

Breast Cancer Classification using Local Directional Ternary Patterns

Mary Mwadulo
Computer Science Department
Masinde Muliro University of
Science and Technology
Kakamega, Kenya

Stephen Mutua, PhD
Computer Science Department
Meru University of Science and
Technology
Meru, Kenya

Raphael Angulu, PhD
Computer Science Department
Masinde Muliro University of
Science and Technology
Kakamega, Kenya

ABSTRACT

Local texture descriptors have outperformed holistic texture descriptors in various pattern recognition applications. However, the local descriptors have limitations that can compromise the data in an image. For instance, the Local Binary Patterns (LBP) are sensitive to noise, Local Ternary Patterns (LTP) use a static threshold for all images making it a challenge to select an optimum threshold for all images in a dataset, and the Local Directional Patterns (LDP) use orientation responses to derive an image gradient disregarding the central pixel and $8-k$ responses. These limitations lead to the loss of subtle texture features while encoding an image. This paper proposes a Local Directional Ternary Patterns (LDTP) texture descriptor, which not only considers the central pixel in encoding image gradient but also takes into account all directional responses and an adaptive threshold. Findings from empirical records for the MIAS breast cancer dataset and using different classifiers show that LDTP attains a higher accuracy level for both normal/abnormal and benign/malignant classification compared to the other local texture descriptors.

Keywords

Breast cancer, Feature extraction, LBP, LDP, LTP, LDTP, Mammogram.

1. INTRODUCTION

Computer vision and pattern recognition applications widely use texture analysis because of its potential in extracting significant image features. Application of texture descriptors in facial image analysis[1], object detection [2], pedestrian detection [3], and Medical image analysis [4] have attracted a lot of research interest because they give deep insights to texture recognition processes. The two broad categories of texture descriptors are; holistic and local. Holistic descriptors describe an image as a whole to generalize the entire object, while local descriptors describe critical points in an image. Further, holistic descriptors base their dimension reduction on techniques such as Principal Component Analysis (PCA)[5] and Linear Discriminant Analysis (LDA)[6]. These descriptors have been proved successful in pattern recognition applications; however, they often fail in small sample size problems[7].

In recent years, much research has focused on the application of local descriptors for texture analysis because of their effectiveness, ease of extraction, robustness to illumination, and pose variation[8]. Local Binary Patterns (LBP) has emerged as one of the most prominent local texture descriptors because of its discriminative power and computational simplicity. Consequently, it has seen successful application in facial expression recognition [9], medical image analysis[10],

and biomedical image analysis [11]. However, a significant limitation of LBP is that it misses the local structure of an image since it does not consider the effect of the central pixel. Furthermore, it is sensitive to noise, and sometimes it may falsely classify two or more different patterns to the same class. Consequently, owing to its low dimensionality and high efficiency[12], LBP resulted in a wide variety of descriptors aimed at improving its discrimination capability, robustness, and applicability. Some of its variants include; Local Ternary Patterns (LTP) [13] and Local Directional Patterns (LDP)[14].

The Local Ternary Patterns introduced by Tan *et al* [13] aimed to make LBP less sensitive to the noise by extending it to a three value code instead of the two value code. Although LTP proved to be more robust to noise than LBP, it does not capture detailed information of the image and sometimes it may falsely classify different patterns to the same class as experienced by LBP. Additionally, it uses a static threshold defined by the user, which makes it subjective and dynamically inappropriate for all images in a given dataset. In a bid to overcome these challenges, Jabid *et al* [14] proposed the Local Directional Patterns (LDP) descriptor that is more consistent in the presences of noise and works by considering edge response value in different directions using compass masks such as Kirsch [15], Robinson[16], Frei-Chen [17]and Sobel[18].

The Kirsch is the most popularly used mask and works by deriving edge responses in eight directions. It uses the top k significant edge responses to derive the LDP code of the referenced pixel while ignoring the remaining $8-k$ responses. Although LDP generates different codes for different patterns, the reliance of top k significant responses leads to loss of information around a local neighborhood. Furthermore, failure to involve the central pixel leads to loss of discriminant information, which may make LDP capture an image edge where there is none. Moreover, Shabat *et al* [19] noted that LDP is computationally expensive compared to other local descriptors.

The motivation of this work thus is two-fold. First, it seeks to improve on the limitations of known local descriptors to enhance the encoding of more salient texture features. As a result, this paper proposes a technique herein referred to as the Local Directional Ternary Patterns (LDTP) texture descriptor that seeks to improve on the limitations of LDP and LTP. This descriptor considers all the directional responses in encoding local texture in an image which enables more subtle texture features to be captured and improve the discriminative power of the extracted features. Additionally, the proposed descriptor uses an adaptive threshold, making it dynamically appropriate for all images in a dataset or diverse datasets.

Second, the work sought to validate the proposed descriptor

using breast cancer data obtained from the MIAS dataset and compare the findings against those obtained when using the LTP and LDP descriptors. The 322 images in the dataset underwent data augmentation in which the original image was mirrored along the x-axis and the y-axis resulting in 966 images. Using the Support Vector Machines (SVM) and k-Nearest Neighbors (KNN) classifiers, empirical findings show that the LDTP method has a better accuracy level than LTP and LDP.

The remainder of the paper is organized as follows; Section II discusses related work, while Section III analyses existing local texture descriptors related to the proposed descriptor. Section IV presents the proposed LDTP descriptor, and Section V presents materials and experiments used in validating the LDTP approach for breast cancer classification. Section VI discusses the results, and finally, section VII presents the conclusion and gives further research insights.

