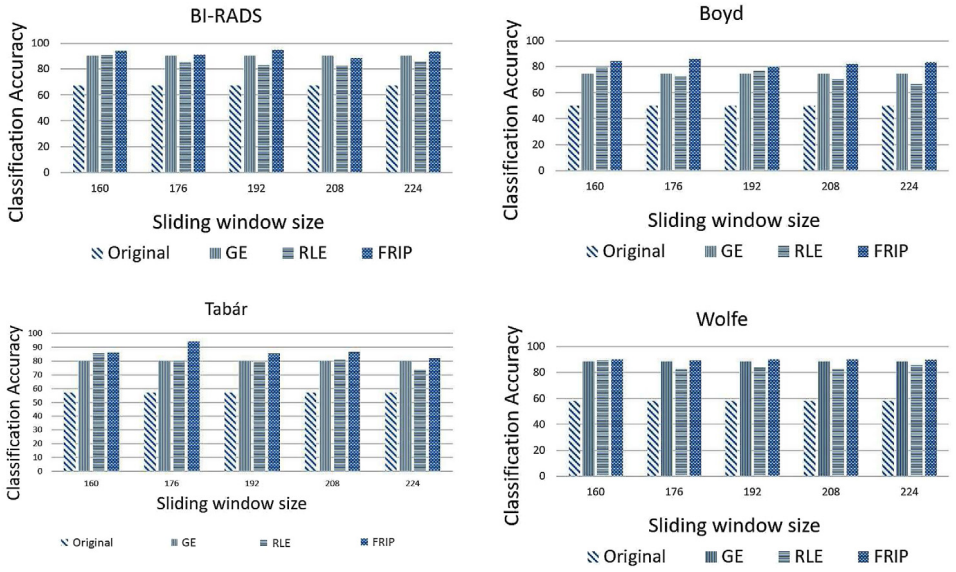


Fig. 10. Classification accuracy using different sliding window sizes for all assessment criteria.

original, RLE and GE strategies while using the same classifiers as illustrated previously. It can be seen that for the BI-RADS, Boyd and Tabár metrics, FRIP leads to a superior performance, although accidentally, for the Wolfe criterion, RLE provides a statistically equivalent performance as compared to FRIP. This observation is probably due to the fact that the block randomly chosen by RLE happens to be closer to the winner block selected by FRIP. Because two adjacent sliding window share a certain amount of identical information, the FPR values of these two blocks may be similar with regard to the Wolfe criterion. Through the use of the multi-round image enhancement strategy, the features extracted from the RLE selected block could be slightly better than those from FRIP selected block by chance.

In particular, for all classification schemes, the values of the kappa coefficient gained by FRIP are consistently, substantially higher than 0.60. This once again demonstrate the superiority of the proposed approach.

5.2. Different sliding window sizes

In this subsection, different sliding window sizes are used to verify further the superiority of the proposed method. The step size is set to 16 pixels. The 5 resulting locally enhanced datasets are implemented with the sliding window sizes which are of 160×160, 176×176, 192×192, 208×208, 224×224 pixels, respectively. For consistency of presenting experimental results, again, the RF classifier with a forest of 100 trees is employed.

As shown in Fig. 10, when compared to the original breast cancer datasets, the feature datasets obtained from the enhanced images systematically receive higher classification accuracies. Occasionally, with a 208×208 window size, the proposed method is slightly underperformed than GE. However, the datasets obtained under the FRIP framework with different sliding window sizes consistently outperform those by the GE and RLE strategies in most cases. In addition, it can be observed that, in some cases, the size of the sliding window will impact the quality of the final feature dataset significantly.

In general, through the experiments conducted, the FRIP framework has improved the classification accuracy by 3%–10% compared to the original, RLE and GE schemes. By checking against the use of different classifiers while comparing the classification accuracy and kappa coefficient in statistics, the superiority of FRIP is clearly demonstrated. Importantly, with the use of differently sized sliding windows, the FRIP datasets achieve consistently better results, showing the robustness of the proposed approach. This experimental result also implies that it is feasible to empirically select important parts to construct FPR. A fundamental discovery is that information obtained from

Table 7
T-test on kappa coefficient.

