International Journal of Engineering & Technology, 7 (4.31) (2018) 322-325

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

HARIRAYA: a Novel Breast Cancer Pseudo-Color Feature for Multimodal Mammogram using Deep Learning

Azree Shahrel Ahmad Nazri[1]*, Olalekan Agbolade[2]

[1] *Department of Computer Science, FSKTM, UPM*
[2] *Institute Bioscience, UPM*
**Corresponding author E-mail: azree@upm.edu.my*

Abstract

Breast cancer is the leading cancer in the world. Mammogram is a gold standard for detecting breast cancer at earlier screening because of its sensitivity. Standard grayscale mammogram images are used by expert radiologists and Computer Aided-Diagnosis (CAD) systems. Yet, this original x-ray color provides little information to human radiologists and CAD systems to make decision. This binary color code thus affects sensitivity and specificity of prediction and subsequently affects accuracy. In order to enhance classifier models' performance, this paper proposes a novel feature-level data integration method that combines features from grayscale mammogram and spectrum mammogram based on a deep neural network (DNN), called HARIRAYA. Pseudo-color is generated using spectrum color code to produce Spectrum mammogram from grayscale mammogram. The DNN is trained with three layers: grayscale, false-color and joint feature representation layers. Empirical results show that the multi-modal DNN model has a better performance in the prediction of malignant breast tissue than single-modal DNN using HARIRAYA features.

Keywords: *Breast Cancer; convolutional neural network; mammogram; multi-modal features; false color*

1. Introduction

Breast cancers are increasing in the alarming rate worldwide. This particular type of cancer is the leading cancer among women [1]. Typically breast tissues are characterized by background tissues: fatty tissues, fatty glandular tissues, and dense glandular tissues. As for breast cancer tissues, they are divided into two severities of abnormalities: benign and malignant [2]. Benign tumor cells are different only slightly in behavior and appearance from their tissue of origin while Malignant on the other hand describes tissues that grow rapidly and capable of metastasizing. Both abnormalities are categorized into six classes: Calcification, Spiculated masses, Ill-defined masses, Architectural distortion, Well-defined/circumscribed masses and Asymmetry. The major classes of abnormalities in breast cancer are calcification and masses [2]. Calcifications grow inside the soft breast tissue as a small lump of calcium deposits. Masses is typically difficulty to detect due to its poor exhibition of image contrast [3]. Architectural distortion and asymmetry are the most difficult tissue abnormalities to detect. These two abnormalities do not have clear features or characteristics as observables. These problems in detecting tissue abnormalities in both abnormal tumors are great challenge to radiologists and in medical image processing. The most operative and low-level ways have been developed by digital mammograms for the detection of abnormal tumors.

Radiologists mainly use their eyes with the help of breast cancer diagnostic to discern cancer they screen the mammograms. Yet, cancer is not easily detected in many cases by the naked eyes as a result of poor imaging conditions. For instance, the appearance of unstable and subtle breast cancer on mammogram in their early stage which may cause the radiologists to misjudge the abnormality if only diagnosis by experiences is used [2], making the performance to vary from 65% to 88%. These challenges faced by human limitations can partially be overcome by using CAD system which typically serves as alternative opinion to radiologist. The major goal of CAD system is to detect and screen the breast tissue abnormalities with improved accuracy and reliability using computer vision and machine learning approaches. The typical system flow of CAD system is shown in Figure 1. Several variants of CAD system for detection of abnormalities have been proposed and developed. A CAD system that utilized Gabor filters and undecimated wavelet transform have been suggested [4], [5].