2. RELATED WORK

Previous studies have used local texture descriptors for extracting breast cancer features. Muramatsu *et al* [20] developed the Radial Local Ternary Patterns (RLTP) that considered not only pattern orientation about the center of a mass but also robustness to image rotation. Using 376 ROI's from Nagoya medical dataset and ANN, SVM and RF classifiers, for benign /malignant classification, they achieved the highest Area Under Curve (AUC) of 0.90 with ANN classifier. Even though the test was performed using Leave-One-Case-Out cross-validation, the training and test dataset were not completely independent. Therefore to test the effectiveness of the proposed method there is a need to validate it with an independent dataset.

Rabidas *et al* [21] compared the performance of LBP, LBP Variance, and Complete LBP extraction techniques for benign/malignant breast mass classification. By using a stepwise logistic regression method for feature selection and the Fisher Linear Discriminant Analysis classifier on 200 mammograms from the DDSM database, they achieved a classification accuracy of 92.95% 87.7% and 90.6% with LBP, LBPV, and CLBP respectively using tenfold cross-validation. Even though LBP achieved the highest accuracy, it was not rotation invariant and therefore not suitable for benign/malignant classification which requires edge orientation at the margin.

In an attempt to achieve a local pattern that is rotation invariant, Rabidas *et al* [22] proposed Discriminative Robust LBP (DRLBP) and Discriminative Robust LTP (DRLTP) texture descriptors for classification of mammographic masses as benign or malignant. Using Fisher Linear Discriminant Analysis classifier, stepwise logistic regression for feature selection, tenfold cross-validation on 58 masses from the Mini MIAS dataset. They achieved the best results with DRLBP having an AUC of 0.98. Even though the results of DRLBP and DRLTP were very close to each other, the authors did not test the statistical significance of that difference.

Ponraj *et al* [23] proposed a Local Binary Textual Patterns that considers the central pixel. The central pixel is compared with the neighboring pixel to generate a binary bit. The binary bit converts to a decimal number by counting the number of zeros and ones. Then textual features are used to classify a mammogram into either a normal or an abnormal class. By using an SVM classifier and 70 images from the MIAS database, they achieved an accuracy of 97.2% and 96.4% for Binary pattern one and Binary pattern two, respectively. The main criticism of the approach is that it leads to long

histograms that increase the computational speed. Further, there was no comparative analysis of the proposed method with Original LBP or existing LBP variants as evidence of significant contribution.

Gardezi and Faye [24] fused Completed Local Binary Patterns (CLBP) with curvelet features for Normal/Abnormal classification using images from MIAS and IRMA datasets. Firstly, they computed CLBP features using rotationally invariant mapping; then, they computed curvelet features from the curvelet sub-band coefficient. The extracted CLBP and curvelet features were fused and passed to a classifier. They obtained a classification accuracy of 96%. The high accuracy is attributed to feature a fusion of different texture features and using images from two different datasets.

Paramkusham *et al* [25] extracted the Local Quinary Patterns to classify a mass into a normal or abnormal class, then geometric features to classify the abnormal tumor into a benign or malignant class. Using images from the IRMA dataset and SVM classifier, they attained classification accuracy of 99.27% and 79.13% for the normal/abnormal and benign/malignant classification, respectively.

From the previous work on the application of local texture features for breast cancer feature extraction in mammogram images and classification, it is evident that the extraction of breast cancer features sparingly employs local texture features, because of low classification accuracy levels achieved especially for benign/malignant classification. Studies that produced high accuracies either used privately owned datasets that cannot be accessed for verification or used a few images for testing. Therefore, there is a need to develop a better texture feature extraction approach for extracting breast cancer features, which consequently will improve classification accuracy.

3. LOCAL FEATURE DESCRIPTORS

3.1 Local Binary Patterns

Local Binary Patterns (LBP) is a texture descriptor based on local pixels. The original LBP proposed by Ojala *et al* [26] forms labels from an image pixel by thresholding a 3×3 circular neighborhood of each pixel with the center pixel. Given an image pixel, LBP code is generated by thresholding the neighborhood of each pixel by following the pixels along a circular path, either clockwise or counter-clockwise on distance R , which is the radius of comparison. Figure 1 shows a sample of 3×3 LBP operation.

47	72	30	0	1	0
98	50	100	1		1
7	44	24	0	0	0
(a)			(b)		

Figure 1: A sample 3 X 3 LBP operation with P=8 and R=1 for (a) original pixel values of the sample image region and (b) the resultant image after thresholding

LBP code is obtained by applying the equation (1) on the image region and then a thresholding function τ as defined in (2);

$$LBP_{P,R}(x_c, y_c) = \sum_{n=0}^{N-1} 2^n \tau(g_n - g_c) \quad (1)$$

$$\tau(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \quad (2)$$

Where N is the number of neighboring pixels, R the distance of neighboring pixels from the center pixel, g_c the gray-value of the center pixel, g_n for $n = 0, 1, 2, \dots, N - 1$ is the gray value of a neighboring pixel on a circular symmetric neighborhood of distance $R > 0$ and the function $\tau(x)$ is a thresholding function that generates a binary bit for a particular pixel.

