Journal of Science and Technology ISSN: 2362-1443
2023, Vol. 3, No. 1, 13-17
DOI: https://doi.org/10.3126/jost.v3i1.69059

Improving Breast Cancer Classification using Local Binary Pattern in Bit Plane Slicing and CNN

Anuj Ghimire
Nepal College of Information Technology
ghimireanuj01@gmail.com

Sanjeeb Prasad Panday*
Institute of Engineering, Tribhuvan University
sanjeeb@ioe.edu.np

*Corresponding Author

Article History:
Received: 7 July 2023
Revised: 9 October 2023
Accepted: 11 December 2023

Keywords—Breast Cancer, Bit Plane Slicing, LBP, CNN, Entropy, Texture Analysis, Recognition Rate

***Abstract*—Breast cancer is a most common women's health problem and considered the most dangerous disease among females. Breast cancer affects 8% of women at some time in their lives and is the main cause of cancer death, according to various statistics. The only way to consistently detect early and potentially curable breast cancer is to use mammography screening. The first step in the identification process is image segmentation. Most of the algorithm results for detecting textures or cancer cells work well on binary images. Texture analysis is crucial for distinguishing between normal and aberrant types. The images' bit planes are examined individually and in different ways. The classification takes into account the attributes found on these bit planes. The Local Binary Pattern operator can also be used to extract texture features. To distinguish between normal and abnormal tumors, the retrieved textural features can be trained and classified using the Convolutional Neural Network (CNN) classifier. The conclusion of the classification model depends on the data set and how the characteristic was extracted from the data set. This method is unbiased by the type of anomaly and can be used to supplement computer-assisted detection based on the detection of specific abnormal structures. Therefore, in many cases of images, the algorithm provides an accuracy of 93.11% or more.**

I. INTRODUCTION

Breast cancer is now a health problem that continues to threaten the health of women around the world. According to a World Health Organization report, around one million women are diagnosed with breast cancer annually, and half of them will die because it is frequently too late by the time professionals recognize disease [1]. Cancer in mammary gland is the most frequently encountered malignant tumor, and there is no way to prevent it. Early detection, on the other hand, can assist to decrease the difficulties caused by cancer and, as a result, the casualty rate can be minimized.

Imaging techniques are important in discovering abnormal spots in the breast that can be seen on a standard mammography. To diagnosis the disorder present in the breast an x-ray examination is done commonly known as mammogram. Mammography is safe technique which utilizes high quality x-rays with less radiation. Furthermore, it is commonly acknowledged that screen film mammography is the only efficient imaging modality for early diagnosis of breast cancer in women. This type of mammography is used to prevent breast cancer in women who have no symptoms of the disease. When any irregular is observed during screening of the breast further diagnosis is most. Among various feature extraction methods Local Binary Pattern (LBP) is one which can also be used for the texture classification [3]. The surrounding neighbor pixels are compared with other neighbor pixel in order to identify the pattern of the grey level changes.

II. RELATED WORK

The authors [6] developed a bit plane slicing approach where input images are divided into eight number of bit images with divergent texture information in every bit plane. They have presented a new method for automatically learning features from datasets of images using different planes of the image. Those bit planes consist different texture information which further used as the feature space to classify the cancer. The research further reduced the dimensions of the feature so that it can easily learned by the machine learning algorithms. The experiment was conducted on the dataset having both training and testing set of mammograms. Each picture class has its own subdirectory. The experiment concludes that every bit of the mammogram image has different accuracy level of cancer classification. Higher bit-planes contain picture structural features, while low order bit bear relatively less image feature and have high pictorial content randomness. On the CNN classifier, image features from one to eight-bit planes are compared to the original image feature. Overall, the experiment demonstrates that the proposed approach may considerably increase identification rate and classification performance on some bit-planes.

In the paper entitled "LBP Features for Breast Cancer Classification" by the mammogram image of breast is used for classification. Researchers have extracted Local Binary Pattern based feature from the image and classify those features using Support Vector Machines. Experiment was performed on the images extracted from MIAS and DDSM database securing 84% of classification accuracy [7].

