Article

Breast Density Classification Using Local Quinary Patterns with Various Neighbourhood Topologies †

Andrik Rampun [1],* , **Bryan William Scotney [1],*** , **Philip John Morrow [1],*** , **Hui Wang [2]** and **John Winder [3]**

[1] School of Computing, Ulster University, Coleraine BT52 1SA, UK
[2] School of Computing, Ulster University, Jordanstown, Newtownabbey BT37 0QB, UK; h.wang@ulster.ac.uk
[3] School of Health Sciences, Ulster University, Newtownabbey BT37 0QB, UK; rj.winder@ulster.ac.uk
* Correspondence: y.rampun@ulster.ac.uk (A.R.); bw.scotney@ulster.ac.uk (B.W.S.);
  pj.morrow@ulster.ac.uk (P.J.M.); Tel.: +44-28-7012-4648 (B.W.S.)
† This paper is an extended version of our paper published in Annual Conference on Medical Image Understanding and Analysis, Edinburgh, UK, 11–13 July 2017.

Abstract: This paper presents an extension of work from our previous study by investigating the use of Local Quinary Patterns (LQP) for breast density classification in mammograms on various neighbourhood topologies. The LQP operators are used to capture the texture characteristics of the fibro-glandular disk region (FGD_{roi}) instead of the whole breast area as the majority of current studies have done. We take a multiresolution and multi-orientation approach, investigate the effects of various neighbourhood topologies and select dominant patterns to maximise texture information. Subsequently, the Support Vector Machine classifier is used to perform the classification, and a stratified ten-fold cross-validation scheme is employed to evaluate the performance of the method. The proposed method produced competitive results up to 86.13% and 82.02% accuracy based on 322 and 206 mammograms taken from the Mammographic Image Analysis Society (MIAS) and InBreast datasets, which is comparable with the state-of-the-art in the literature.
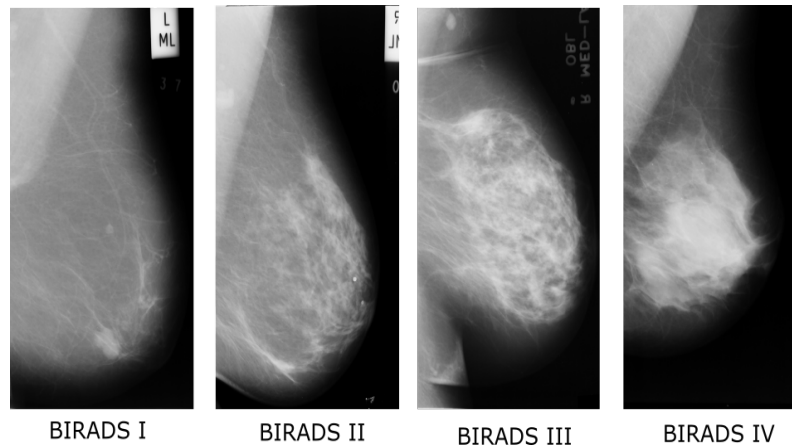
Keywords: breast density classification; computer aided diagnosis; local quinary patterns; breast mammography

1. Introduction

In 2014, there were more than 55,000 malignant breast cancer cases diagnosed in the United Kingdom (UK), with more than 11,000 mortalities [1]. In the United States (US), it was estimated that more than 246,000 malignant breast cancer cases were diagnosed in 2016, with approximately 16% of women expected to die [2]. Many studies have indicated that breast density is a strong risk factor for developing breast cancer [3–12] because breast cancer has a very similar appearance to dense tissues, which makes it difficult to detect in mammograms. Keller et al. [13] investigate the associations between breast cancer and the Laboratory for Individualized Breast Radiodensity Assessment (LIBRA) tool. Their study found that there is a significant association between breast cancer and the Gail risk factor plus Body Mass Index (BMI).

Therefore, an accurate breast density estimation is an important step during the screening procedure because women with dense breasts can be six times more likely to develop breast cancer [1,2]. Although most experienced radiologists can do this task, manual density classification is impractical, tiring, time-consuming and often results vary between radiologists. Computer Aided Diagnosis (CAD) systems can reduce these problems providing robust, reliable, fast and consistent diagnosis results. Currently, in the US, breast density is visually assessed by radiologists who classify density according to the fourth edition of Breast Imaging-Reporting and Data System (BI-RADS) into four classes based

on the following criteria: (a) BI-RADS I (0–25% dense tissues, predominantly fat); (b) BI-RADS II (26–50% dense tissues, fat with some fibro-glandular tissue); (c) BI-RADS III (51–75% dense tissues, heterogeneously dense); and (d) BI-RADS IV (above 75% dense tissues, extremely dense) used in this study. Figure 1 shows four examples of breasts with different BI-RADS classifications.

BIRADS I   BIRADS II   BIRADS III   BIRADS IV

Figure 1. Examples of breast density according to Breast Imaging-Reporting and Data System classes (BI-RADS).

In this paper, we present a fully automatic method for breast density classification using local pattern information extracted using Local Quinary Patterns (LQP) operators from the fibro-glandular region instead of from the whole breast area as all of the studies in the literature have done. In particular, this paper extends our previous work in [14] by investigating the various neighbourhood topologies and different dominant local patterns. The contributions of our study are:

1. To the best of our knowledge, this is the first study attempting to use local patterns extracted using the LQP operators in the application of breast density classification.
2. We investigate the effects of the local pattern information when obtained using different resolutions, various neighbourhood topologies, different orientations and different dominant local patterns.
3. We show the importance of extracting features from the fibro-glandular region instead of from the whole breast area as all of the current studies in the literature have done.

This paper adds the following extensions to our initial work in [14]:

1. We have made a significant extension to the literature review of this study.
2. We have extended the evaluation results for the circle topology, which was originally presented in [14] covering different aspects such as parameter selection, orientations and quantitative comparison.
3. We investigate the topology aspect of the proposed method such as ellipse, parabola and hyperbola.
4. We further evaluated the proposed method on the InBreast [15] dataset.

The remainder of the paper is organised as follows: Section 2 summarises related work in the literature; we present the technical details of the proposed method in Section 3 and present experimental results in Section 4 including the quantitative evaluation and comparisons with existing studies in the literature. Finally, we discuss and present possible future work in Section 5 and conclude the study in Section 6.

2. Literature Review

One of the earliest approaches to breast density assessment was a study by Boyd et al. [16] using interactive thresholding known as Cumulus, where regions with dense tissue were segmented by

manually tuning the grey-level threshold value. The most popular approaches are based on the first and second-order (e.g., Grey Level Co-occurrence Matrix) statistical features as used by Oliver et al. [3], Bovis and Singh [4], Muštra et al. [17] and Parthaláin et al. [6]. Texture descriptors such as local binary patterns (LBP) were employed in the study of Chen et al. [7] and Bosch et al. [8] and Textons were used by Chen et al. [7], Bosch et al. [8] and Petroudi et al. [12]. Other texture descriptors also have been evaluated such as fractal-based [5,10], topography-based [9], morphology-based [3] and transform-based features (e.g., Fourier, Discrete Wavelet and Scale-invariant feature) [4,8].

There are many breast density classification methods in the literature, but only a few studies have achieved accuracies above 80%. The methods of Oliver et al. [3] and Parthaláin et al. [6] extract a set of features from dense and fatty tissue regions segmented using a fuzzy c-means clustering technique for input into the classifier. Oliver et al. [3] achieved 86% accuracy, and Parthaláin et al. [6], who used a sophisticated feature selection framework, achieved 91.4% accuracy based on the Mammographic Image Analysis Society (MIAS) database [18]. Bovis and Singh [4] produced 71.4% accuracy based on a combined classifier paradigm in conjunction with a combination of the Fourier and Discrete Wavelet Transforms, and first and second-order statistical features. Chen et al. [7] made a comparative study on the performance of local binary patterns (LBP), local grey-level appearance (LGA), Textons and basic image features (BIF) and reported accuracies of 59%, 72%, 75% and 70%, respectively. Later, they proposed a method by modelling the distribution of the dense region in topographic representations and reported a slightly higher accuracy of 76%. Petroudi et al. [12] implemented the Textons approach based on the Maximum Response 8 (MR8) filter bank. The χ^2 distribution was used to compare each of the resulting histograms from the training set with all the learned histogram models from the training set and reported 75.5% accuracy. He et al. [11] achieved an accuracy of 70% using the relative proportions of the four Tabár building blocks. Muštra et al. [17] captured the characteristics of the breast region using multi-resolution of first and second-order statistical features and reported 79.3% accuracy.

Tamrakar and Ahuja [19] investigate the study of two-class classification (benign versus malignant) patch-based density problem by considering each of the BI-RADS classes. They investigated various texture descriptors and proposed a new feature extraction method called Histogram of Orientated Texture to achieve more than 92% classification accuracy. Ergin and Kilinc [20] used the Histogram of Oriented Gradients (HOG), Dense Scale Invariant Feature Transform (DSIFT), and Local Configuration Pattern (LCP) methods to extract the rotation- and scale invariant features for all tissue patches to achieve above 90% accuracy. Gedik [21] used fast finite Shearlet transform as a feature extraction technique and used *t*-test statistics as a feature selection procedure. The author reported an average of more than 97% accuracy for two-class classification across two different datasets.

With the advent of deep learning techniques achieving performance similar to human readers [22], they are gaining attention from computer scientists across different research fields. In medical image analysis, deep learning techniques have been used in segmentation and classification problems related to the brain, eye, chest, digital pathology, breast, cardiac, abdomen and musculoskeletal imaging [23]. In the application of breast imaging, particularly in mammography, Kallenberg et al. [24] showed the potential of unsupervised deep learning methods applied to breast density segmentation and mammographic risk scoring on three different clinical datasets, and a strong positive correlation was found compared with manual annotations from expert radiologists. In a study of Ahn et al. [25], a Convolutional Neural Network (CNN) was used to learn the characteristics of dense and fatty tissues from 297 mammograms. Based on the 100 mammograms used as a testing dataset, the authors reported the correlation coefficient between the breast estimates by the CNN and those by the expert's manual measurement to be 0.96. Arevalo et al. [26] proposed a hybrid CNN method to segment breast mass and reported 0.83 in the area under the curve (AUC) value. Qui et al. [27] proposed three pairs of convolution-max-pooling layers that contain 20, 10, and five feature maps to predict short-term breast cancer risk and achieved 71.4%. Using a deeper network approach, Cheng et al. [28] proposed an eight multi-layer deep learning algorithm to employ three pairs of convolution-max-pooling layers

for mammographic masses classification and reported AUC of 0.81. In a similar problem domain (mass classification), Jiao et al. [29] used CNN and a decision mechanism along with intensity information and reported an average of 95% accuracy on different sizes of datasets.