Concatenating all the eight bits results in a binary number that is converted to its decimal equivalent and allocated to the central pixel as its LBP code. For instance, the resultant LBP code for Figure. 1 is 010100012=8110. Then it generates a histogram of an LBP-encoded image that represents micro-pattern structures in an image whose histogram is defined by equation (3):

$$H_i = \sum_{x,y} I(\text{LBP}_{p,R}(x_c, y_c), i), i = 0, 1, 2, \dots, 2^p - 1 \quad (3)$$

Where p is the number of texture patterns that LBP operator can encode, and I is defined by equation (4):

$$I(a) = \begin{cases} 1 & \text{if } a \text{ is TRUE} \\ 0 & \text{Otherwise} \end{cases} \quad (4)$$

Even though the LBP descriptor is computationally straightforward and robust to monotonic gray scale changes, it is sensitive to noise because it relies on the central pixel, and therefore, inadequately dependable in capturing breast cancer discriminant features.

3.2 Local Ternary Patterns

Reliance on the central pixel as a threshold in LBP makes it sensitive to noise. A small change of the central pixel greatly changes the LBP code. To overcome this challenge, LTP extends LBP by thresholding the pixels into (0, +1,-1) bits instead of (0, +1) bits. Using three value bits makes LTP robust to noise than LBP. Consider threshold constant k, center pixel c, and a neighboring pixel p. The LTP formula is given by equation (5):

$$S(P) = \begin{cases} 1, & \text{if } p > c + k \\ 0 & \text{if } c - k < p < c + k \\ -1 & \text{if } p < c - k \end{cases} \quad (5)$$

Where S (P) is the p^{th} neighbor containing the LTP code.

Once pixels are split into (0, +1,-1) bits, the negative values are eliminated by dividing the ternary pattern into upper and lower patterns. Extending LTP to a three-bit code makes it robust to noise and encodes more patterns; however, it utilizes one static threshold defined by the user for all datasets or all images in a dataset. Consequently, the chosen threshold, k is invariant to grey value variation, and there is no defined approach of selecting an optimum value for the threshold. This makes LTP highly subjective and biased to the selected threshold. Figure 2 shows a sample 3 x 3 image region for LTP code.

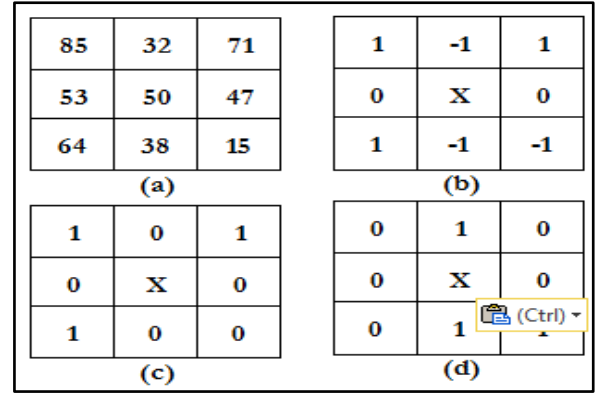


Figure 2: Generation of an LTP code with (a) showing the Original image pixels, (b) the LTP code with $k=\pm 5$ (c) Upper LTP pattern and (d) Lower LTP pattern

3.3 Local Directional Patterns

Both LBP and LTP use pixel intensity values to encode an image gradient, which makes feature discrimination inadequate and may lead to similar codes for two or more different patterns. Jabid *et al* [16] proposed Local Directional Patterns (LDP), which computes an 8-bit binary code for every pixel in the image by comparing edge responses of each pixel in different directions instead of raw pixel intensities like LBP and LTP. Some of the edge detectors used to encode an image gradient in LDP include; Kirsch [17], Robinson [16], and Sobel [20]. Among these edge detectors, Kirsch is the most popularly used mask because it is robust in directional edge detection since it considers all eight neighbors[27].

3.3.1 Kirsch edge detector

This is a first-order derivative edge detector that gets an image gradient by convolving 3×3 image regions with a set of masks. Kirsch [27] defined a nonlinear edge detector technique as defined in equation (6);

$$P_{(x,y)} = \max\{1, \max_{k=0}^7 [|5S_k - 3T_k|\}] \quad (6)$$

In which;

$$S_k = P_k + P_{k+1} + P_{k+2}$$

and

$$T_k = P_{k+3} + P_{k+4} + P_{k+5} + P_{k+6} + P_{k+7}$$

Where $P_{(x,y)}$ is the kirsch gradient, the subscripts P_k are evaluated modulo 8 and P_k [$k: 0, 1, 2, \dots, 7$] are eight neighboring pixels of $P(x, y)$ shown in Figure 3a.

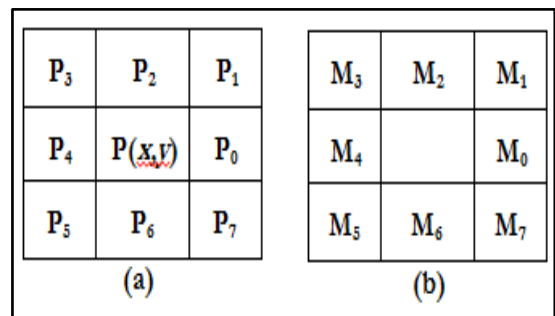


Figure 3: Illustration of (a) the Eight neighbors of pixel $P(x, y)$, and (b) their corresponding Kirsch mask

For each center pixel $p_{x,y}$, there are eight directional response values. High (absolute) response values depict the presence of

an edge or a corner in that particular direction. The interest of LDP is to determine k significant directional responses and set their corresponding bit-value to 1 and the remaining $8-k$ bits to 0. This binary code is converted to decimal and assigned to $p_{x,y}$ as its LDP code. This process is repeated for all pixels in an image to obtain an LDP-encoded image. The LDP descriptor then finds the image gradient towards a particular direction by convolving 3×3 image regions with the respective mask M_i as illustrated in Figure 4.