Another study on breast cancer classification is done using multilayer perception algorithms where the classification of the mammary gland tumor is done on the basis of recurrences-event and no-recurrences events. Image database available by University of Medicine Center, Institute Of Oncology, Ljubljana, Yugoslavia was used which consists of two hundred eighty six datasets of 2 classes with two hundred one non-recurrence and eighty five recurrences-events classes. To fill the missing values on database approach named data cleaning is applied. The concept of data cleaning is to average attribute values of data which are on the same class. The main study of the researchers is to implement MLP for cancer diagnosis in breast [8].

Breast tumor is also classified using two alternative classifiers: Naive Bayes (NB) and K-Nearest Neighbor (KNN). The test was conducted using the Breast Cancer Dataset (BCD) from the University of California, Irvine (UCI). A comparison of the two new methods was done, and accuracy was assessed via cross validation. In terms of accuracy, the KNN classifier (97.51 percent) outperforms the NB classifier (96.19 percent) [9].

The textural properties of digital mammographic pictures are extracted using the LBP technique and the Gabor Filter. The textural features are fed into a SVM classifier, which correctly identifies benign, malignant, and regular breast cases with 96.72 percent, 84 percent, and 81.90 percent accuracy, respectively [10].

III. RESEARCH METHODOLOGY

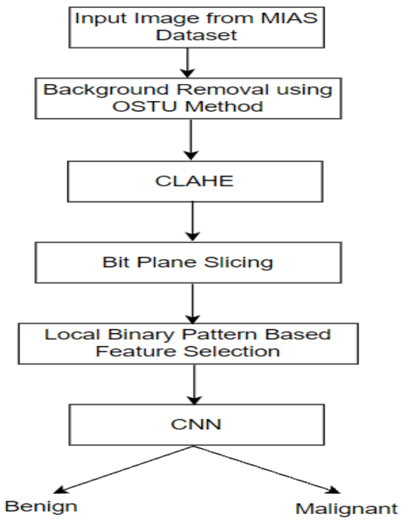

Fig. 1. Block Diagram

By introducing Contrast Limited Adaptive Histogram Equalization (CLAHE) and Local Binary Pattern (LBP) on individual bit planes of the imagery, the suggested framework aims to increase the efficiency of the CNN algorithm's categorization as shown in Fig 1. A series of photos from MIAS files is used to verify the model. The processing of the image in each step is explained as:

A. Image Pre-Processing

The breast and some surroundings make up the mammographic image. To isolate the breast image exclusively, we must first eliminate the background. Simple thresholding can be applied because the background is essentially uniform (black in color). The Otsu method [15] can be utilized for thresholding because of its excellent results in various image processing fields. The separation of the breast region and the removal of the pectoral muscle are the two main phases of pre-processing. The very first phase of the mammary region extraction procedure is to detach the breast from the surroundings, and phase two is responsible to eliminate the bounding boxes first from mammogram. (shown in Fig. 2).

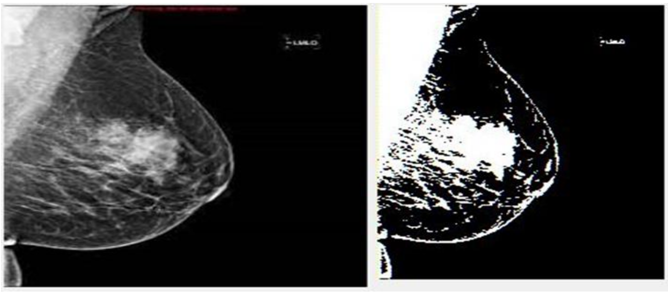

Fig. 2. Original image from MIAS dataset (left) output of Otsu method (right)

B. CLAHE

After the backdrop has been removed, the picture is further processed to account for the illumination restrictions encountered throughout the image capturing process. To solve the problem of uneven lighting conditions at the time of picture capture, CLAHE is applied to the image acquired using the Ostu technique.

The main focus of histogram equalization is to uniformly distribute the contrast in an image. This approach is typically used to increase the overall sharpness of a wide number of pixels, particularly only when image's usable information is captured by near intensity levels.