Based on the results reported in the literature, the majority of the proposed methods has used traditional machine learning algorithms tested on 4-class breast density classification and achieved below 80% accuracy. This indicates that breast density classification is a challenging task due to the complexity of tissue appearance in the mammograms such as wide variations in texture and ambiguous texture patterns within the breast region. Although numerical results from deep learning based methods are quite promising, none of these studies has attempted the 4-class density classification problem. In fact, the majority of the deep learning based methods are based on 2-class classification problems (e.g., malignant versus benign, and fatty versus dense tissues). In this paper, texture features were extracted from the fibro glandular region (FGD_{roi}) only instead of from the whole breast region in order to obtain more specific descriptive information. The motivation for this approach is that, in most cases, the non-FGD_{roi} contains mostly fatty tissues regardless of its BI-RADS class, and most dense tissues are located and start to develop within the FGD_{roi}. Therefore, extracting features from the whole breast region means obtaining multiclass texture information, which makes the extracted features less discriminant for the BI-RADS classes.

3. Methodology

Figure 2 shows an overview of the proposed methodology. Firstly, we segment the breast area and estimate the FGD_{roi} area. Subsequently, we use a simple median filter using a 3×3 sliding window for noise reduction and employ the multiresolution LQP operators with different neighbourhood topologies and orientations to capture the microstructure information within the FGD_{roi}. For dimensionality reduction, we select only dominant patterns in the feature space to remove redundant or unnecessary information. Finally, we train the Support Vector Machine (SVM) classifier to build a predictive model and use it to test each unseen case.
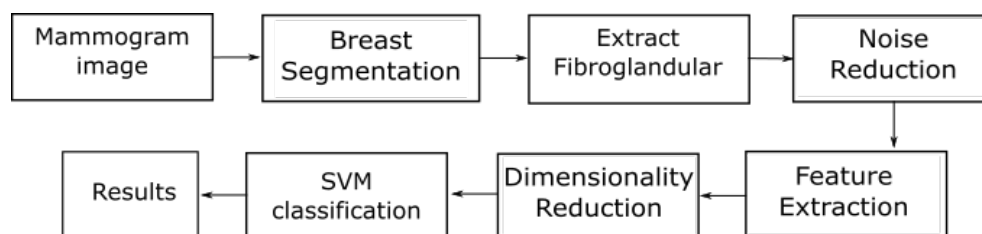
Figure 2. An overview of the proposed breast density methodology.

3.1. Pre-Processing

To segment the breast and pectoral muscle region, we used the method in [30], which is based on Active Contours without edges for the breast boundary estimation and contour growing with edge information for the pectoral muscle boundary estimation. The leftmost image in Figure 3 shows the estimated FGD_{roi} area. To extract FGD_{roi}, we find B_w, which is the longest perpendicular distance between the y-axis and the breast boundary (magenta line in Figure 3). The width and the height of the square area of the FGD_{roi} (amber line Figure 3) can be computed as $B_w \times B_w$ with the centre located at the intersection point between B_h and B_w lines. B_h is the height of the breast, which is the longest perpendicular distance between the x-axis and the breast boundary. B_h is then relocated to the middle of B_w to get the intersection point. The size of the FGD_{roi} varies depending on the width of the breast. Figure 3 shows examples of segmentation results using our method in [30].
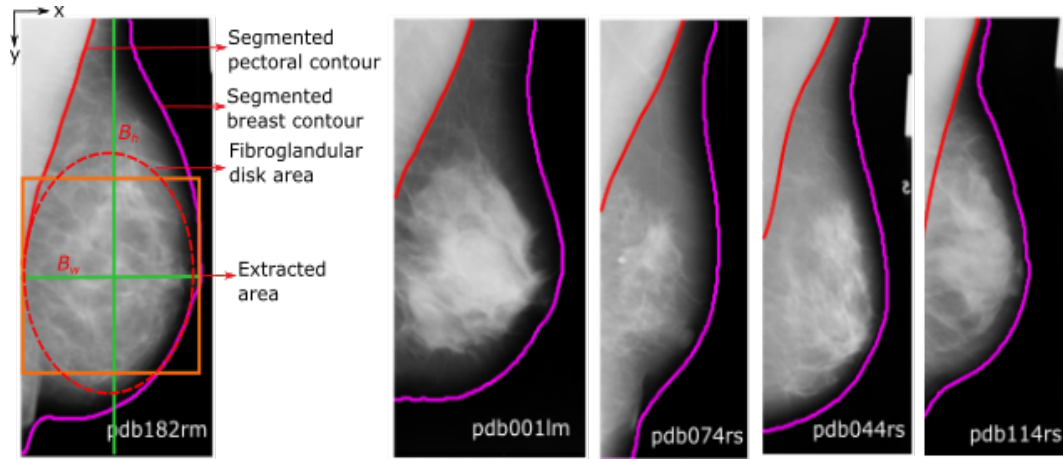
Figure 3. Example of segmentation results using our method in [30].

3.2. Feature Extraction

In the feature extraction stage, we used the LQP mathematical operators to extract local pattern information within the FGD_{roi}, which is a variant of Local Binary Pattern (LBP). The LBP operators were first proposed by Ojala et al. [31,32] to encode pixel-wise information. The value of the LBP code of the pixel (i, j) is given by:

$$LBP_{(P,R)}(i,j) = \sum_{p=0}^{(P-1)} s(g_p - g_c)2^p,$$
(1)

$$s(x) = \begin{cases} 1, & x \geq 0, \\ 0, & \text{otherwise,} \end{cases}$$
(2)

where R is the radius of the circle that form the neighbourhood, P is the number of pixels in the neighbourhood, g_c is the grey level value of the centre pixel, and g_p is the grey level value of the pth neighbour. Later, Tan and Triggs [33] modified the approach by introducing Local Ternary Pattern (LTP) operators, which threshold the neighbouring pixels using a three-value encoding system based on a constant threshold set by the user. In our previous study [34], experimental results suggest that the LTP operators using a combination of uniform and nonuniform mapping patterns ('*riu2*') produced robust texture descriptors that can be used for density classification. Once the LTP code is generated, it is split into two binary patterns by considering its positive and negative components:

$$s(x) = \begin{cases} 1, & x > g_c + \tau_1, \\ 0, & g_c - \tau_1 < x < g_c + \tau_1, \\ -1, & x < g_c + \tau_1, \end{cases}$$
(3)

where τ_1 is a threshold value set by the user. Nanni et al. [35] introduced a five-value encoding system called LQP. The LBP, LTP and LQP are similar in terms of architecture as each is defined using a circle centred on each pixel and the number of neighbours. The main difference is that the LBP, LTP and LQP threshold the neighbouring pixels into two (1 and 0), three (-1, 0 and 1) and five (2, 1, 0, -1 and -2) values, respectively. This means that, for LQP, the difference between the grey-level value of the centre pixel (g_c) and a neighbour's grey level (g_p) can assume five values, whereas LBP and LTP consider two and three values, respectively. The value of LQP code of the pixel (i, j) is given by:

$$LQP_{(P,R)}^{pattern}(i,j) = \sum_{p=0}^{(P-1)} s_{pattern}(g_p)2^p,$$
(4)

$$s(x) = \begin{cases} 2, & x \geq g_c + \tau_2, \\ 1, & g_c + \tau_1 \leq x < g_c + \tau_2, \\ 0, & g_c - \tau_1 \leq x < g_c + \tau_1, \\ -1, & g_c - \tau_2 \leq x < g_c - \tau_1, \\ -2, & \text{otherwise,} \end{cases} \tag{5}$$

where τ_1 and τ_2 are threshold values set by the user, and *pattern* $\in \{1, 2, 3, 4\}$. Once the LQP code is generated, it is split into four binary patterns by considering its positive, zero and negative components, as illustrated in Figure 4, using the following conditions:

$$s_1(x) = \begin{cases} 1, & \text{if } s(x) = 2, \\ 0, & \text{otherwise,} \end{cases} \tag{6}$$

$$s_2(x) = \begin{cases} 1, & \text{if } s(x) = 1, \\ 0, & \text{otherwise,} \end{cases} \tag{7}$$

$$s_3(x) = \begin{cases} 1, & \text{if } s(x) = -1, \\ 0, & \text{otherwise,} \end{cases} \tag{8}$$

$$s_4(x) = \begin{cases} 1, & \text{if } s(x) = -2, \\ 0, & \text{otherwise.} \end{cases} \tag{9}$$

In this case, feature extraction is the process of computing the frequencies of all binary patterns and presenting the occurrences in a histogram, which represents the number of appearances of edges, corners, spots, and lines within the FGD_{roi}. This means that the size of the histogram is 2^P. To enrich texture information, we compute texture information at different resolutions, which can be achieved by concatenating histogram features extracted using different values of R and P in the LQP operators. In this paper, resolution is controlled by the radius of the circle (e.g., different window sizes) and a different number of neighbours. Figure 4 shows an example of converting neighbouring pixels to an LQP code and binary code, resulting in four binary patterns. In the LQP code, there are five numerical values generated based on Equation (5), whereas binary patterns 1, 2, 3 and 4 are generated based on Equations (6)–(9), respectively.
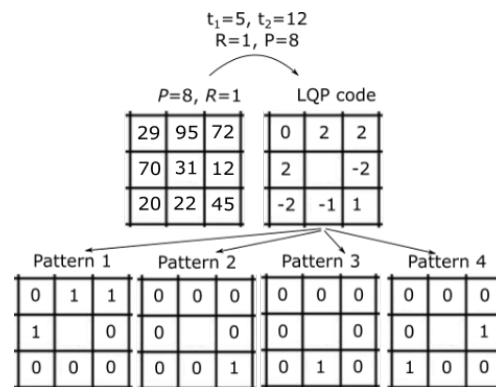

Figure 4. An illustration of computing the Local Quinary Pattern (LQP) code using $P = 8$ and $R = 1$, resulting in four binary patterns.

Figure 5 shows an example of the feature extraction process using multiresolution LQP operators. Note that each resolution produces four binary patterns (see Figure 4) resulting in four histograms. Subsequently, these four histograms are concatenated into a single histogram, which represents the

feature occurrences of a single resolution. The illustration in Figure 5 uses three resolutions, producing three primary histograms (note that each main histogram is a result of concatenation from four histograms of four binary patterns) that are merged into a single histogram as the actual representation of the feature occurrences in the FGD_{roi}.
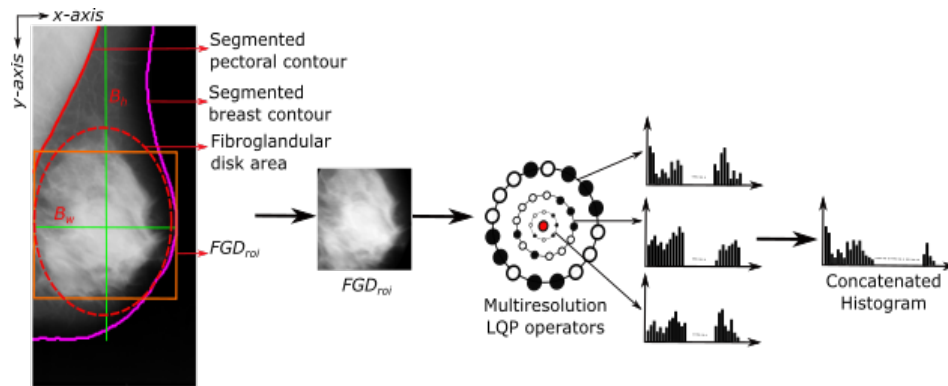
Figure 5. An overview of the feature extraction using multiresolution LQP operators. Black dots in multiresolution LQP operators mean neighbours with less value than the central pixel (red dot). Note that the outer circle with bigger dots represent a larger *R* value.