$$\begin{array}{cc}
 \begin{bmatrix} -3 & -3 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & 5 \end{bmatrix} & \begin{bmatrix} -3 & 5 & 5 \\ -3 & 0 & 5 \\ -3 & -3 & -3 \end{bmatrix} \\
 \text{(a) East } M_0 & \text{(b) North East } M_1 \\
 \\
 \begin{bmatrix} 5 & 5 & 5 \\ -3 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} & \begin{bmatrix} 5 & 5 & -3 \\ 5 & 0 & -3 \\ -3 & -3 & -3 \end{bmatrix} \\
 \text{(c) North } M_2 & \text{(d) North West } M_3 \\
 \\
 \begin{bmatrix} 5 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & -3 & -3 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ 5 & 0 & -3 \\ 5 & 5 & -3 \end{bmatrix} \\
 \text{(e) West } M_4 & \text{(f) South West } M_5 \\
 \\
 \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & -3 \\ 5 & 5 & 5 \end{bmatrix} & \begin{bmatrix} -3 & -3 & -3 \\ -3 & 0 & 5 \\ -3 & 5 & 5 \end{bmatrix} \\
 \text{(g) South } M_6 & \text{(h) South East } M_7
 \end{array}$$

Figure 4: Kirsch edge response masks in eight directions

For k significant responses, it calculates the LDP code shown in Figure 5 as;

$$LDP_k = \sum_{j=0}^7 b((m_j - m_k) \times 2^j) \quad (7)$$

In which b is:

$$b(a) = \begin{cases} 1, & \text{if } a \geq 0 \\ 0, & \text{Otherwise} \end{cases}$$

Where m_k is the k th significant directional response, and m_j is the response of Kirsch mask M_j . LDP operator generates C_k^8 distinct patterns in the LDP encoded image and uses a histogram $H(i)$ with C_k^8 bins to represent the input image of size $M \times N$ as;

$$H(i) = \sum_{m=0}^M \sum_{n=0}^N f(LDP_k(m, n), i) \quad (8)$$

Where f is:

$$f(p, i) = \begin{cases} 1, & \text{if } p = i \\ 0, & \text{Otherwise} \end{cases} \quad (9)$$

Where $f(p, i)$ is a logical function that compares if the LDP code at location $p(m, n)$ of the LDP-encoded image is equal to current LDP pattern i for all i in the range $0 \leq i \leq C_k^8$. The resultant histogram with dimensions $1 \times C_k^8$ represents the image. The resultant feature has spots, corners, edges, and texture information about the image. The limitation of LDP is not considering all responses in generating the LDP code. The discarded responses though not among top k could contribute to making LDP robust and pattern discriminative. Furthermore, some of the top k directional responses could be

in particular orientation like west-east.

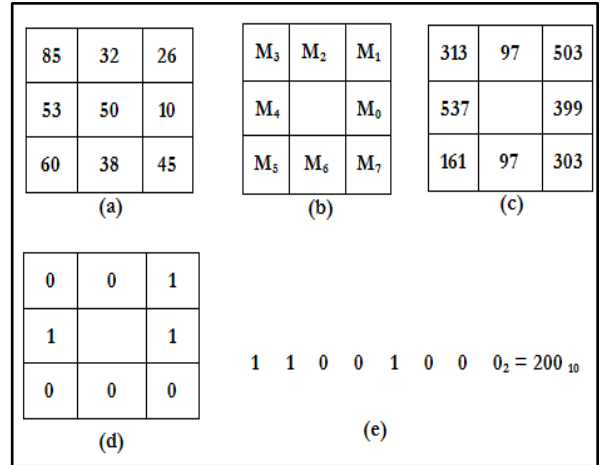


Figure 5: Process of encoding an image with an LDP operator when $k = 3$ for (a) Original Image region (b) Kirsch masks as presented in Figure 4, (c) Result of convolving each pixel in (a) with 8 Kirsch masks; (d) Pick top $k = 3$ significant responses, set their corresponding bit to 1 and the rest to 0 and finally (e) Resultant LDP code

4. METHODS AND MATERIALS

4.1 Materials

This researcher used mammogram images obtained from the Mammographic Image Analysis Society (MIAS) breast cancer dataset. A UK researcher group interested in understanding mammograms created the publically accessible dataset for scientific research. It has 322 digitized mammographic images of both breasts from 161 patients and provides background information on the class of abnormalities present in a mammogram. The class of abnormalities consists of normal/abnormal class, and based on the severity of abnormality; it splits the abnormal class into a malignant and benign class. Out of 322 images, 207 images are normal, and 115 images are abnormal. Among the 115 abnormal images, 64 are benign, and 51 are malignant[28]. Because the database has few images, there was a need to increase the number of images. This was achieved through data augmentation in which an original image was mirrored along the x-axis and the y-axis. These ensured the images are increased from 322 to 966 images.