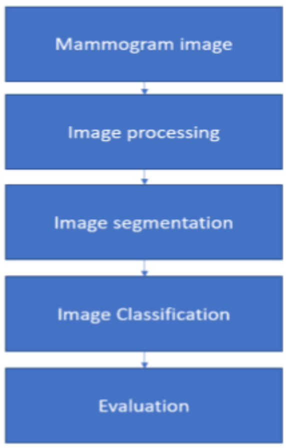

Fig. 1: Block diagram of a CAD system

The extraction of multi-resolutional features which quantify the mass shapes was performed by the proposition of DWT modulus maxima technique with wavelet based sub-band image decomposition [6], [7]. One study introduced the combination of supervised learning with a decimated WT and another study proposed a CAD variant that based on curvelet transform [8], [9]. A variant of CAD system used WT based on statistical methods and the Daubechies wavelet family such as db2, db4, db8 and db16 was proposed by [10]. SVM has been used for classification and CT for feature extraction while another study proposed PCA for feature extraction method to determine breast cancer [11], [12]. Micro-calcification (MCC) was detected by a CAD system using SVM classifier [13]. This study proposes a breast cancer recognition based on two combined features. The first feature is based on grayscale mammogram and the second feature is generated from spectrum mammogram. This multi-modal fusion approach helps improve the performance of breast cancer recognition accuracy and the proposed DNN model with feature-level concatenation is based on three-layer groups: grayscale layer, false-color layer and joint feature representation layer.

2. Main Body

The schematic diagram in Figure 2 shows a method of joining feature representation with the proposed DNN model. In Figure 2, features from grayscale and false-color mammogram are extracted using pre-trained convolutional neural network (CNN) VG16 architecture. This VGG16 architecture produces two feature vectors: grayscale feature vector and spectrum feature vector. These two feature vectors are inputted into the proposed DNN model. The DNN model uses these features for breast cancer classification.

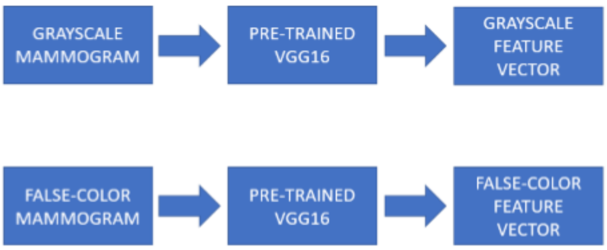

Fig. 2: Overview of the proposed method

Figure 1 shows a CAD system flow. The process of transforming grayscale mammogram into spectrum images is called pseudo-color technique. Mammography is a technique for recording x-ray images. The process of transforming a grayscale into a spectrum for a pseudo-coloring operation is done by a process that replaces the gray-scale values with color using pre-defined lookup tables (LUTs). An image is represented in form of matrix in computer vision and a row of pixel is transformed into number matrix with values between 0 and 255. The brightness value in 8-bit image ranging from 0 (black) to 255 (white) is assigned by default gray-scale LUT. For pseudo-colorizing, the chosen LUT's color code is "spectrum", which 0 indicates red and 100 represents green. The "spectrum" color code is shown in Figure 3.
Further, Figure 3 shows two mammogram samples: normal and malignant. The normal and malignant breast samples are shown both in grayscale and spectrum. The pseudo-colorizing transforms grayscale image into "spectrum" image. An important point that can be discerned from Figure 3 is the spectrum color clearly exposes the lesion location on the image compared to grayscale image.

2.1. Dataset Information

The mammogram images used in this study are collected from public domain database; Mammographic Image Analysis Society (MIAS). These datasets contain 322 images that composes of normal, benign and malignant. Benign and malignant are abnormal breast tissues. There are 207 mammogram images for normal breast tissues and 52 mammogram samples for malignant breast tissues. This study does not consider benign breast tissue case. The composition of the MIAS dataset creates imbalance dataset. The imbalance dataset affects many empirical measurements such as accuracy. Many techniques can be used to deal with imbalance dataset and this study chose to use ratio technique that will be discussed later.