Figure 6 shows four images of binary patterns generated based on the conditions in Equations (6)–(9). The LQP code image is generated using the term in Equation (5). All histograms from binary patterns are concatenated to produce a single histogram. This process is repeated depending on the number of resolutions chosen by the user. In the proposed approach, this process was repeated three times as we used three different resolutions such as $[P = 5, R = 10]$, $[P = 7, R = 14]$, $[P = 9, R = 18]$.
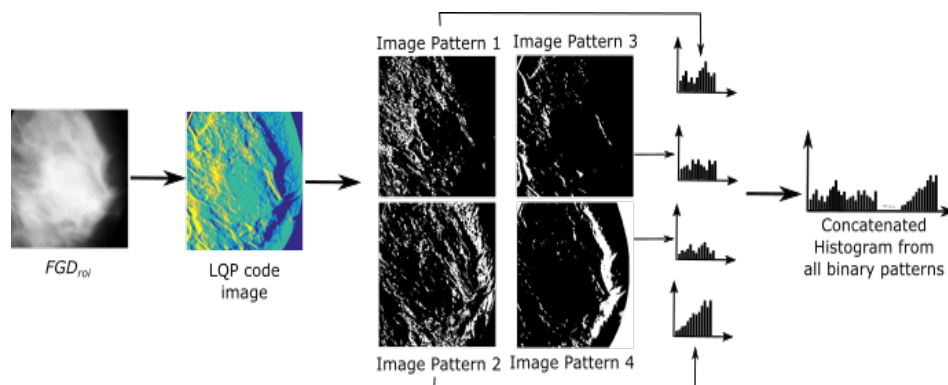
Figure 6. Four images of binary patterns generated from the LQP code image.

3.3. Neighbourhood Topology

The typical neighbourhood topology used in the LBP operators is a circle. In this study, we are interested in investigating the effects of the proposed method when using different neighbourhood topologies. For this purpose, we employed three further topologies, namely ellipse, parabola and hyperbola, as illustrated in Figure 7, where eight neighbours (black dots) were used and distributed equally. The central point is represented by the red dot in each topology. The value for each neighbour is the gray level value of a pixel that is located nearest to it. We summarise the equation and parameters involved in Table 1 for each topology as defined by Nanni et al. [35].
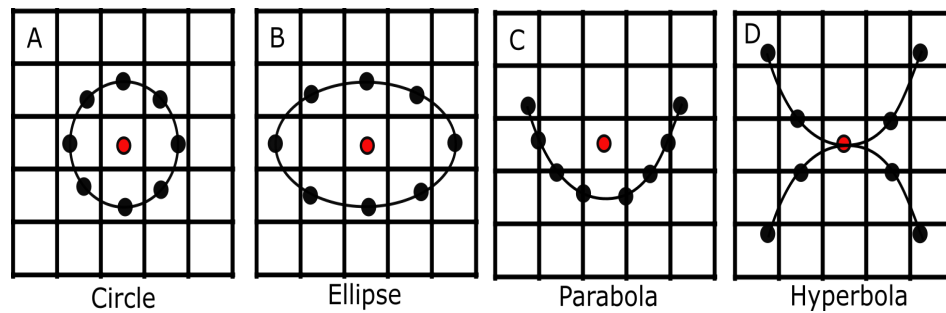
Figure 7. Five different neighbourhood topologies employed in our study.

Table 1. Parameter descriptions for all neighbourhood topologies. Note that *a* and *b* are the semimajor and semiminor axis lengths, and *c* is the distance between vertex and focus.

| Topologies	Equations	Parameters
Circle (*ci*)	$x^2 + y^2 = R^2$	R is the radius of the circle
Ellipse (*el*)	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$	a and b are the semimajor and semiminor axis lengths, where $a \neq b$
Parabola (*par*)	$y = -\frac{1}{c}x^2 + 2c$	c is the distance between vertex and focus
Hyperbola (*hy*)	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$	a and b are the semimajor and semiminor axis lengths

3.4. Topology Orientations

In this study, we further investigate the effects on the classification results when using local pattern information extracted at different orientations (θ). For this purpose, we investigated three other directions: 45°, 90° and 135° (note that $\theta = 0°$ is the default orientation). Furthermore, these orientations also can be combined to create a new topology as illustrated in the rightmost image in Figure 8.
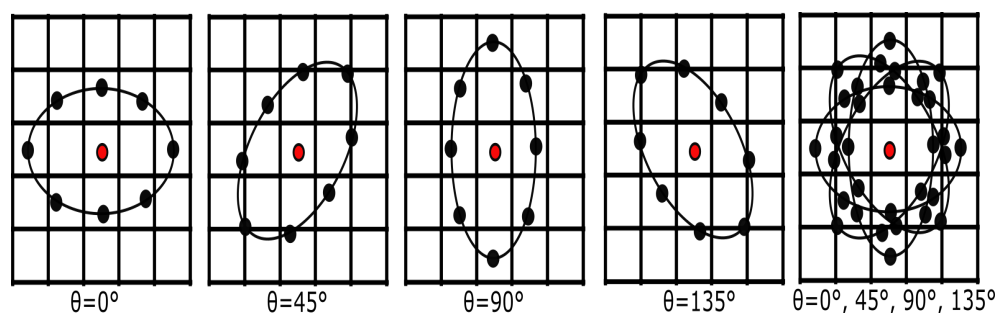
Figure 8. Ellipse topology at different orientations and its combination.

Topology orientation can be achieved through spatial rotation of the binary pattern. To illustrate this process, consider Figure 9, which represents a binary pattern in an ellipse topology (for simplicity, assume that the pattern was generated after the process in Figure 6 was implemented). In this example, the initial decimal value for the binary pattern 01100001 is 134 but rotating the pattern at $\theta = 45°$ produces a new decimal value of 67 (binary pattern becomes 11000010). Similarly, rotation at $\theta = 90°$ and $\theta = 135°$ result in 161 and 208 decimal values, respectively. The same rotation process applies to the circle, parabola and hyperbola topologies.
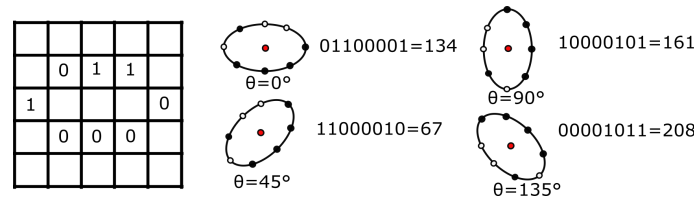
Figure 9. Spatial rotation in an ellipse topology resulting in different decimal values. The decimal value is calculated in a clockwise direction.

3.5. Dominant Patterns

Selecting dominant patterns is an important process for dimensionality reduction purposes due to the large number of features (e.g., concatenation of several histograms). According to Guo et al. [36], a set of dominant patterns for an image is the minimum set of pattern types that can cover $n(0 < n < 100)$ of all patterns of an image. In other words, dominant patterns are patterns that occur frequently in training images. Therefore, to find the dominant patterns, we apply the following procedure. Let $I_1, I_2, ..., I_j$ be images in the training set. Firstly, we compute the multiresolution histogram feature ($H_{I_j}^{LQP}$) for each training image. Secondly, we perform a bin-wise summation for all the histograms to find the pattern's distribution from the training set. Subsequently, the resulting histogram (H^{LQP}) is sorted in descending order, and the patterns corresponding to the first D bins are selected, where D can be calculated using the following equation:

$$D = \arg\min_N \frac{\sum_{i=1}^{N-1} H^{LQP}(i)}{\sum_{i=1}^{2^P} H^{LQP}(i)} > 0.01 \times n. \tag{10}$$

Here, N is the total number of patterns and n is the threshold value set by the user. For example, $n = 97$ means removing patterns that have less than 3% occurrence in H^{LQP}. This means only the most frequently occurring patterns will be retained for training and testing. The smaller the value of n, the smaller the number of patterns selected.

3.6. Classification

Once the feature extraction is completed, the SVM is employed as our classification approach using a polynomial kernel. The GridSearch technique is used to explore the best two parameters (complexity (C) and exponent (e)). We test all possible values of C and e ($C = 1, 2, 3, ..., 10$ and $e = 1, 2, 3, ..., 5$ with interval 1.0) and select the best combination based on the highest accuracy in the training phase. The SVM classifier was trained, and, in the testing phase, each unseen FGD_{roi} from the testing set is classified as BI-RADS I, II, III or IV.

4. Experimental Results

To test the performance of the method, we used the Mammographic Image Analysis Society (MIAS) dataset [18], which consists of 322 mediolateral-oblique (MLO) mammograms of 161 women. The spatial resolution of the images is 50×50 and quantised to eight bits with a linear optical density in the range 0 to 3.2. The density distribution for BI-RADS classes are as follows: 60 (BI-RADS I), 105 (BI-RADS II), 129 (BI-RADS III) and 31 (BI-RADS IV). The second database used in this study is the InBreast dataset [15], which consists of 410 images of 115 women. Since our study concentrates on breast representation in MLO views, only 206 MLO mammograms were taken into our experiment. The pixel size of all images is 70 mm (microns) and 14-bit contrast resolution. The image matrix was 3328×4084 or 2560×3328 pixels. The density distribution for BI-RADS classes are as follows: 69 (BI-RADS I), 74 (BI-RADS II), 49 (BI-RADS III) and 14 (BI-RADS IV). Each image contains BI-RADS information (e.g., BI-RADS class I, II, III or IV) provided by an expert radiologist. A stratified ten runs 10-fold cross validation scheme was employed, where the patients were randomly split into

90% for training and 10% for testing, and repeated 100 times. Accuracy (*Acc*) is used to measure the performance of the method, which represents the total number of correctly classified images as a proportion of the total number of images. Both datasets were annotated by expert radiologists based on the fourth edition of the BI-RADS system. Therefore, evaluation of the proposed method is based on the fourth edition guidelines. All images in both datasets are in *.tif* format (8 bits) with greylevel range from 0 to 255.

4.1. Results on Different Multiresolutions

This section presents quantitative results for the proposed method when using different parameter values (e.g., *P*, *R*, *a*, *b* and *c*) in the LQP operators. We evaluated the method using the following three different multiresolutions and tested it on various neighbourhood topologies:

1. Small multiresolution (LQP^{small}),
2. Medium multiresolution (LQP^{medium}),
3. Large multiresolution (LQP^{large}).

All associated parameters are summarised in Table 2. Note that these parameter values were determined empirically, as the aim of our study is not to optimise the LQP parameters but to investigate the effects of neighbourhood topology in the feature extraction process using the LQP operators. Moreover, since there are many possible combinations of parameter values in each topology, testing all the values is impractical and time-consuming. Nevertheless, we selected several values empirically to show the effects of these parameters in our study.