Dataset		AdaboostM1	JRip	PART	J48	RF	Summary
BI-RADS	FRIP	0.88	0.85	0.88	0.88	0.91	(v/ /*)
	RLE	0.83 *	0.79 *	0.82 *	0.82 *	0.86 *	(0/0/5)
	GE	0.83 *	0.81 *	0.82 *	0.82 *	0.85 *	(0/0/5)
	Original	0.43 *	0.49 *	0.40 *	0.82 *	0.51 *	(0/0/5)
Boyd	FRIP	0.79	0.76	0.78	0.79	0.80	(v/ /*)
	RLE	0.70 *	0.68 *	0.68 *	0.69 *	0.74 *	(0/0/5)
	GE	0.65 *	0.59 *	0.64 *	0.66 *	0.68 *	(0/0/5)
	Original	0.34 *	0.22 *	0.32 *	0.31 *	0.36 *	(0/0/5)
Tabár	FRIP	0.82	0.73	0.80	0.81	0.81	(v/ /*)
	RLE	0.68 *	0.61 *	0.66 *	0.66 *	0.71 *	(0/0/5)
	GE	0.70 *	0.66 *	0.68 *	0.70 *	0.73 *	(0/0/5)
	Original	0.36 *	0.28 *	0.34 *	0.36 *	0.41 *	(0/0/5)
Wolfe	FRIP	0.84	0.82	0.85	0.85	0.87	(v/ /*)
	RLE	0.86 v	0.84 v	0.84 *	0.85	0.86 *	(2/1/2)
	GE	0.80 *	0.75 *	0.79 *	0.79 *	0.84 *	(0/0/5)
	Original	0.41 *	0.42 *	0.39 *	0.42 *	0.42 *	(0/0/5)

the local evaluation can be more representative than the global information.

6. Conclusion

Diagnosis of breast cancer using digital mammograms is an important practical area of research. Positive results may affect human mortality. This paper has presented a fuzzy-rough refined image processing (FRIP) framework to improve the quality of mammographic image features hierarchically. In particular, in this work, ROI is segmented and enhanced locally according to the highest value of the fuzzy positive regions (FPR) of the blocks of potential interest. An object with a higher value of FPR is more affirmatively related to the decisions on risk assessment. In so doing, the features of a mammographic image are extracted and represented according to a locally enhanced ROI. The final feature dataset is generated as a fusion of such resulting features, having the best FPR from a multi-round IE pool.

The proposed framework has proven to give improved mammographic risk assessment results. Nevertheless, as pointed out previously, this is a very difficult application domain. There is no actual ground-truth to ensure which classification result is to be ultimately correct in the first place. Therefore, the experimental results achieved should be used with care, treating them as providing a useful reference aid for human decision making. The eventual task of deciding on the actual mammographic risk is up to human radiologists. Nevertheless, the present approach has shown to offer a good candidate for playing such a supportive role.

The current work can be brought for classifying mammograms into normal and abnormal breast tissue in order to support radiologists for visual diagnosis, and it can be extended to other mammogram databases. Currently, GLCM and PCNN are utilised to implement the FE and IE methods. It would be interesting to investigate any other novel methods that can obtain more representative features (e.g., [55,56] which is also based on fuzzy-rough sets or [57] which covers a range of nature-inspired feature selection approaches) may be employed as their replacement. For the sliding window algorithm, both the step size of each sliding and the sliding window size play an important role in impacting the final results. Consequently, in addition to the verification of different sliding window sizes as shown in this paper, sensitivity analysis with an incremental step size with respect to each sliding is crucial for further model validation. This remains an active research. Moreover, since the important part of a medical image may not be regular, the method focused on selecting irregular important areas is a worthwhile avenue of exploration.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Yanpeng Qu: Conceptualization, Methodology, Writing - original draft, Funding acquisition. **Qilin Fu:** Software, Investigation, Visualization, Formal analysis. **Changjing Shang:** Writing - review & editing, Visualization, Investigation. **Ansheng Deng:** Resources, Data curation. **Reyer Zwiggelaar:** Resources, Validation, Writing - review & editing. **Minu George:** Investigation. **Qiang Shen:** Funding acquisition, Supervision, Project administration.