C. Bit Plane Slicing

Once OSTU and CLAHE is performed in the input image, then the next step involved in this study is to split in image into eight different bit planes using Bit Plane Slicing (BPS). BPS is method of dividing image into bit planes. In general, 256-gray level image is divided into eight-bit planes [16]. The main idea of BPS is to identify the relative significance of the all those bit planes. Also, different bit planes carry different level of information so to capture all the discriminating features of each bit BPS is obvious need.

D. *Feature Extraction using Local Binary Pattern*

Two-dimensional texture analysis has been studied in a variety of applications, including imagery recognition. The irregularity in the images is caused by factors such as orientation, scale, and lighting conditions. The way texture variations are encoded also affects how the image is analyzed. The LBP is a versatile picture pattern description that is both simple and effective. To capture the pattern of intensity of every bit level change, LBP based features are utilized [14].

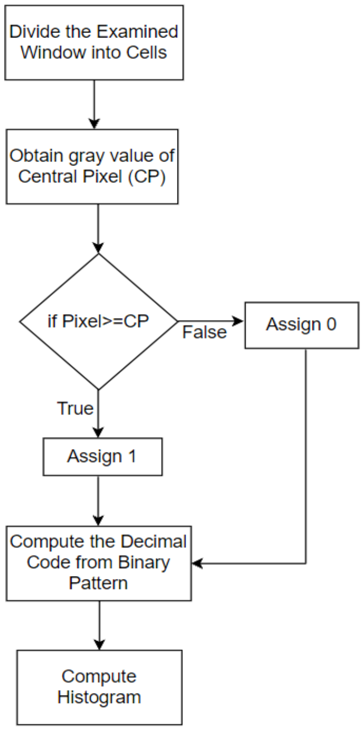

Fig. 4. Flowchart for LBP Computation

E. *Convolutional Neural Network*

Convoluting an image and looking for patterns is what image patterns do. CNN is variant of deep learning neural network. Such networks are equipped with different layers with very first layer capable to recognize edges and angles in the image. When we move on deeper in the network there will be transfer of the distinct and irregular patterns as well as complicated traits. CNNs are exceptionally good at spotting objects in images because of this characteristic. In this experiment the role of CNNs is to learn the features obtained by LBP procedure to distinguish between benign and malignant tumors present in women's mammary gland.

IV. EXPERIMENT AND RESULT

This chapter discusses the experiment and result i.e. the output achieved after doing the research.

A. *Dataset*

The MIAS dataset includes 6,000 mammography, each with its own resolution of 300 x 300 pixels with annotations such as back - ground matter composition, category, intensity, irregularity location, and abnormalities circumference. This database comprises 3000 harmless and 3000 cancerous specimens that were selected for screening.

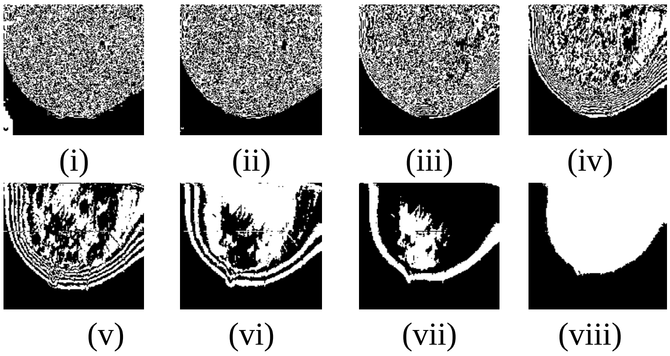

Fig. 3. Different Bit Planes of Mammogram

The images from the MIAS database were breakdown into train and test datasets. For training the model 80% of the images were considered and remaining 20% is used for testing purpose. The training dataset tuned the presented framework, while test model has been used to estimate performance indicators like precision, recall, and accuracy, as well as the F1-score. Following table 4.1 describes how datasets were divided.

TABLE I. DESCRIPTION OF TRAINING AND TESTING DATASETS

Ratio	Phase	Split Percent (%)	Calculation	Dataset Size
80:20	Training	80%	80% of 6000	4800
	Testing	20%	20% of 6000	1200

B. *Validation Criteria*

The validation of the model is performed using K-Fold cross validations (CV). For this the available datasets are sub-divided into two groups: training data and testing data.