4.2 Methods

In section 3, the existing local texture descriptors (LBP, LDP, and LTP) were presented. This study proposes a Local Directional Ternary Patterns (LDTP) texture descriptor as presented in subsection 4.2.1;

4.2.1 Local Directional Ternary Patterns

As earlier indicated, the Local Ternary Patterns (LTP) uses static threshold τ defined by a user for all datasets or images in a dataset. The chosen threshold is invariant to grey-value variations, static and lacks a defined technique of selecting an optimum value for the threshold. The Local Directional Patterns (LDP) only considers top k directional responses when encoding local texture in an image. Although the difference between a reference pixel and its neighbors derives an image gradient, it does not consider the central pixel, thereby deteriorating the discriminative power of the features extracted. Further, LDP disregards $8-k$ responses, which leads to a loss of subtle texture features. Herein, a Local Directional Ternary Patterns (LDTP) texture descriptor is proposed.

Unlike LDP and LTP, the LDTP descriptor considers the central pixel, takes into account all directional responses, and uses an adaptive threshold for 3×3 image regions when deriving an image gradient.

The LDTP applies Kirsch masks to compute responses for all the eight directions for central reference pixel in a 3×3 image region. It uses the sign of the directional responses to increasing the discriminative nature of the encoded image gradient. Given a 3×3 image region, LDTP first finds the absolute difference in grey-values between neighboring pixels and central reference pixel as defined by (10);

$$P_{i,j} = |P_{i,j} - P_c| \quad (10)$$

Where $P_{i,j}$, is the gray value at row i , column j , and P_c is the gray value of the central pixel.

The absolute differences are then convolved with the Kirsch masks to get directional responses. Figure 6 illustrates a 3×3 image region, the absolute difference in grey values, and corresponding directional responses.

85	32	26	35	18	24	7	175	215
53	50	10	3	0	40	-57	0	111
60	38	45	10	12	5	-241	-225	15
(a)			(b)			(c)		

Figure 6: LDTP process (a) Image region (b) Absolute differential values (c) Directional responses using kirsch masks

The directional responses are then normalized using the min-max normalization technique as shown in (11);

$$x_{i,j}^{norm} = \frac{x_{i,j} - \min}{\text{Max} - \text{Min}} \quad (11)$$

Where $x_{i,j}$ is the absolute response at index i, j and Max and Min are the maximum and minimum responses respectively, and $x_{i,j}^{norm}$ is the normalized value for $x_{i,j}$ responses.

The normalized values are in the range of $0 \dots 1$, but they add up to a value of more than 1. However, to ensure the values add to a maximum value of 1, the normalized values are passed through a softmax function given by (12) and illustrated in Figure 7(b) in comparison to the min-max technique in 7(a);

$$P_{i,j} = \frac{e^{x_{i,j}^{norm}}}{\sum e^{x_{i,j}^{norm}}} \quad (12)$$

Where $P_{i,j}$ is the likelihood for the presence of an edge towards a given direction, $e^{x_{i,j}^{norm}}$ is the exponential value of the normalized absolute responses at index i, j and $\sum e^{x_{i,j}^{norm}}$ is the summation of all the exponential values

0.5318	0.9091	1.0000	0.1192	0.1739	0.1904
0.3818	0	0.7636	0.1026	0	0.1503
0.0364	0.0000	0.5455	0.0726	0.0700	0.1209
(a)			(b)		

Figure 7: Normalization process with (a) Min-Max Normalization and (b) Soft-max technique

Once the normalization is done, the probability space is split into three parts for $-1, 0, +1$ bits for the generated ternary pattern and calculates two thresholds, T_p and T_n as shown in (13);

$$T_p = \frac{1}{4} \sum_{i=0}^3 L_i^T \quad (13)$$

$$T_n = \frac{1}{4} \sum_{i=0}^3 L_i^B$$

Where L_i^T is the i th top likelihood, and L_i^B is the i th bottom likelihood. For the likelihoods shown in Figure 7 (b), the thresholds are $T_p=0.1589$ and $T_n =0.0911$, respectively.

Where T_p is calculated as the average of the four highest values given by $0.1739, 0.1904, 0.1503$ and 0.1209 , while T_n is the average of the four lowest values given by $0.1192, 0.1026, 0.0726$ and 0.0700

LDTP code is calculated as;

$$\text{LDTP} = \sum_{i=0}^7 p(L_i - (T_p | T_n)) \times 2^i \quad (14)$$

$$p(\dots) = \begin{cases} 1, & \text{if } L_i - T_p \geq 0 \\ -1, & \text{if } L_i - T_n \leq 0 \\ 0, & \text{Otherwise} \end{cases}$$

Where L_i is a likelihood i as shown in Figure 7 (b)

For the likelihoods in figure 7(b), Figure 8 shows the resultant ternary pattern and the corresponding positive and negative LTP codes. The LDTP code is split into a positive and negative LTP code, as shown in Figure 8(b) and Figure 8(c), respectively

0	+1	+1	0	1	1	0	0	0
0		0	0		0	0		0
-1	-1	0	0	0	0	1	1	0
(a)			(b)			(c)		

Figure 8: Resultant Ternary Pattern with (a) Resultant LDTP code (b) Positive LTP code (c) Negative LTP code

The positive and negative gradients of an $M \times N$ image shown in Figure 8(b) and Figure 8(c) are represented as H_p and H_n

and defined as expressed in equation (15);

$$H_p(i) = \sum_{m=0}^M \sum_{n=0}^N f(LDP_p(m, n), i) \quad (15)$$

$$H_n(i) = \sum_{m=0}^M \sum_{n=0}^N f(LDP_n(m, n), i)$$

$$f(x, i) = \begin{cases} 1, & \text{if } x = i \\ 0 & \text{Otherwise} \end{cases}$$

Where $f(x, i)$ is a logical function that compares if the LDP code at location $x(m, n)$ of the LDP encoded image is equal to current LDP pattern i for all i in the image $0 \leq i \leq C_k^8$.