Acknowledgments

This work is jointly supported by the Innovation Support Plan for Dalian High-level Talents, China (No. 2018RQ70) and a Sêr Cymru II COFUND Fellowship, UK. The authors would like to thank the anonymous referees for their constructive comments which have been very helpful in revising this work.

References

[1] I. Rubio, S. Diaz-Botero, A. Esgueva, R. Rodriguez, T. Cortadellas, O. Cordoba, M. Espinosa-Bravo, The superparamagnetic iron oxide is equivalent to the tc99 radiotracer method for identifying the sentinel lymph node in breast cancer, Eur. J. Surg. Oncol. 41 (1) (2015) 46–51.

[2] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA: Cancer J. Clin. 68 (6) (2018) 394–424.

[3] N. Krishnamoorthy, R. Asokan, I. Jones, Classification of malignant and benign micro calcifications from mammogram using optimized cascading classifier, Curr. Signal Transduct. Therapy 11 (2) (2016) 98–104.

[4] M. Tortajada, A. Oliver, R. Marts, S. Ganau, L. Tortajada, M. Sentss, J. Freixenet, R. Zwiggelaar, Breast peripheral area correction in digital mammograms, Comput. Biol. Med. 50 (4) (2014) 32–40.

[5] Y. Qu, G. Yue, C. Shang, L. Yang, R. Zwiggelaar, Q. Shen, Multi-criterion mammographic risk analysis supported with multi-label fuzzy-rough feature selection, Artif. Intell. Med. 100 (101722) (2019) 1–14.

[6] Y. Qu, C. Shang, Q. Shen, N.M. Parthaláin, W. Wu, Kernel-based fuzzy-rough nearest-neighbour classification for mammographic risk analysis, Int. J. Fuzzy Syst. 17 (3) (2015) 471–483.

[7] P.E. Freer, Mammographic breast density: impact on breast cancer risk and implications for screening, Radiographics 35 (2) (2015) 302–315.

[8] A. Oliver, J. Freixenet, R. Marts, J. Pont, E. PsSrez, E.R. Denton, R. Zwiggelaar, A novel breast tissue density classification methodology, IEEE Trans. Inf. Technol. Biomed. 12 (1) (2008) 55–65.