Table 2. Parameters values for different multiresolutions on different neighbourhood topologies.

Topologies	Multiresolution	Parameters
Circle (R, P)	LQP_{ci}^{small} LQP_{ci}^{medium} LQP_{ci}^{large}	$(1, 8) + (2, 12) + (3, 16)$ $(5, 10) + (7, 14) + (9, 18)$ $(11, 16) + (13, 20) + (15, 24)$
Ellipse (a, b, P)	LQP_{el}^{small} LQP_{el}^{medium} LQP_{el}^{large}	$(1, 3, 10) + (2, 5, 14) + (3, 7, 18)$ $(5, 10, 10) + (7, 14, 14) + (9, 18, 18)$ $(11, 15, 10) + (13, 17, 14) + (15, 19, 18)$
Parabola (c, P)	LQP_{par}^{small} LQP_{par}^{medium} LQP_{par}^{large}	$(1, 10) + (2, 14) + (3, 18)$ $(5, 10) + (7, 14) + (9, 18)$ $(11, 10) + (13, 14) + (15, 18)$
Hyperbola (a, b, P)	LQP_{hy}^{small} LQP_{hy}^{medium} LQP_{hy}^{large}	$(1, 3, 10) + (2, 5, 14) + (3, 7, 18)$ $(5, 10, 10) + (7, 14, 14) + (9, 18, 18)$ $(11, 15, 10) + (13, 17, 14) + (15, 19, 18)$

We present quantitative results for the proposed method in Figure 10 using small, medium and large multiresolutions covering four different neighbourhood topologies. Experimental results suggest that the LQP^{medium} outperformed LQP^{small} and LQP^{large} regardless of its topology and the number of dominant patterns selected, with the best accuracy of 84.91% ($n = 99.2$), followed by the LQP_{el}^{medium} approach at 83.03%($n = 93.8$). The LQP_{hy}^{medium} achieved the third best accuracy and the LQP_{par}^{medium} outperformed both LQP_{par}^{large} and LQP_{par}^{small}. This suggests that using a combination of medium sizes of *R* captures more discriminant features regardless of the topology. Previous studies [37–39] on texture descriptors across different window sizes found that using a small value of *R* does not capture sufficient information about the regions due to the limited intensities and grey level variations. On the other hand, using a large *R* tends to alter the actual representation of the area, especially when one class dominates over another class. For example, if there is a small dense region

within the large circle area that mostly contains fatty tissues, combining different multi-resolutions such as $LQP_{ci}^{small}(1,8) + LQP_{ci}^{medium}(5,10) + LQP_{ci}^{large}(11,16)$ do not provide significant improvement on the classification accuracy ($Acc = 79.65\%$) due to insufficient textural information in $LQP^{small}(1,8)$ and over-representation of the actual textural information from $LQP^{large}(11,16)$. We further tested $LQP_{el}^{small}(1,3,10) + LQP_{el}^{medium}(7,14,14) + LQP_{el}^{large}(15,19,18)$ and achieved an accuracy of 78.62%.
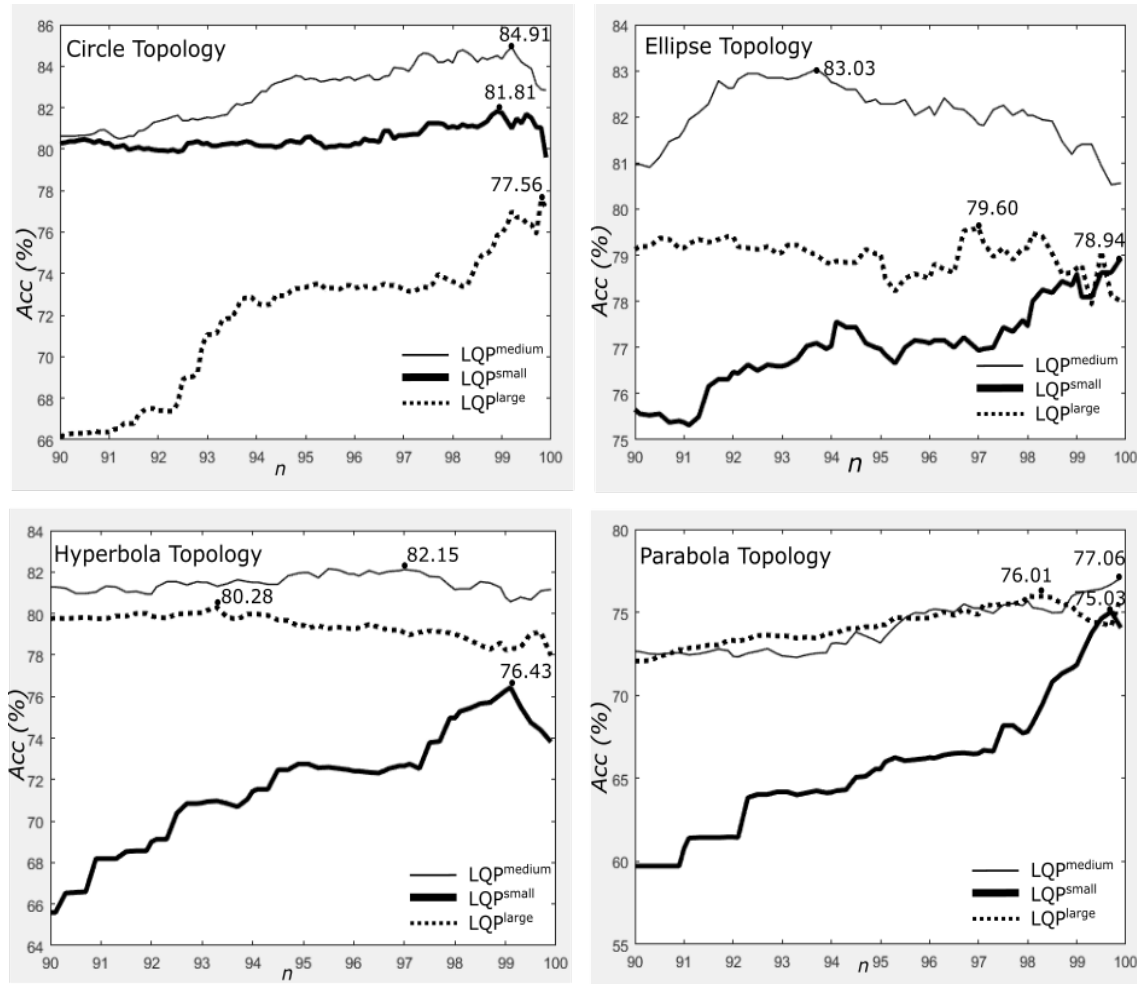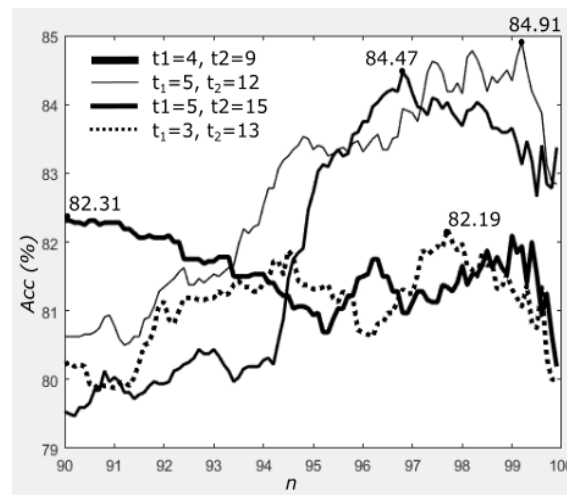

Figure 10. Quantitative results using different multiresolutions based on different neighbourhood topologies. Note that the *x*-axis and *y*-axis represent the accuracy (percentage of correctly classified cases) and the percentage of dominant patterns, respectively.

The results in Figure 10 also indicate that the parameter *n* plays an important role in the performance of the method. It can be observed that the best classification results were achieved when $98.5 < n < 99.5$ for $LQP_{ci}^{small,medium,large}$. For LQP_{el}^{small} and LQP_{hy}^{small}, the best classification results were achieved when $92.5 < n < 94$ and $95 < n < 97.5$, respectively. Moreover, selecting $n < 96$ reduces the classification accuracy significantly, especially for LQP_{ci}^{large}, LQP_{el}^{small} and LQP_{hy}^{large} compared with their best accuracies. Nevertheless, LQP^{medium} produced on average $Acc > 80\%$ across different values of *n* ($90 < n < 99.9$) and different topologies. The LQP_{ci}^{small} produced better results (on average $Acc > 80\%$ across different values of *n*) than LQP_{el}^{small}, LQP_{hy}^{small}, and LQP_{par}^{small}. Regarding accuracy, the LQP_{el}^{medium} produced just under 2% higher compared to the LQP_{ci}^{medium} but approximately 1% higher than the result of LQP_{hy}^{medium}. For the parabola topology, none of the multiresolution approaches achieved $Acc > 80\%$, and the highest accuracy was obtained using LQP_{par}^{small}, with just over 77% correct classification.

4.2. Results on Different Thresholds

For the threshold values, we have investigated several combinations of $\tau_{\in\{1,2\}}$ and found that $\tau_1 = 5$ and $\tau_2 = 12$ produced the best accuracy for all experiments. However, for comparison, we report the performance of the method using $[\tau_1 = 4, \tau_2 = 9]$, $[\tau_1 = 3, \tau_2 = 13]$ and $[\tau_1 = 5, \tau_2 = 15]$ for the circle topology. We also show the effect on performance when varying n from 90 to 99.9 with intervals of 0.1.