C. *K-Fold CV*

Every instance of the large datasets is partitioned into folds, which are groups of comparable statistics (here, K=10). These folds are often used to fit the classifier, while the fold that is omitted out is being used to validate it. Each fold comprises a range of data, including one which serves as a training example and the subsequent as a validation set. Python modules indexing are used to construct these training and validation groupings.

D. *Result and Analysis*

BPS is used to separate the mammography image into eight different planes once it has been subdivided. The LBP is determined for every bit plane, and quantitative features are then extracted from the LBP representations. The feature map constructed by the above-mentioned process is 24-pixels long for every picture. R, G, and B signals are more difficult to handle and interpret than gray scale. So, to transform an image to 1s and 0s, choose a specified threshold, then transform pixels with levels more than or equal to the predefined threshold to white pixels, and pixels with values or less than equal to the threshold level to black pixels.

We bypassed the first step to compute the LBP because the bit planes are binary in nature. It's a valuable descriptor for texture-based classification problems since it characterizes local variance in gray levels. Local bit level changes are included in the LBP pictures computed from eight different bit planes. Bit level texture present in the mammogram can also be analyzed from statistical methods. Randomness and variance are major calculation during statistical processing which is very good at covering all the discriminating features. Entropies and fractal dimensions are computed to find randomness and variance in the mammogram sample.

TABLE II. RESULTS ACHIEVED ON THE TEST SET

Image Size	Sample Type	Obtained Accuracy	Precision	Recall	F1 Score
224 x 224 pixels	Benign	0.93	0.90	0.96	0.93
	Malignant		0.96	0.92	0.94

The verification of the model is done by comparing the model with other similar models and for the validation K-fold CV is done. Like other approaches measurement of specificity, sensitivity and accuracy is made for performance evaluation of the model.

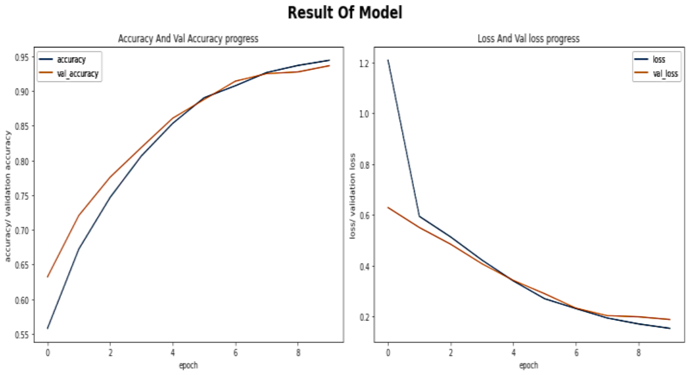

Figure 1Accuracy and Loss Curve of the Training Model

To examine how the model responds with image of different size, the model are tested with various image size on the mammography database. When the size of input image seems to be small the size of batch is chosen big i.e., of 64. For training ImageNet is used for 10 epochs. Following table shows the actual result obtained in train set. The curve analysis and classification report of the proposed model are also shown below.

TABLE III. CLASSIFICATION OF THE MODEL

Accuracy: 93.11

Classification Report

	Precision	Recall	F1-Score	Support
B	0.89	0.96	0.92	3865
M	0.96	0.91	0.94	4919
Accuracy			0.93	8784
Macro Avg	0.93	0.93	0.93	8784
Weighted Avg	0.93	0.93	0.93	8784

E. Evaluation using 10-fold Cross Validation

The proposed model is self-evaluated using the K-fold CV (Cross Validation). For this sample of the data sets are divided into group of equal sample sizes where the value of K is used as 10 i.e there are 10 folds of the dataset. These divided datasets are being used to validate the algorithm, whereas the fold that is kept out is being used to test it. Each fold comprises a variety of data, one of which serves as a training set and the other as a test set. Results of K-fold CV are shown in in table here in under:

TABLE IV. PRECISION, RECALL, ACCURACY AND LOSS RATE (IN %) FROM 10-FOLD CV

Obtained Values in each fold				
Folds	***Precision***	***Recall***	***Accuracy***	***Loss***
K1	96.98%	92.94%	99.09%	11.80%
K2	97.74%	98.28%	99.76%	5.27%
K3	98.71%	98.84%	99.84%	3.39%
K4	99.06%	99.35%	99.32%	2.41%
K5	97.66%	98.13%	99.75%	5.14%
K6	97.99%	98.52%	99.79%	4.44%
K7	90.91%	95.52%	97.77%	18.18%
K8	91.77%	94.47%	94.87%	16.41%
K9	94.56%	95.92%	99.33%	10.11%
K10	93.55%	95.98%	99.20%	11.06%
Average	**95.853%**	**96.725%**	**98.872%**	**8.722%**

V. CONCLUSION AND FUTUREWORK

This study aimed to improve the recognition rate of the conventional binary model that were used to analysis medical mammogram images by reducing the false positives. As a result, the introduction of a few new rotationally invariant binary patterns reduces false positives and generates grouped histogram feature vectors. There will be variance in the intensity level between close pixels in any image, depending on its resolution. The texture of an image refers to the characteristics of the image's intensity surfaces. One of the most essential applications of texture patterns is rotational invariance. One such pattern is proposed in this work that is unaffected by physical rotation of the image.

The focus is to enhance the net performance of the models by distinguishing malignancy using BPS and LBP. The presented binary patterns have a higher consistency rate than that of other binary patterns, with an accuracy of almost 93.11 percent. The idea of this method is to give the ultimate relevant advice, not to make a final conclusion about the presence of malignant alterations in an image. This method has the potential to boost medical diagnosis accuracy. When compared to other existing methods, this study seeks to provide a visual overview in tumor segmentation. Following is the comparative study with other findings.

TABLE V. COMPARATIVE ANALYSIS WITH OTHER FINDINGS

Authors	*Extracted Features*	*Classifier*	*Accuracy*
Guoming Chen, Yongchang Chen, and Zeduo Yuan[6]	BPS	CNN	84%
Pavel Kral and Ladislav Lenc[7]	BPS, LBP	SVM	84%
Nam S.H and Choi J.Y[22]	Multi-resolution LBP	SVM-RFE	90%
Proposed Approach	CLAHE, BPS, LBP	CNN	93.11 %

VI. Recommendation and Future work

The proposed outcome is unsatisfactory due to the low profile of the rendering GPU and the usage of Google photos that are modest in size and compressed in a way that is only suited for webpages. The majority of web photos are low-resolution. As a result, the visuals appear jagged, pixelated, and rough. When analyzing them, the system is unable to generate enough histogram. A greater resolution means more pixels, and more pixels equals more information, which is important in medical imaging for diagnosis. However, in non-commercial tasks, high-resolution photographs are not always available. As a result, the outcome of the investigation is affected.

This research can be extended to classify the mammogram images of other datasets as well as with the MRI images. Also, different variant of LBP such as rotation invariant can also be used to extract the feature present in the bit level to more precisely classify the malignant tumor. Furthermore, this experiment can also be tested with other classifiers to justify the importance of CLAHE in image enhancement and LBP in feature description.

Reference

[1] L.A. Altonen, R. Saalovra., P. Kristo, F. Canzian, A. Hemminki, Peltomaki P, R. Chadwik, A. De La Chapelle, "Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease", N Engl J Med, Vol. 337, pp. 1481–1487, 1998.

[2] Mehul P Sampat, Mia K Markey, Alan C Bovik, et al., "Computer-aided detection and diagnosis in mammography," Handbook of image and video processing, vol. 2, no. 1, pp. 1195–1217, 2005.

[3] Timo Ojala, Matti Pietikainen, and David Harwood, "A comparative study of texture measures with classification based on featured distributions," Pattern recognition, vol. 29, no. 1, pp. 51–59, 1996.

[4] Timo Ahonen, Abdenour Hadid, and Matti Pietikinen, "Face recognition with local binary patterns," in Computer Vision - ECCV 2004, vol. 3021 of Lecture Notes in Computer Science, pp. 469–481. Springer Berlin Heidelberg, 2004.

[5] Loris Nanni, Alessandra Lumini, and Sheryl Brahnam, "Local binary patterns variants as texture descriptors for medical image analysis," Artificial intelligence in medicine, vol. 49, no. 2, pp. 117–125, 2010.