The resultant histogram has dimensions $1 \times C_k^8$, which represent the image. The two histograms could be fused and used for any pattern recognition application.

4.2.2 Classification

Once features were extracted, the researchers implemented different kernels for Support Vector Machines (SVM) and K-Nearest Neighbors (KNN) classifiers for the classification of two different labels. Level one classification classifies a tumor into a normal or abnormal class, and level two classifies an abnormal tumor into a benign or malignant class. This was done for the various local texture descriptors including LDTP and the results recorded for comparative purposes.

4.2.3 Validation and evaluation protocol

The proposed feature extraction technique is validated using ten-fold cross-validation and the SVM and KNN classifiers. The performance of the model is then mapped into a True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) using a confusion matrix from which performance metrics Accuracy, Sensitivity and Specificity are calculated. Accuracy defines the proportion of accurate results on the total number of instances tested, calculated by (16);

$$\text{Accuracy} = \frac{TN + TP}{FN + FP + TN + TP} \quad (16)$$

The sensitivity metric evaluates the effectiveness of a classifier by obtaining the proportion of positive examples correctly predicted as positive defined as in (17);

$$\text{Sensitivity} = \frac{TP}{FN + TP} \quad (17)$$

Specificity measures the effectiveness of a classifier by obtaining the proportion of negative examples correctly predicted as unfavorable defined as shown in (18);

$$\text{Specificity} = \frac{TN}{FP + TN} \quad (18)$$

5. RESULTS AND DISCUSSION

This research molded the breast cancer problem as a two-class classification problem in which it classified the breast tumor into a normal or abnormal class and then classified the abnormal tumor into a benign or malignant class. To visualize the performance of the classifiers, Table 1 and Table 2 shows the accuracy comparison of LDP, LTP, and LDTP with different kernels for SVM and KNN classifiers.

Table 1 and Table 2 show Normal/Abnormal classification accuracy comparison between LDP, LTP, and LDTP using different kernels for SVM and KNN classifiers. It is observed that in all instances LDTP outperformed LDP and LTP except for SVM with fine Gaussian. From the results achieved, it is

also evident that the highest classification accuracy of 91% was achieved by the weighted KNN.

5.1 Accuracy comparison of LDP, LTP, and LDTP

Table 1 and Table 2 presents the comparative results of the different local texture descriptors that exist against the proposed local directional ternary pattern. As observed in both tables, LDTP outperforms all the existing descriptors in classifying normal/abnormal breast cancer mammograms.

Table 1: Normal/Abnormal accuracy comparison of LDP, LTP, and LDTP using SVM classifier

Feature descriptor	SVM					
	Linear	Quadratic	Cubic	Fine Gaussian	Medium Gaussian	Course Gaussian
LDP	64.9	69.8	73.5	69.9	65.0	64.9
LTP	74.2	81.4	84.0	88.2	71.1	66.9
LDTP	80.7	82.7	84.2	80.7	80.7	80.7

Table 2: Normal/Abnormal accuracy comparison of LDP, LTP, and LDTP using KNN classifier

Feature descriptor	KNN					
	Fine	Medium	Course	Cosine	Cubic	Weighted
LDP	82.1	61.3	64.9	59.7	59.8	81.7
LTP	77.1	71.4	68.4	71.2	70.3	79.2
LDTP	88.7	79.5	80.4	79.1	79.1	91.0

Tables 3 and 4 show Benign/ Malignant classification accuracy comparison between LDP, LTP, and LDTP using different kernels for SVM and KNN classifiers. The LDTP performed better than LDP and LTP in all instances except for SVM with fine Gaussian where the LTP outperformed the LDTP. The highest classification accuracy was 90.6% achieved by the weighted KNN classifier.

Table 3: Benign/Malignant accuracy comparison of LDP, LTP, and LDTP using SVM classifier

Feature descriptor	SVM					
	Linear	Quadratic	Cubic	Fine Gaussian	Medium Gaussian	Course Gaussian
LDP	64.9	65.8	71.5	70.6	64.7	64.9
LTP	67.4	83.1	81.8	86.9	67.0	66.9
LDP	80.7	82.2	85.4	81.2	80.7	80.4

Table 4: Benign/Malignant accuracy comparison of LDP, LTP, and LDTP using KNN classifier

Feature descriptor	KNN					
	Fine	Medium	Course	Cosine	Cubic	Weighted
LDP	80.5	61.4	64.9	61.8	61.4	82.8
LTP	75.9	64.9	66.9	65.4	64.9	81.6
LDTP	88.7	79.1	80.4	79.5	79.7	90.6

Generally from Tables 1, 2, 3, and 4, it is evident that in terms of classification accuracy, the weighted KNN outperformed the SVM classifier in both instances of normal/abnormal and benign/malignant classification. However, the LDTP showed consistency in all the cases hence an indicator that a feature extractor plays a significant role in classification.

5.2 Sensitivity and Specificity comparison of LDTP using Weighted KNN and Cubic SVM

Figure 9 shows the sensitivity of the LDTP using weighted KNN and Cubic SVM. For the normal abnormal class, weighted KNN outperformed SVM, while for the Benign/Malignant class Cubic SVM and the Weighted KNN classifiers had equal sensitivity of 99.5%.