Figure 11 shows quantitative results for LQP_{ci}^{medium} using the following thresholds: $([\tau_1, \tau_2])$: $[5, 12]$, $[4, 9]$, $[3, 13]$ and $[5, 15]$. Note that these parameters are determined empirically, and there are many possible values that can be tested. It can be observed that threshold values of $[\tau_1 = 5, \tau_2 = 12]$ produced the best accuracy of 84.91% with $n = 99.2$ followed by $[\tau_1 = 5, \tau_2 = 15]$ with $n = 96.8$. The other threshold values still produced good results compared with most of the proposed methods in the literature. Once again, it can be observed that selecting dominant patterns play a significant role in the classification accuracy. In Figure 11, the proposed method produced very good results when $95.5 < n < 99.5$ for both $[\tau_1 = 5, \tau_2 = 12]$ and $[\tau_1 = 5, \tau_2 = 15]$.
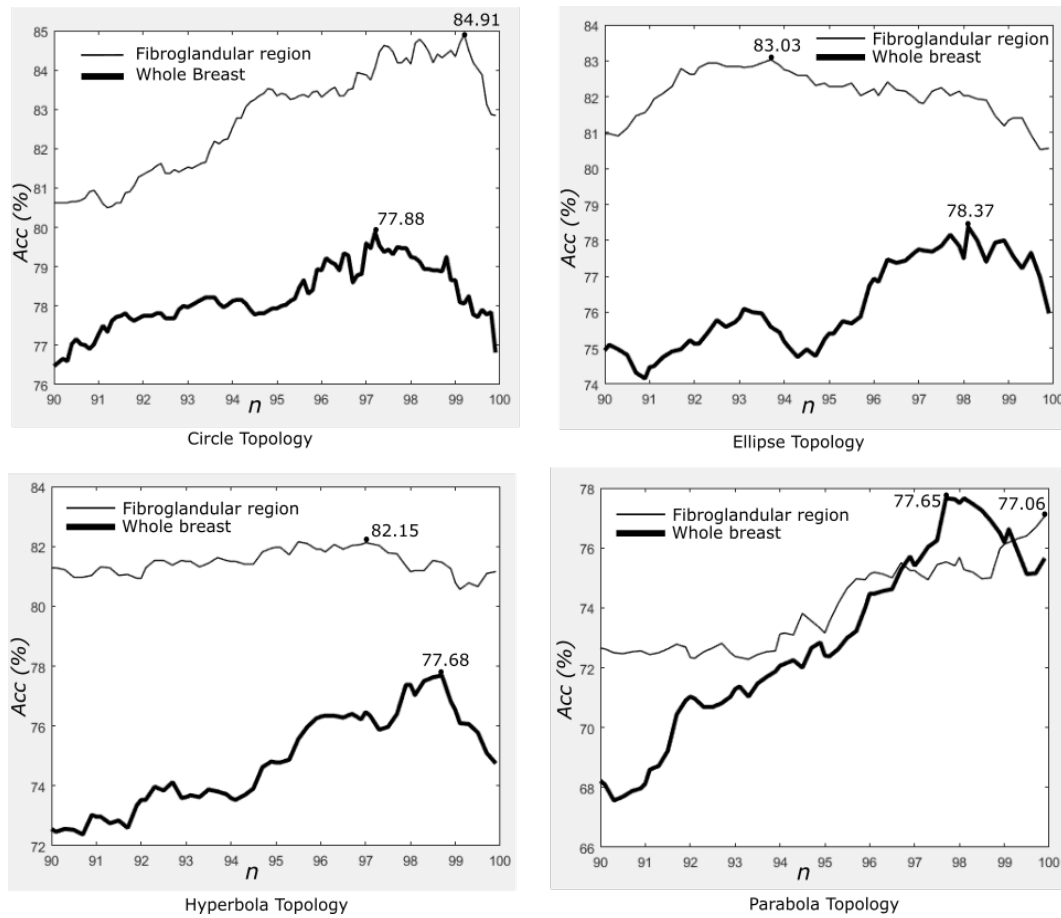

Figure 11. Quantitative results using different values of τ_1 and τ_2 on circle neighbourhood topology. Note that the *x*-axis and *y*-axis represent the accuracy (percentage of correctly classified cases) and the percentage of dominant patterns, respectively.

## 4.3. Fibro-Glandular Region versus Whole Breast Region

In this section, we present quantitative results when classifying unseen cases based on local pattern information extracted from the entire breast area (as all of the current studies in the literature have done) versus pattern information obtained from the FGD_{roi} only.

For this purpose, we conducted an experiment using LQP_{ci}^{medium} with $[\tau_1 = 5, \tau_2 = 12]$. We selected medium multiresolution, as this approach produces the best results compared with small and large multiresolution methods. Results can be seen in Figure 12 (top left) that suggest that textures from the fibro-glandular disk region are sufficient to differentiate breast density. Extracting features from the whole breast produced up to 77.88% accuracy with $n = 97.2$, which is 7% lower than when features are extracted from the FGD_{roi} only. Secondly, to further support these results, we conducted additional experiments using LQP_{el}^{medium}, LQP_{hy}^{medium} and LQP_{par}^{medium}. As can be observed in Figure 12, regardless of topology, features extracted from the FGD_{roi} once again produced more discriminant descriptors than the ones obtained from the whole breast region.

Figure 12. Quantitative results based on features extracted from the whole breast versus fibro-glandular disk region covering four different topologies. Note that the *x*-axis and *y*-axis represent the accuracy (percentage of correctly classified cases) and the percentage of dominant patterns, respectively.

Our explanation for these results is that, in many cases, non-fibro-glandular disk areas predominantly contain fatty tissues regardless of their BI-RADS class because most dense tissues start to develop within the FGD_{roi}. Therefore, capturing texture information outside the FGD_{roi} means extracting similar information resulting in less discriminative features across BI-RADS classes [40]. In cases where the non-fibro-glandular disk region is dominated by dense tissue, the FGD_{roi} mostly contains dense masses. For example, Figure 13 shows histograms extracted from the whole breast regions (*wb*) and the FGD_{roi}. To measure the difference quantitatively, we used the χ^2 distance (*d*) to gauge the difference between these histograms and found $d = 0.0705$ for H_1^{wb} and H_2^{wb}, and $d = 0.122$ for H_1^{roi} and H_2^{roi}. This means that H_1^{wb} is more similar to H_2^{wb} than H_1^{roi} and H_2^{roi}.
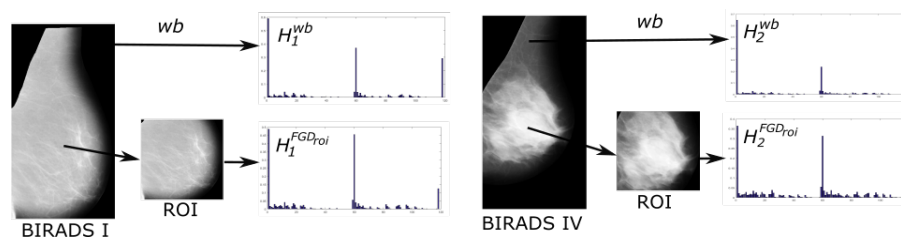
**Figure 13.** Histograms extracted from the *wb* versus region of interest (ROI) with BIRADS class I and IV.

4.4. Results on Different Neighbourhood Topologies

To summarise the performance of the proposed method using different neighbourhood topologies, we present the results evaluated based on the MIAS dataset [18], all in a single graph as shown in Figure 14. The results are based on medium size multiresolution (e.g., LQP_{ci}^{medium}). All associated parameters remain the same as in Table 2. Experimental results in Figure 14 suggest that three of the topologies employed, namely circle, ellipse and hyperbola, produced consistent results ($Acc > 80\%$) across different values of n ($90 < n < 99.9$), whereas LQP_{par}^{medium} failed to achieved accuracy above 80%. The circle topology achieved the highest accuracy of 84.91% followed by ellipse and hyperbola with $Acc = 83.7\%$ and 82.15%, respectively. In terms of accuracy, the main difference between the LQP_{ci}^{medium} and the LQP_{el}^{medium} is that the LQP_{ci}^{medium} tends to produce higher results when $n > 95$ and $n < 95$ for LQP_{el}^{medium}.
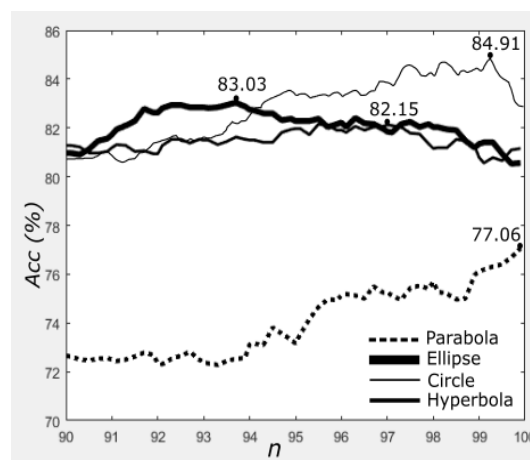
Figure 14. Quantitative results (MIAS dataset [18]) of $LQP_{ci,el,hy,par}^{medium}$ based on features extracted from the fibro-glandular disk region on different neighbourhood topologies. Note that the *x*-axis and *y*-axis represent the accuracy (percentage of correctly classified cases) and the percentage of dominant patterns, respectively.

We further tested the proposed method using the InBreast dataset [15] and present the results in Figure 15. All associated parameters are the same as in Table 2. It can be observed that the ellipse topology performed the best with over 82% accuracy at $n = 99.7$, whereas the circle topology produced close to 80% at $n = 94.4, 94.8$ and 99.5. The parabola topology achieved the highest accuracy (78.5%) at $n = 99.9$ and 78% at $n = 95.1$ for the hyperbola topology. On average, the proposed method achieved 75.1% , 74.6%, 73.9% and 72.5% accuracy for the circle, ellipse, hyperbola and parabola, respectively, across different n.
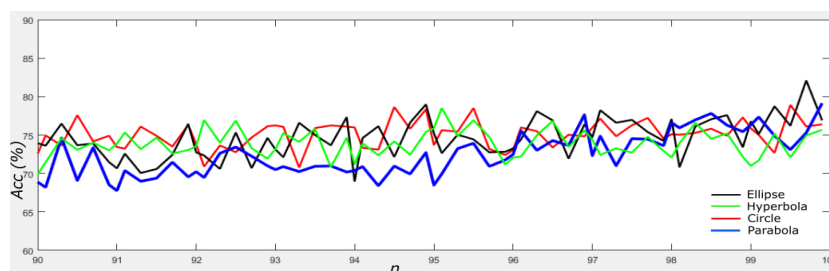
Figure 15. Quantitative results (InBreast dataset [15]) of $LQP_{ci,el,hy,par}^{medium}$ based on features extracted from the fibro-glandular disk region on different neighbourhood topologies. Note that the *x*-axis and *y*-axis represent the accuracy (percentage of correctly classified cases) and the percentage of dominant patterns, respectively.
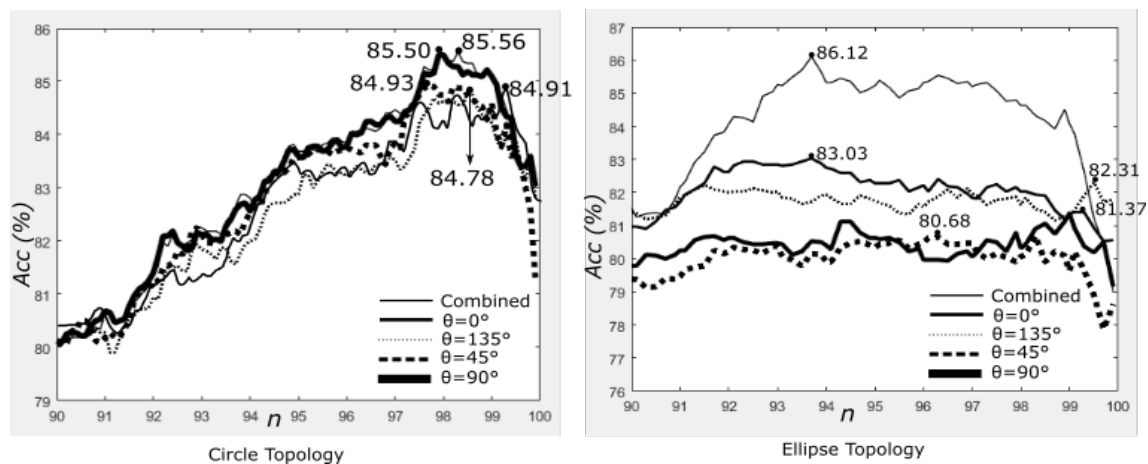
4.5. Results on Different Orientations