[9] D.K. Jalal, R. Ganesan, A. Merline, Fuzzy-c-means clustering based segmentation and cnn-classification for accurate segmentation of lung nodules, Asian Pac. J. Cancer Prev. Apjcp 18 (7) (2017) 1869–1874.
[10] M.D. Harrison, J.C. Campos, P. Masci, Reusing models and properties in the analysis of similar interactive devices, Innov. Syst. Softw. Eng. 11 (2) (2015) 95–111.
[11] A. Khokhar, Breast cancer literacy amongst office going women of delhi, Clin. Oncol. 3 (2018) 1430.
[12] C.I. Gonzalez, P. Melin, J.R. Castro, O. Castillo, O. Mendoza, Optimization of interval type-2 fuzzy systems for image edge detection, Appl. Soft Comput. 47 (2016) 631–643.
[13] S. Lin, T. Xu, X. Yin, Region of interest extraction for palmprint and palm vein recognition, in: International Congress on Image and Signal Processing, Biomedical Engineering and Informatics, 2017, pp. 538–542.
[14] A. Elmoufidi, K.E. Fahssi, S. Jai-Andaloussi, A. Sekkaki, Q. Gwenole, M. Lamard, Anomaly classification in digital mammography based on multiple-instance learning, IET Image Process. 12 (3) (2018) 320–328.
[15] T. Prabhakar, S. Poonguzhali, Denoising and automatic detection of breast tumor in ultrasound images, Asian J. Inf. Technol. 15 (18) (2016) 3506–3512.
[16] K. Fujimoto, Y. Ueda, S. Kudomi, T. Yonezawa, Y. Fujimoto, K. Ueda, Automatic roi construction for analyzing time–signal intensity curve in dynamic contrast-enhanced mr imaging of the breast, Radiol. Phys. Technol. 9 (1) (2016) 30–36.
[17] A. Khambampati, D. Liu, S. Konki, K. Kim, An automatic detection of the roi using otsu thresholding in nonlinear difference eit imaging, IEEE Sens. J. 18 (12) (2018) 5133–5142.
[18] Y. Zhang, F. Guo, G. Zhao, Q. Liu, X. Zhang, A comprehensive review of medical image enhancement technologies, Comput. Aided Draft. Des. Manuf. 22 (3) (2012) 1–11.
[19] H. Lidong, Z. Wei, W. Jun, S. Zebin, Combination of contrast limited adaptive histogram equalisation and discrete wavelet transform for image enhancement, IET Image Process. 9 (10) (2015) 908–915.
[20] K. Singh, R. Kapoor, Image enhancement via median-mean based sub-image-clipped histogram equalization, Optik 125 (17) (2014) 4646–4651.
[21] A. Makandar, B. Halalli, Breast cancer image enhancement using median filter and clahe, Int. J. Sci. Eng. Res. 6 (4) (2015) 462–465.
[22] J. Gilles, Empirical wavelet transform, IEEE Trans. Signal Process. 61 (16) (2013) 3999–4010.
[23] K. Lloyd, P.L. Rosin, D. Marshall, S.C. Moore, Detecting violent and abnormal crowd activity using temporal analysis of grey level co-occurrence matrix (glcm)-based texture measures, Mach. Vis. Appl. 28 (3–4) (2017) 361–371.
[24] T. Koyanagi, N.K. Kumar, T. Hwang, L.M. Garrison, X. Hu, L.L. Snead, Y. Katoh, Microstructural evolution of pure tungsten neutron irradiated with a mixed energy spectrum, J. Nucl. Mater. 490 (2017) 66–74.
[25] S. Tian, X. Cui, Y. Gong, An approach to generate spatial voronoi treemaps for points, lines, and polygons, J. Electr. Comput. Eng. 2015 (2015) 41.
[26] L. Zhang, L. Kong, X. Han, J. Lv, Structural class prediction of protein using novel feature extraction method from chaos game representation of predicted secondary structure, J. Theoret. Biol. 400 (2016) 1–10.
[27] L. Sun, Z. Wu, J. Liu, L. Xiao, Z. Wei, Supervised spectral spatial hyperspectral image classification with weighted markov random fields, IEEE Trans. Geosci. Remote Sens. 53 (3) (2015) 1490–1503.
[28] F. Wang, Y. Wu, Q. Zhang, P. Zhang, M. Li, Y. Lu, Unsupervised change detection on sar images using triplet markov field model, IEEE Geosci. Remote Sens. Lett. 10 (4) (2013) 697–701.
[29] K. Layek, S. Samanta, A. Sadhu, S.P. Maity, A. Barui, Classification of sonoelastography images of prostate cancer using transformation-based feature extraction techniques, in: Soft Computing Based Medical Image Analysis, 2018, pp. 245–269.
[30] M. George, E. Denton, R. Zwiggelaar, Mammogram breast density classification using mean-elliptical local binary patterns, in: 14th International Workshop on Breast Imaging, 2018.
[31] R. Membarth, O. Reiche, F. Hannig, J. Teich, M. Korner, W. Eckert, Hipacc: A domain-specific language and compiler for image processing, IEEE Trans. Parallel Distrib. Syst. 27 (1) (2016) 210–224.