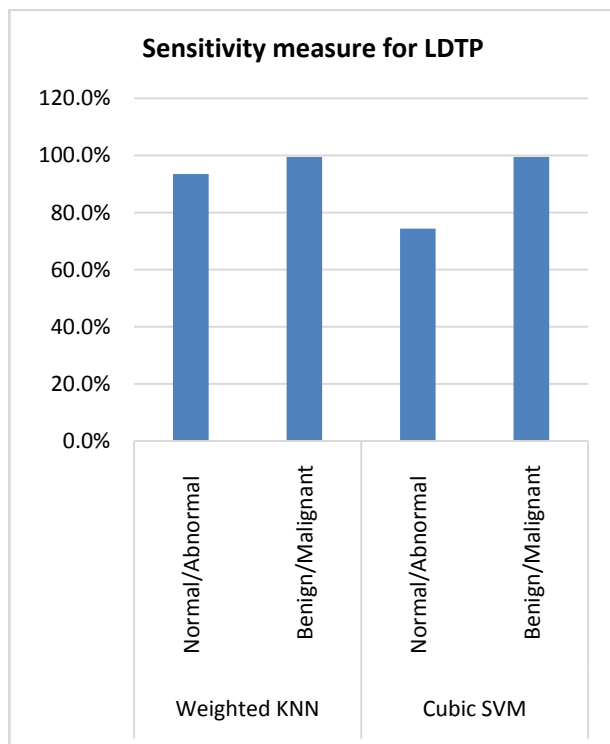


Figure 9: Sensitivity comparison of LDTP using Weighted KNN and Cubic SVM

Figure 10 presents the specificity of LDTP using weighted KNN and Cubic SVM. For both the Normal/Abnormal and Benign/Malignant classification, Cubic SVM achieved a higher specificity than weighted KNN.

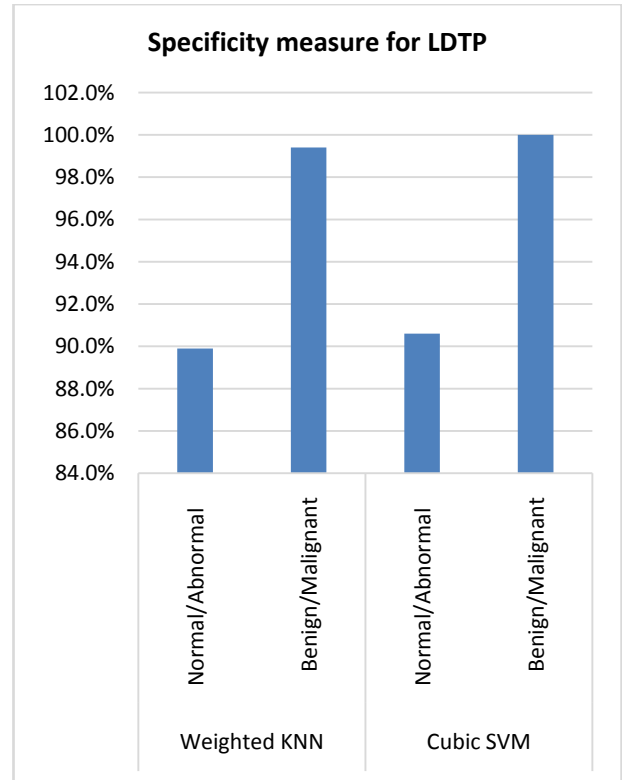


Figure 10: Specificity comparison of LDTP using Weighted KNN and Cubic SVM

Even though the weighted KNN performed better than Cubic SVM in terms of classification accuracy, Cubic SVM proved to be more effective at predicting positive and negative examples as positive and negative respectively than weighted KNN.

6. CONCLUSION

This study proposed a local texture feature extraction approach implemented for breast cancer classification problem. The researchers molded the breast cancer problem as a two-level classification problem in which it first classifies the breast tumor into a normal or abnormal class; then, it classifies the abnormal tumor into a benign or malignant class. The features are extracted through the LDTP approach using different kernels for SVM and KNN classifiers on 966 mammogram images from the MIAS dataset. For the normal/abnormal classification, the model achieved the highest accuracy of 91%, sensitivity of 93.5%, and specificity of 89.9% with the weighted KNN classifier. For the benign/malignant classification, the model attained the highest accuracy of 90.6%, the sensitivity of 99.5% with the weighted KNN, and specificity of 100% with Cubic SVM classifier. The LDTP approach has shown impressive results and can effectively be used to predict breast cancer tumor type with higher precision than LDP and LTP descriptors. The experiment performed reveals the validity of the approach using images from the MIAS dataset.

The prospect of this work is to evaluate the success of the developed approach by comparing how closely its findings match with studies conducted by other researchers in a similar area. Also, to use alternative classifiers and datasets to provide a full comparison between the effect of various classification approaches and the number of mammogram images selected. Further, to use a database with many images to demonstrate more feasibility of the approach.

7. ACKNOWLEDGMENTS

I wish to thank God Almighty for giving me a sane mind as I prepared the paper. Special appreciation to my supervisors Dr. Stephen Mutua and Dr. Raphael Angulu for their unlimited guidance and consistent counsel during the entire research.