To investigate the effect on the performance at various orientations (θ), we conducted several experiments by varying θ with the following values: $0°$ (which is the default orientation), $45°$, $90°$ and $135°$. The parameter values remain the same for all medium size multiresolution approaches with $[\tau_1 = 5, \tau_2 = 12]$. Therefore, the new parameter settings are presented in Table 3. Note that only the circle and ellipse topologies are tested due to their promising results presented in Section 4.4. Combining different orientations for the ellipse results in the neighbourhood architecture becoming more complex as presented in the rightmost image in Figure 8. To assess the effect of a single orientation, first, we set a constant value of θ across different resolutions. For example, for the circle (R, P, θ), $(1, 8, 0°) + (2, 12, 0°) + (3, 16, 0°) + (4, 18, 0°)$ (from now on, we call it $LQP^{medium}_{\{ci,el\}}S(\theta)$, where $\theta \in \{0°, 45°, 90°, 135°\}$). Secondly, we vary the value of θ (as presented in Table 3), hence the parameters are $(1, 8, 0°) + (2, 12, 45°) + (3, 16, 90°) + (4, 18, 135°)$ (from now on, we call it $LQP^{medium}_{\{ci,el\}}S(*)$, where $*$ indicates all four orientations).

**Table 3.** Parameter values for different neighbourhood topologies and orientations.

Topologies	Multiresolution	Parameters
Circle (R, P, θ)	LQP^{medium}_{ci}	$(5, 10, 0°) + (7, 14, 45°) + (9, 16, 90°) + (9, 18, 135°)$
Ellipse (a, b, P, θ)	LQP^{medium}_{el}	$(5, 10, 10, 0°) + (7, 14, 14, 45°) + (9, 18, 18, 90°) + (11, 18, 20, 135°)$

The left image of Figure 16 shows classification accuracies using the circle neighbourhood topology. Individually, experimental results suggest that the $LQP^{medium}_{ci}S(90°)$ produced the best accuracy of 85.50%, which indicates density patterns are more visible at this orientation followed by $\theta = 45°$ with accuracy 84.93%. Overall results suggest that multiresolution $LQP^{medium}_{ci}S(\theta)$ can produce consistent results (>83%) regardless of the orientation with $97 \leq n \leq 99.5$. Combining all orientations suggests that $LQP^{medium}_{ci}S(*)$ produced 85.56%, which is very similar to using a single orientation. On the other hand, the right image in Figure 16 shows that $LQP^{medium}_{el}S(*)$ produced 86.12% accuracy, which is noticeable compared to the second best accuracy of $LQP^{medium}_{el}S(0°)$ (83.03%). The $LQP^{medium}_{el}S(45°)$, $LQP^{medium}_{el}S(90°)$ and $LQP^{medium}_{el}S(135°)$ produced the best accuracy of 80.68%, 81.37% and 82.31%, respectively, when tested using $90 \leq n \leq 100$.

Figure 16. Quantitative results using different orientations evaluated on the MIAS dataset [18]. Note that the *x*-axis and *y*-axis represent the accuracy (percentage of correctly classified cases) and the percentage of dominant patterns, respectively.

Figure 17 shows classification accuracies using the circle and ellipse neighbourhood topologies for the InBreast [15] dataset. Note that we used the same parameter values as presented in Table 3. It can be observed that combining all orientations ($\theta = \{0°, 45°, 90°, 135°\}$) produced the highest accuracy of just over 80% for both topologies. However, each orientation tends to produce an accuracy between 74% to 77%, which is lower than the results when tested on the MIAS dataset [18]. This might be due to a smaller number of samples and imbalance number of instances across different classes, particularly BI-RADS IV (only 6.7%). The accuracy variations shown in Figure 17 indicate that the selection of n is crucial in order to optimise the performance of the proposed method. Individually, the ellipse topology produced the second best accuracy of 80.02% at $n = 99.6$, whereas the circle topology produced an accuracy of 79.99% at $n = 92.8$.
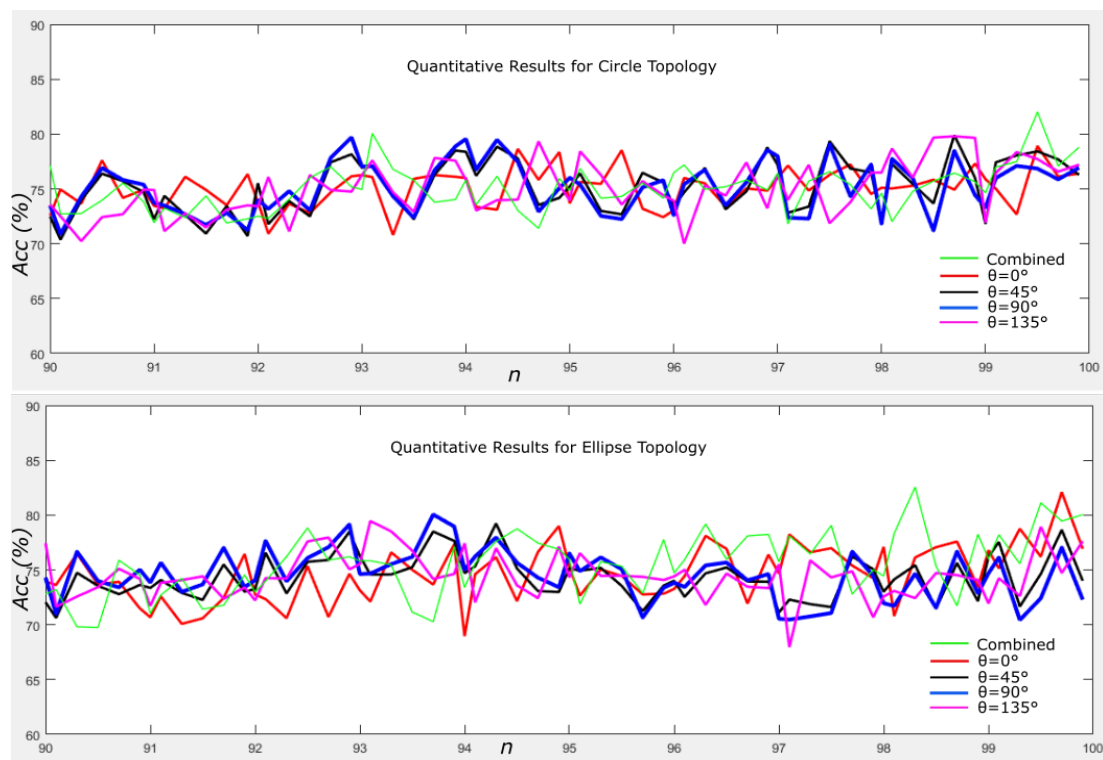
Figure 17. Quantitative results using different orientations evaluated on the InBreast dataset [15]. Note that the *x*-axis and *y*-axis represent the accuracy (percentage of correctly classified cases) and the percentage of dominant patterns, respectively.

4.6. Summary of Quantitative Results

The quantitative results for the sensitivity (*Sen*), specificity (*Spe*) and area under the curve (*AUC*) are summarised in Tables 4 and 5 for both MIAS [18] and InBreast [15] datasets, respectively. We only tested the LQP^{medium} for circle and ellipse topologies because they are proven to be more effective based on the results presented in Sections 4.1–4.5 compared to the other topologies. Since the classification procedure in this study is a non-binary or multiclass [41] problem, we used the pairwise analysis approach [42] to calculate the *Sen*, *Spe* and *AUC* values to further evaluate the proposed method. The *Sen* measures the proportion of actual true positives that are correctly identified and the *Spe* represents the proportion of actual true negatives correctly identified. The *AUC* (or also known as the A_z value) indicates the trade-off between the true positive rate against the false positive rate.

Table 4. Overall quantitative results for MIAS [18] database using circle and ellipse topologies. The parameters *R* and *P* are based on the values in Table 3.

Operators	Sen(%)	Spe(%)	Area Under the Curve(%)	Best *n*
$LQP_{ci}^{medium}(\theta = 0°)$	80.50 ± 17.71	92.30 ± 7.53	88.41 ± 9.12	99.3
$LQP_{ci}^{medium}(\theta = 45°)$	80.01 ± 17.29	92.21 ± 7.78	88.51 ± 9.31	97.6
$LQP_{ci}^{medium}(\theta = 90°)$	79.94 ± 17.81	92.77 ± 7.90	88.86 ± 9.23	97.9
$LQP_{ci}^{medium}(\theta = 135°)$	80.38 ± 18.12	92.13 ± 8.05	87.89 ± 9.92	98.5
$LQP_{ci}^{medium}(All)$	80.67 ± 17.51	92.81 ± 7.91	89.12 ± 9.12	98.3
$LQP_{el}^{medium}(\theta = 0°)$	77.78 ± 20.18	91.03 ± 8.60	87.78 ± 9.62	93.7
$LQP_{el}^{medium}(\theta = 45°)$	76.01 ± 19.97	90.29 ± 9.06	87.03 ± 9.38	96.2
$LQP_{el}^{medium}(\theta = 90°)$	76.38 ± 20.03	90.53 ± 8.77	87.98 ± 9.56	99.2
$LQP_{el}^{medium}(\theta = 135°)$	77.64 ± 20.65	91.20 ± 8.51	86.71 ± 9.92	99.4
$LQP_{el}^{medium}(All)$	78.86 ± 18.27	91.65 ± 7.82	89.92 ± 9.01	93.7

Table 5. Overall quantitative results for InBreast [15] database using circle and ellipse topologies. The parameters *R* and *P* are based on the values in Table 3.