[32] R. Jensen, Q. Shen, Computational Intelligence and Feature Selection: Rough and Fuzzy Approaches, John Wiley and Sons, 2008.
[33] R. Jensen, Q. Shen, New approaches to fuzzy-rough feature selection, IEEE Trans. Fuzzy Syst. 17 (4) (2009) 824–838.
[34] R. Zhang, G. Ding, F. Zhang, J. Meng, The application of intelligent algorithm and pulse coupled neural network in medical image process, J. Med. Imaging Health Inform. 7 (4) (2017) 775–779.
[35] C.C. Aggarwal, C.X. Zhai, Mining text data, Data Min. 45 (5) (2015) 429–455.
[36] B. Qin, Y. Xia, S. Prabhakar, Rule induction for uncertain data, Knowl. Inf. Syst. 29 (1) (2011) 103–130.
[37] M. Lydia, A.I. Selvakumar, S.S. Kumar, G.E.P. Kumar, Advanced algorithms for wind turbine power curve modeling, IEEE Trans. Sustain. Energy 4 (3) (2013) 827–835.
[38] K. Kang, J. Michalak, Enhanced version of adaboostm1 with j48 tree learning method, ArXiv Preprint arXiv:1802.03522.
[39] A. Cutler, D. Cutler, J. Stevens, Random forests, Mach. Learn. 45 (1) (2011) 157–176.
[40] M.L. Mchugh, Interrater reliability: the kappa statistic, Biochem. Med. 22 (3) (2012) 276.
[41] W. Choi, S. Riyahi, C.J. Liu, W. Lu, Robust normal lung ct texture features for the prediction of radiation-induced lung disease, Med. Phys. 99 (2) (2017) 196–197.
[42] S. Pare, A.K. Bhandari, A. Kumar, G.K. Singh, An optimal color image multilevel thresholding technique using grey-level co-occurrence matrix, Expert Syst. Appl. 87 (2017) 335–362.
[43] K. Zhan, J. Shi, H. Wang, Y. Xie, Q. Li, Computational mechanisms of pulse-coupled neural networks: a comprehensive review, Arch. Comput. Methods Eng. 24 (3) (2017) 573–588.
[44] C. Huang, G. Tian, Y. Lan, Y. Hao, Y. Cheng, Y. Peng, W. Che, A new pulse coupled neural network (pcnn) for brain medical image fusion empowered by shuffled frog leaping, Front. Neurosci. 13 (2019) 210.
[45] D. Zhou, H. Zhou, C. Gao, Y. Guo, Simplified parameters model of pcnn and its application to image segmentation, Pattern Anal. Appl. 19 (4) (2016) 939–951.
[46] R. Nie, M. He, J. Cao, D. Zhou, Z. Liang, Pulse coupled neural network based mri image enhancement using classical visual receptive field for smarter mobile healthcare, J. Ambient Intell. Humaniz. Comput. (2018) 1–12.
[47] K.P. Pramod, P. Vadakkepat, P.L. Ai, Fuzzy-rough discriminative feature selection and classification algorithm, with application to microarray and image datasets, Appl. Soft Comput. 11 (4) (2011) 3429–3440.
[48] C. Ancy, L.S. Nair, An efficient cad for detection of tumour in mammograms using svm, in: 2017 International Conference on Communication and Signal Processing (ICCSP), IEEE, 2017, pp. 1431–1435.
[49] L. Laura, J.H. Menell, Breast imaging reporting and data system (bi-rads), Radiol. Clin. North Amer. 40 (3) (2002) 409–430.
[50] N. Boyd, J. Byng, R. Jong, E. Fishell, L. Little, A. Miller, G. Lockwood, D. Tritchler, M.J. Yaffe, Quantitative classification of mammographic densities and breast cancer risk: results from the canadian national breast screening study, J. Natl. Cancer Inst. 87 (9) (1995) 670–675.
[51] L. Tabár, T. Tot, P.B. Dean, Breast cancer : the art and science of early detection with mammography : perception, interpretation, histopathologic correlation, Jama J. Amer. Med. Assoc. 300 (300) (2008) 1822.
[52] J.N. Wolfe, Risk for breast cancer development determined by mammographic parenchymal pattern, Cancer 37 (5) (1976) 2486–2492.
[53] P.B. Lam, P.M. Vacek, B.M. Geller, H.B. Muss, The association of increased weight, body mass index, and tissue density with the risk of breast carcinoma in vermont, Cancer 89 (2) (2015) 369–375.
[54] R.B. Vallabhaneni, V. Rajesh, Brain tumour detection using mean shift clustering and glcm features with edge adaptive total variation denoising technique, Alexandria Eng. J. 57 (4) (2018) 2387–2392.
[55] R. Jensen, A. Tuson, Q. Shen, Finding rough and fuzzy-rough set reducts with sat, Inform. Sci. 255 (2014) 100–120.
[56] Y. Qu, R. Li, A. Deng, C. Shang, Q. Shen, Non-unique decision differential entropy-based feature selection, Neurocomputing (2019) http://dx.doi.org/10.1016/j.neucom.2018.10.112, (in press).
[57] R. Diao, Q. Shen, Nature inspired feature selection meta-heuristics, Artif. Intell. Rev. 44 (3) (2015) 311–340.