8. REFERENCES

- [1] T. Ahonen, A. Hadid, and M. Pietika, "Face Description with Local Binary Patterns: Application to Face Recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037–2041, 2006.
- [2] V. Rao, V. Prasad, and M. Sugumaran, "Real-time video object detection and classification using hybrid texture feature extraction," *Int. J. Comput. Appl.*, pp. 1–8, 2018.
- [3] Y. Cao, S. Pranata, and H. Nishimura, "Local Binary Pattern Features for Pedestrian Detection at Night/Dark Environment," *IEEE Int. Conf. Image Process.*, pp. 2053–2056, 2011.
- [4] Z. Lai and H. Deng, "Medical Image Classification Based on Deep Features Extracted by Deep Model and Statistic Feature Fusion with Multilayer Perceptron," *Comput. Intell. Neurosci.*, p. 13, 2018.
- [5] M. Turk and A. Pentland, "Eigenfaces for Recognition," *J. Cogn. Neurosci.*, vol. 3, no. 1, pp. 71–86, 1991.
- [6] P. Belhumeur, J. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, 1997.
- [7] R. Angulu, J. R. Tapamo, and A. O. Adewumi, "Age Estimation with Local Ternary Directional Patterns," Paul M., Hitoshi C., Huang Q. *Image Video Technol. Image Video Technol.*, vol. 10749, pp. 421–434, 2018.
- [8] M. Eisa, A. Elgamal, R. Ghoneim, and A. Bahey, "Local Binary Patterns as Texture Descriptors for User Attitude Recognition," *Int. J. Comput. Sci. Netw. Secur.*, vol. 10, no. 6, pp. 222–229, 2010.
- [9] N. Kauser and J. Sharma, "Facial Expression Recognition using Lbp Template of Facial Parts and Multilayer Neural Network," in *IEEE International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)*, 2017, pp. 445–449.
- [10] Z. Camlica, H. R. Tizhoosh, and F. Khalvati, "Medical Image Classification via SVM using LBP Features from Saliency-Based Folded Data," *IEEE 14th Int. Conf. Mach. Learn. Appl.*, pp. 128–132, 2015.
- [11] P. Jasmine and R. Kumar, "Multi-resolution Joint LBP Histograms for Biomedical Image Retrieval," *Int. J. Comput. Appl.*, vol. 95, no. 3, pp. 23–27, 2014.
- [12] A. Suruliandi, K. Meena, and R. Rose, "Local binary pattern and its derivatives for face recognition," *IET Comput. Vis.*, vol. 6, no. 5, pp. 480–488, 2012.
- [13] X. Tan and B. Triggs, "Enhanced Local Texture Feature Set for Face Recognition Under Difficult Lighting Conditions," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1635–1650, 2010.
- [14] T. Jabid, H. Kabir, and O. Chae, "Local Directional Pattern (LDP) for Face Recognition," *Proc. IEEE Int. Conf. Consum. Electron.*, pp. 329–330, 2010.
- [15] R. A. Kirsch, "Computer Determination of the Constituent structure of Biological Images," *Comput. Biomed. Res.*, vol. 328, pp. 315–328, 1971.
- [16] G. Robinson, "Edge Detection by Compass Gradient Masks," *Comput. Graph. Image Process.*, vol. 6, pp. 492–501, 1977.
- [17] R. Park, "A Fourier interpretation of the Frei-Chen edge masks," *Pattern Recognit. Lett.*, vol. 11, pp. 631–636, 1990.
- [18] I. Sobel, "An Isotropic 3x3 Image Gradient Operator," *Res. Gate*, 2015.
- [19] A. M. Shabat and J. Tapamo, "A comparative study of the use of local directional pattern for texture-based informal settlement classification," *J. Appl. Res. Technol.*, vol. 15, no. 3, pp. 250–258, 2017.
- [20] C. Muramatsu, T. Hara, T. Endo, and H. Fujita, "Breast mass classification on mammograms using radial local ternary patterns," *Comput. Biol. Med.*, vol. 72, pp. 43–53, 2016.
- [21] R. Rabidas, A. Midya, J. Chakraborty, and W. Arif, "A Study of Different Texture Features Based on Local Operator for Benign-malignant Mass Classification," in *6th International Conference on Advances in Computing & Communications, ICACC*, 2016, pp. 389–395.
- [22] R. Rabidas, A. Midya, A. Sadhu, and J. Chakraborty, "Benign-Malignant Mass Classification in Mammogram using Edge Weighted Local Texture Features," in *SPIE*, 2016, vol. 9785, pp. 1–6.
- [23] N. Ponraj, J. Winston, Poongodi, and M. Mercy, "Novel Local Binary Textural Pattern for Analysis and classification of mammogram using Support Vector Machine," in *International Conference on Signal Processing and Communication (ICSPPC'17)*, 2017, pp. 380–383.
- [24] S. Jamal, S. Gardezi, and I. Faye, "Fusion of Completed Local Binary Pattern Features with Curvelet Features for Mammogram Classification," *Appl. Math. Inf. Sci. An Int. J.*, vol. 12, no. 6, pp. 1–12, 2015.
- [25] S. Paramkusham, K. Rao, and P. Rao, "Novel technique for the detection of abnormalities in Mammograms using texture and geometric features," in *International Conference on Microwave, Optical and Communication Engineering, ICMOCE*, 2016, pp. 150–153.
- [26] T. Ojala, M. Pietikainen, and D. Harwood, "A Comparative Study on Texture Measures with Classification Based on Feature Distributions," *Pattern Recognit.*, vol. 29, no. 1, 1996.
- [27] S. Lee, "Multilayer Cluster Neural Network for Totally Unconstrained Handwritten Numeral Recognition," *Neural Networks*, vol. 8, no. 5, pp. 783–792, 1995.
- [28] J. Sunkling, "The mammographic image analysis society digital mammogram database." 2014.