Oprators	Sen(%)	Spe(%)	AUC(%)	Best *n*
$LQP_{ci}^{medium}(\theta = 0°)$	74.67 ± 13.25	83.53 ± 14.76	84.18 ± 9.12	99.5
$LQP_{ci}^{medium}(\theta = 45°)$	71.25 ± 15.29	80.61 ± 16.67	81.71 ± 9.31	98.7
$LQP_{ci}^{medium}(\theta = 90°)$	70.45 ± 14.19	82.91 ± 14.77	82.86 ± 8.93	92.9
$LQP_{ci}^{medium}(\theta = 135°)$	73.24 ± 14.12	83.57 ± 14.35	82.89 ± 8.15	98.7
$LQP_{ci}^{medium}(All)$	76.67 ± 17.65	85.73 ± 17.12	86.89 ± 9.67	94.5
$LQP_{el}^{medium}(\theta = 0°)$	74.85 ± 18.13	85.77 ± 15.09	85.67 ± 9.68	99.7
$LQP_{el}^{medium}(\theta = 45°)$	75.88 ± 18.47	87.06 ± 15.87	84.25 ± 9.38	94.3
$LQP_{el}^{medium}(\theta = 90°)$	73.24 ± 19.25	81.18 ± 16.75	83.56 ± 10.51	93.7
$LQP_{el}^{medium}(\theta = 135°)$	74.14 ± 18.65	85.96 ± 15.85	84.91 ± 9.45	93.1
$LQP_{el}^{medium}(All)$	71.56 ± 18.88	83.65 ± 17.12	84.93 ± 9.61	98.3

Note that the standard deviations in Tables 4 and 5 are quite large due to the effect of imbalance in the number of instances across different classes. For example, the proportion of samples for BI-RADS IV are quite small in both datasets. In a case where the testing set contains only a small number (e.g., 4 samples) of samples from BI-RADS IV and only two of the samples are correctly classified, this results in 50% sensitivity that not only increases the standard deviation but the average *Sen*, *Spe* and *AUC*. Overall, the proposed method achieved the highest $Sen = 80.50\%$ ($LQP_{ci}^{medium}S(0)$) with $Spe = 92.81\%$ ($LQP_{ci}^{medium}S(0)$) and $AUC = 89.92\%$ ($LQP_{el}^{medium}S(All)$) when evaluated on the MIAS dataset [18]. Nevertheless, the results are lower when tested on the InBreast dataset [15] with the best values of 76.67%, 87.06% and 86.89% for *Sen*, *Spe* and *AUC*, respectively.

5. Discussion

We present this section as three subsections that cover quantitative comparison with current methods in the literature, effects of parameter selection on the proposed method and future work.

5.1. Quantitative Comparisons

For quantitative comparison with other methods in the literature, to minimise bias, we selected those studies that have used the MIAS database [18], four-class classification, and used the same evaluation technique (10-fold cross validation) as in this study. Table 6 shows studies that have been conducted using the same evaluation technique on four-class classification. Note that all studies have used all of the 322 mammograms in the MIAS database [18].

Table 6. Quantitative comparison of the current studies in the literature.

Authors	Summary	Accuracy
Parthaláin et al. [6]	First and second-order statistical features and morphological features extracted from fatty and non-fatty tissues. The Fuzzy C-means (FCM) clustering approach was used to segment fatty and non-fatty tissues. Fuzzy-rough approaches were employed for feature selection and classification. The feature extraction pipeline is the same as the study of Oliver et al. [3].	91.4%
Our method	The Local Quinary Pattern (LQP) operators were used to extract local features within the fibro-glandular region using multiresolution and multi-orientation approaches. Only dominant patterns were used for classification using the support vector machine classifier.	86.13% 82.02%
Oliver et al. [3]	First and second-order statistical features and morphological features extracted from fatty and non-fatty tissues. The Fuzzy C-means clustering approach was used to segment fatty and non-fatty tissues. This is the same as the study of Parthaláin et al. [3]. Several machine learning algorithms (e.g., *k*-Nearest Neighbours, decision tree and Bayes classifiers) were combined as a classification approach.	86%
Rampun et al. [40]	The LTP operators were used to extract local features within the fibro-glandular region at different orientations. Feature from all orientations were concatenated to create a long feature vector. Subsequently, the support vector machine (SVM) classifier was employed in the testing phase.	82.33%
Muštra et al. [17]	The first and second order statistical features were extracted at different orientations and resolutions. A combined paradigm of feature selection approach was employed to simplify complexity in the feature space. Finally, the *k*-Nearest Neighbours and Naive Bayesian classifiers were used to classify unseen cases.	79.3%
Chen et al. [7,9]	Several texture descriptors based on the following feature extraction techniques were tested: the local binary pattern, local grey-level appearance, basic image features and Texton approaches. The *k*-NN classification approach was used to differentiate unseen cases into four BI-RADS classes.	59–76%
Petroudi et al. [12]	The Maximum Response 8 (MR8) filter bank approach was used to extract the breast local information. Subsequently, the χ^2 distribution was used to compare each of the resulting histograms from the training set to all the learned histogram models from the training set.	75.5%
Bovis and Singh [4]	Texture features were extracted using Fourier and Discrete Wavelet transforms, and first and second-order statistical features. In the classification phase, a combined classifier paradigm (e.g., the SVM, Random Forest and *k*-NN) was used as a classification approach.	71.4%
He et al. [4]	Class density was determined based on the relative proportions of the four Tabár building blocks within the whole breast region.	70%

Based on the classification results presented in Table 6, Parthaláin et al. [6] achieved the best accuracy of 91.4% using a combination of first and second-order statistical features and morphological features. Our proposed method achieved up to 86.13% accuracy, which is better than the majority of

the methods presented in Table 6. Although Parthaláin et al. [6] achieved the best accuracy, our method is simpler because we use only the LQP operators to extract local pattern information, whereas the method in [6] used several feature extraction techniques. In fact, the Fuzzy C-means clustering is time-consuming, as every pixel needs to be classified as to whether it belongs to the fatty or non-fatty class. Similar to the study of Oliver et al. [3], the FCM clustering technique was employed and the sequential forward selection (SFS) algorithm was used to select a set of discriminant features. Once again, this process is more complicated because each feature will be compared with the other available features when calculating discriminant features. In contrast, our method selected dominant patterns that are based on the number of occurrences. It does not quantitatively compare each feature vector to the other features, which reduces computational time.

Previously, Rampun et al. [40] used the LTP operator to extract local information from the FDG_{roi} and achieved a promising result of 82.33% accuracy. However, this method suffers from having to deal with a large number of features due to the multi-orientation approach (e.g., eight histograms from eight orientations were concatenated). The main difference compared with the proposed method is that the method in [40] concatenated eight orientations and no feature selection was performed, whereas our method performed feature selection by selecting dominant patterns only and investigated different neighbourhood topologies and different resolutions. Muštra et al. [17] developed a method using first and second-order statistical features and employed a complex feature selection framework (e.g., combining several feature selection techniques). Nevertheless, considering the types of features used, they achieved comparable results of almost 80% accuracy. Chen et al. [7,9] investigated several 'bag-of-words' techniques (e.g., LBP, BIF, Textons and LGA) and reported that the Textons (based on MR8) and the LGA approaches achieved the highest accuracies of 76% and 75%, respectively.

Bovis and Singh [4] reported 71.4% accuracy based on a combined classifier paradigm and He et al. [4] produced 70% accuracy using the proportion of four Tabár building blocks. In an unsupervised approach, Petroudi et al. [12] achieved 75.5% classification accuracy using a Textons (MR8) approach to characterise breast appearance. Based on the numerical results presented in the literature, there is still scope for improvement regarding accuracy, which indicates that four-class classification is a challenging task. In our case (based on MIAS database [18]), this might be due to: (a) strong artefacts because images were scanned which also may have altered the actual representation of the breast tissues and (b) an imbalanced number of data across different class. The MIAS database [18] contains 87, 103, 95, 37 cases for BI-RADS I, II, III and IV, respectively. The number of cases for BI-RADS IV is small, which may cause the prediction model to under represent this class.

5.2. Effects of Parameter Selection

There are many parameters involved in this study, and all of them were determined empirically. The aim of our study is not to optimise these parameters in our classification method but to investigate whether the LQP operators can produce discriminant features that can represent breast density based on different neighbourhood topologies. Based on the experimental results shown in Sections 4.1–4.5, all parameters have a significant effect on the performance results. However, the LQP_{ci}^{medium} and LQP_{el}^{medium} produced better results than the other topologies. This may be due to the texture appearance in breasts being mostly anisotropic, and these topologies are better at dealing with these types of structural information (due to their anisotropic shapes) than the parabola and hyperbola. In terms of the selection of R, a large value of R (e.g., 15 or 17) caused one class to be over represented compared with the other classes, whereas a small value of R (e.g., 1 or 2) limits the texture information within the region. On the other hand, a large value of P (e.g., 22) increases the number of patterns but at the same time increases the information complexity in the feature space, and small P (e.g., 4 or 6) results in insufficient texture patterns. In terms of τ_1 and τ_2 values, we found that the results in our study are higher than the ones found in the study of Nanni et al.[35].

Using a multiresolution approach increases the opportunity for capturing textural information that is not represented on a single scale. This is similar when dealing with textures that are more visible in multi-orientations rather than a single orientation. Nevertheless, multiresolution and multi-orientation approaches also have their disadvantages such as increasing the number of features, which also means it increases the possibility of capturing similar information. One way to reduce this problem is to select dominant patterns by taking features that have high occurrences only, but this still does not solve the problem of retaining two sets of similar features with high occurrences. In this study, before selecting dominant features, the number of features is approximately 200. Performing feature selection results in a much smaller number of features (between 50 to 80). For example, selecting $n = 99.2$ results in 70 to 80 dominant features, whereas $n = 95$ retains 55 to 65 dominant features. The main reason for the significant reduction is due to sparse features (large percentage of zeros) that has been reported in [7,9,35]. Since the LQP operators produced sparse features in our case, this has significant variation on the classification accuracies. As can be observed in Figures 10–12, even a small n can change the results significantly because a small n can remove a large number of many features.

When combining all four orientations, we found that the $LQP_{el}^{medium}S(*)$ produced the highest accuracy, over 3% better than using a single orientation. In contrast, the $LQP_{ci}^{medium}S(*)$ produced only 0.06% improvement over $LQP_{el}^{medium}S(90°)$. This is probably because the topology of neighbourhood distribution has changed to more flexible to a 'circle-like' and 'ellipse-like' (see the rightmost image in Figure 8, where the central region is 'circle-like' and the outer shape is 'ellipse-like'). In this case, the topology is able to capture information using both circle and ellipse neighbourhood distribution.

5.3. Future Work

For future work, we plan to develop a method that can automatically estimate τ_1 and τ_2 as well as combine multiresolution LQP features with 'bag-of-words' features such as Textons and MR8. Furthermore, we plan to use a deep learning approach to find discriminant features and combine them with handcrafted features (e.g., local patterns from LQP operators). Although several studies [22–29] have investigated the use of deep learning in breast density classification, to the best of our knowledge, the separation was only based on two- or three-class classification (none of the deep learning based approaches has been applied to four-class classification). In fact, no study has been conducted combining non-handcrafted and handcrafted features in breast density classification.

6. Conclusions

In conclusion, we have developed a breast density classification method using multiresolution LQP operators applied only within the fibro-glandular disk area, which is the most prominent region of the breast instead of the whole breast region as used in other current studies [3–9,11,12]. Experimental results indicate that the multiresolution LQP features are robust compared with the other methods such as LBP, Textons based approaches and LTP due to the five encoding system that generates more texture patterns. Moreover, the multiresolution approach provides complementary information from different parameters that cannot be captured in a single resolution. We further studied the effects on performance when using different neighbourhood topologies (e.g., ellipse, parabola and hyperbola), and found that the LQP_{ci}^{medium} outperformed LQP_{el}^{medium}, LQP_{hy}^{medium} and LQP_{par}^{medium}. Nevertheless, when combining four orientations ($\theta \in \{0°, 45°, 90°, 135°\}$), the proposed method achieved the best accuracy—up to 86.12% (LQP_{el}^{medium}) based on the MIAS dataset [18], which is better than the other LQP approaches employed in this study. Finally, the proposed method produced competitive results based on two datasets (MIAS [18] and InBreast [15]) compared with some of the best accuracies reported in the literature.