[6] Guoming Chen, Yongchang Chen, and Zeduo Yuan, "Breast Cancer Image Classification based on CNN and Bit-Plane slicing", International Conference of Electrical Engineering and Computer Science, 2018.

[7] Pavel Kral and Ladislav Lenc "LBP Feature for Breast Cancer Detection", International Conference on Image Processing (ICIP), 2016

[8] Jasmir, Siti Nurmaini and Reza Firsandaya Malik, "Breast Cancer Classification Using Deep Learning ", International Conference of Electrical Engineering and Computer Science, 2018

[9] Meriem Amrane, Saliha Oukidi, Ikram Gagaoua, and Tolga Ensar, "Breast Cancer Classification Using Machine Learning ", IEEE, 2018

[10] Shishir Maheshwari, Saliha OukVivek Kanhangad, Sulatha V. Bhandary and U. Rajendra Acharya, "Automated Glaucoma Diagnosis Using Bit-plane Slicing and Local Binary Pattern Techniques ",Computers in Biology and Medicine, 2019

[11] Radia Touahri; Nabiha AzizI, Nacer Eddine Hammami, Monther Aldwairi, Farid Benaida. "Automated Breast Tumor Diagnosis Using Local Binary Patterns (LBP) Based on Deep Learning Classification", IEEE, International Conference on Computer and Information Sciences (ICCIS), (2019).

[12] J. Liu, X. Liu, J. Chen, and J. Tang. "Improved local binary patterns for classification of masses using mammography". In IEEE International Conference on Systems, Man, andCybernetics, pages 2692–2695, 2011..

[13] T. Ojala, M. Pietikäinen, and D. Harwood. "A comparative study of texture measureswith classification based on featured distributions". Pattern Recognition, 29(1):51–59,1996.

[14] Pelin Gorgel, Ahmet Sertbas, Niyazi Kilic, Osman N.Ucan and Onur Osman. "Mammographic mass classification using wavelet based Support VectorMachine", Journal of Electrical & Electronics Engineering,Vol.9, No.1, pp.867-875, (2009).

[15] Nobuyuki Otsu, "A threshold selection method from gray-level histograms," Automatica, vol. 11, no. 285-296, pp. 23–27, 1975.

[16] Gonzalex R.C and Woods R.E., "Digital Image Processing", Second Edition,Prentice-Hall, Englewood Cliffs, New Jersey, (2002).

[17] Gautam, Vijayshree & Lakhwani, Kamlesh. "Implementation of Non Shannon Entropy Measures for Color Image Segmentation and Comparison with Shannon Entropy Measures". International Journal of Science and Research (IJSR). 2. 391-394, 2013.

[18] Fan J, Zeng G, Body M and Hacid M.,"Seeded region growing: and extensive and comparative study, Pattern Recognition", Vol.26, pp.1139-1156, 2005.

[19] A.T. Azar, S.A. El-Said, "Performance analysis of support vector machines classifiers in breast cancer mammography recognition", Neural Comput. Appl. 24 (5) (2014) 1163–1177.

[20] Bird R, Wallace T and Yankaskas B. ,"Analysis of cancers missed at screening mammography", Radiology, Vol. 184, pp. 613-617, (1992).

[21] Yin F.F, Giger M.L, K. Doi K, Metz C.E, Vyborny C.J and Schmidt R.A. "Computerized Detection of Masses in Digital Mammograms: Analysis of Bilateral Subtraction Images", Medical Physics, Vol.18, No.5, pp.955-963,(1991).

[22] Nam S.H and Choi J.Y. "A method of image enhancement and fractal dimension for detection of microcalcifications in mammogram", In Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society 20, Vol.2, pp.1009-1012, (1998).

[23] Pragathi J and Patil H.T. "Segmentation Method for ROI Detection in Mammographic Images using Wiener Filter and Kittler's Method", International Journal of Computer Applications (0975 – 8887), pp.27-33, (2013).

[24] Li H, "Markov random field for tumor detection in digital mammography"IEEE Transactions on Medical Imaging, Vol.14(3),565–576, (1995).

[25] DDSM. Digital database for screening mammography.http://marathon.csee.usf.edu/Mammography/Database.html.