Author Contributions: In this study, Andrik Rampun did all the experiments and evaluations discussed. Bryan W. Scotney, Philip J. Morrow, Hui Wang and John Winder supervised the project and contributed equally in preparation of the final version of the paper.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

AUC	Area Under the Curve
BIF	Basic Image Features
BI-RADS	Breast Imaging-Reporting and Data System
CAD	Computer Aided Diagnosis
CNN	Convolutional Neural Network
FCM	Fuzzy C-means
LBP	Local Binary Pattern
LGA	Local Grey-level Appearance
LTP	Local Ternary Pattern
LQP	Local Quinary Pattern
MR8	Maximum Response 8 filter bank
SFS	Sequential Forward Selection
SVM	Support Vector Machine
FGD_{roi}	Fibro Glandular Region

References

1. Cancer Research UK. Breast cancer statistics. 2014. Available online: http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer (accessed on 6 January 2017).
2. Breast Cancer. U.S. Breast Cancer Statistics. 2016. Available online: http://www.breastcancer.org/symptoms/understand_bc/statistics (accessed on 6 January 2017).
3. Oliver, A.; Freixenet, J.; Martí, R.; Pont, J.; Perez, E.; Denton, E.R.E.; Zwiggelaar, R. A Novel Breast Tissue Density Classification Methodology. *IEEE Trans. Inf. Technol. Biomed.* **2008**, *12*, 55–65.
4. Bovis, K.; Singh, S. Classification of Mammographic Breast Density Using a Combined Classifier Paradigm. In Proceedings of the 4th International Workshop on Digital Mammography, Nijmegen, Netherlands, 7–10 June 2002; pp. 177–180.
5. Oliver, A.; Tortajada, M.; Lladó, X.; Freixenet, J.; Ganau, S.; Tortajada, L.; Vilagran, M.; Sentś, M.; Martí, R. Breast Density Analysis Using an Automatic Density Segmentation Algorithm. *J. Digit. Imaging* **2015**, *28*, 604–612.
6. Parthaláin, N.M.; Jensen, R.; Shen, Q.; Zwiggelaar, R. Fuzzy-rough approaches for mammographic risk analysis. *Intell. Data Anal.* **2010**, *14*, 225–244.
7. Chen, Z.; Denton, E.; Zwiggelaar, R. Local feature based mamographic tissue pattern modelling and breast density classification. In Proceedings of the 4th International Conference on Biomedical Engineering and Informatics (BMEI), Shanghai, China, 15–17 October 2011; pp. 351–355.
8. Bosch, A.; Munoz, X.; Oliver, A.; Martí, J. Modeling and Classifying Breast Tissue Density in Mammograms. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06), New York, NY, USA, 17–22 June 2006; pp. 1552–1558.
9. Chen, Z.; Oliver, A.; Denton, E.; Zwiggelaar, R. Automated Mammographic Risk Classification Based on Breast Density Estimation. In *Pattern Recognition and Image Analysis; Volume 7887 of the series Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 237–244.
10. Wolfe, J.N. Risk for breast cancer development determined by mammographic parenchymal pattern. *Cancer* **1976**, *37*, 2486–2492.
11. He, W.; Denton, E.; Stafford, K.; Zwiggelaar, R. Mammographic Image Segmentation and Risk Classification Based on Mammographic Parenchymal Patterns and Geometric Moments. *Biomed. Signal Process. Control* **2011**, *6*, 321–329.

12. Petroudi, S.; Kadir, T.; Brady, M. Automatic Classification of Mammographic Parenchymal Patterns: A Statistical Approach. In Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat. No.03CH37439), Cancun, Mexico, 17–21 September 2003; Volume 1, pp. 798–801.

13. Keller, B.M.; Chen, J.; Daye, D.; Conant, E.F.; Kontos, D. Preliminary evaluation of the publicly available Laboratory for Breast Radiodensity Assessment (LIBRA) software tool: Comparison of fully automated area and volumetric density measures in a case-control study with digital mammography. *Breast Cancer Res.* **2015**, *17*, 1–17.

14. Rampun, A.; Morrow, P.J.; Scotney, B.W.; Winder, R.J. Breast density classification in mammograms using local quinary patterns. In Proceedings of the Annual Conference on Medical Image Understanding and Analysis MIUA 2017: Medical Image Understanding and Analysis, Edinburgh, UK, 11–13 July 2017; pp. 365–376.

15. Moreira, I.C.; Amaral, I.; Domingues, I.; Cardoso, A.; Cardoso, M.J.; Cardoso, J.S. INbreast: Toward a full-field digital mammographic database. *Acad. Radiol.* **2011**, *19*, 236–428.

16. Boyd, J.W.; Boyd, N.F.; Fishell, E.; Jong, R.A.; Yaffe, M.J. Automated analysis of mammographic densities. *Phys. Med. Biol.* **1996**, *41*, 909–923.

17. Muštra, M.; Grgić, M.; Delać, K. A Novel Breast Tissue Density Classification Methodology. Breast Density Classification Using Multiple Feature Selection. *Automatika* **2012**, *53*, 362–372.

18. Suckling, J.; Parker, J.; Dance, D.; Astley, S.; Hutt, I.; Boggis, C.; Ricketts, I. The mammographic image analysis society digital mammogram database. *Proc. Excerpta Med. Int. Congr. Ser.* **1994**, 375–378.

19. Tamrakar, D.; Ahuja, K. Density-Wise Two Stage Mammogram Classification Using Texture Exploiting Descriptors. *arXiv* **2017**, arXiv:1701.04010.

20. Ergin, S.; Kilinc, O. A new feature extraction framework based on wavelets for breast cancer diagnosis. *Comput. Biol. Med.* **2015**, *51*, 171–182.

21. Gedik, N. A new feature extraction method based on multiresolution representations of mammograms. *Appl. Soft Comput.* **2016**, *44*, 128–133.

22. Kooi, T.; Litjens, G.; van Ginnekena, B.; Gubern-Méridaa, A.; Sánchez, C.I.; Manna, R.; den Heeten, A.; Karssemeijer, N. Large scale deep learning for computer aided detection of mammographic lesions. *Med. Image Anal.* **2017**, *35*, 303–312.

23. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.W.M.; van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88.

24. Kallenberg, M.; Petersen, K.; Nielsen, M.; Ng, A.; Diao, P.; Igel, C.; Vachon, C.; Holland, K.; Karssemeijer, N.; Lillholm, M. Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring. *IEEE Trans. Med. Imaging* **2016**, *35*, 1322–1331.

25. Ahn, C.K.; Heo, C.; Jin, H.; Kim, J.H. A Novel Deep Learning-based Approach to High Accuracy Breast Density Estimation in Digital Mammography. In Proceedings of the SPIE Medical Imaging 2017: Computer-Aided Diagnosis, Orlando, FL, USA, 3 March 2017; Volume 10134.

26. Arevalo, J.; González, F.A.; Ramos-Pollán, R.; Oliveira, J.L.; Lopez, M.A.G. Representation learning for mammography mass lesion classification with convolutional neural networks. *Comput. Methods Program. Biomed.* **2016**, *127*, 248–257.

27. Qiu, Y.; Wang, Y.; Yan, S.; Tan, M.; Cheng, S.; Liu, H.; Zheng, B. An initial investigation on developing a new method to predict short-term breast cancer risk based on deep learning technology. In Proceedings of the SPIE Medical Imaging 2016: Computer-Aided Diagnosis, San Diego, CA, USA, 24 March 2016; Volume 9785.

28. Cheng, J.Z.; Ni, D.; Chou, Y.H.; Qin, J.; Tiu, C.M.; Chang, Y.C.; Huang, C.S.; Shen, D.; Chen, C.M. Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans. *Sci. Rep.* **2016**, *15*, 24454.

29. Jiao, Z.; Gao, X.; Wang, Y.; Li, J. A deep feature based framework for breast masses classification. *Neurocomputing* **2016**, *197*, 221–231.

30. Rampun, A.; Morrow, P.J.; Scotney, B.W.; Winder, R.J. Fully Automated Breast Boundary and Pectoral Muscle Segmentation in Mammograms. *Artif. Intell. Med.* **2017**, *79*, 28–41.

31. Ojala, T.; Pietikainen, M.; Maenpaa, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 971–987.

32.   Hadid, A.; Pietikainen, M.K.; Zhao, G.; Ahonen, T. *Computer Vision Using Local Binary Patterns*; Springer: London, UK, 2011; pp. 13–47.

33.   Tan, X.; Triggs, B. Enhanced local texture feature sets for face recognition under difficult lighting conditions. In *Analysis and Modelling of Faces and Gestures*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 168–182.

34.   Rampun, A.; Morrow, P.J.; Scotney, B.W.; Winder, J. A Quantitative Study of Local Ternary Patterns for Risk Assessment in Mammography. In Proceedings of the International Conference on Innovation in Medicine and Healthcare, Vilamoura, Portugal, 21–23 June 2017; pp. 283–286.

35.   Nanni, L.; Luminia, A.; Brahnam, S. Local binary patterns variants as texture descriptors for medical image analysis. *Artif. Intell. Med.* **2010**, *49*, 117–125.

36.   Gio, Y.; Zhao, G.; Pietikäinen, M. Discriminative features for feature description. *Pattern Recognit.* **2012**, *45*, 3834–3843.

37.   Rampun, A.; Wang, L.; Malcolm, P.; Zwiggelaar, R. A quantitative study of texture features across different window sizes in prostate t2-weighted mri. *Procedia Comput. Sci.* **2016**, *90*, 74–79.

38.   Rampun, A.; Tiddeman, B.; Zwiggelaar, R.; Malcolm, P. Computer aided diagnosis of prostate cancer: A texton based approach. *Med. Phys.* **2016**, *43*, 5412–5425.

39.   Rampun, A.; Zheng, L.; Malcolm, P.; Tiddeman, B.; Zwiggelaar, R. Computer-aided detection of prostate cancer in T2-weighted MRI within the peripheral zone. *Phys. Med. Biol.* **2016**, *61*, 4796–4825.

40.   Rampun, A.; Morrow, P.J.; Scotney, B.W.; Winder, R.J. Breast density classification in mammograms using local ternary patterns. In Proceedings of the International Conference Image Analysis and Recognition ICIAR 2017: Image Analysis and Recognition, Montreal, QC, Canada, 5–7 July 2017; pp. 463–470.

41.   Aly, M. Survey on Multiclass Classification Methods. 2005. Available online: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.175.107 (accessed on 4 December 2017).

42.   Landgrebe, T.C.W.; Duin, R.P.W. Approximating the multiclass ROC by pairwise analysis. *Pattern Recognit. Lett.* **2007**, *28*, 1747–1758.