2.2. Convolution Neural Network and Deep Learning Network

There are two problems when using imbalance dataset which are over-fitting and entropic capacity problem. The exposure of a model to a very few samples is known as Overfitting when the learning patterns do not have a new data generalization. While the capacity of the entropy reveals the amount of information the model is capable of storing. Several ways to handle these challenges are proposed in machine learning: (1) the number parameters used in the model, (2) the use of weight regulation such as L1 and L2 and (3) the leverage of a pre-trained network on a large dataset. This study opts to use ratio technique. In the other words, this study leverages a network pre-trained architecture on a large dataset to investigate the effect by feeding the architecture with data in ratio. In convolutional neural network (CNN), there are different pre-trained CNN architectures that can be used and, this study utilizes VGG16 pre-trained CNN architectures. Figure 4 shows the architecture of VGG16. VGG16 CNN architecture is used in this study to overcome over-fitting and entropic capacity issues that originate from small dataset used in this study. There are 322 breast samples (normal and malignant) investigated in this study which is too small to train VGG16. Figure 4 shows VGG16 architecture consists of 6 convolutional blocks and one fully connected block. To train VGG16 with 322 MIAS dataset, 4 convolutional blocks will be frozen, and the last block will be fine-tuned and trained with 322 MIAS dataset while FOC block will be fine-tuned and reduced to 2 dense layers. The feature vectors of grayscale and false color mammogram are extracted from the last VGG16 block as shown in Figure 4.

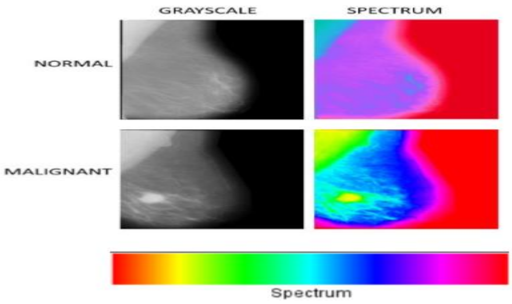

Fig. 3: Normal and malignant breast cancer are shown in grayscale and Spectrum codes. The Spectrum color is generated from pseudo-colorizing technique.

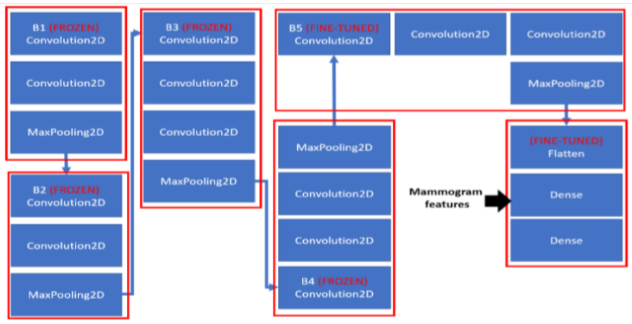

Fig. 4: Transfer learning VGG16 for feature extraction from grayscale and false-color mammogram. Red rectangles are CNN blocks that composes of CNN components. The first fully connected layer produces a mammogram feature vector.

The feature vectors extracted from VGG16 are then used in a DNN model. A DNN model used for learning purpose from multiple datasets in multi-modal data integration approaches. A DNN model with feature-level data integration method generally performs more satisfactorily than a single DNN [14]–[16]. Figure 5 shows the structure of the proposed DNN. The DNN model consists of 4 layers: grayscale layer, Spectrum layer, and joint feature representation layer. Both the grayscale and the spectrum layers have independent operation. To perform multilevel feature abstraction and representation, each feature layer uses the corresponding feature group consisting of data with properties similar to its input. These layers contribute greatly to the extraction of features from mammogram images by concatenating the feature vectors of the layers together with joint feature representation layer to integrate the features into a unified feature. These two layers are composed of 256, 256 and 128 neurons and the joint layers have three layers with sizes of 1024, 1024 and 2. For dropouts and activation function grayscale, spectrum and joint feature layers applied rectified linear units. The SoftMax with Loss Layer, from a Caffe framework, is used as the loss layer because the DNN is a binary classification

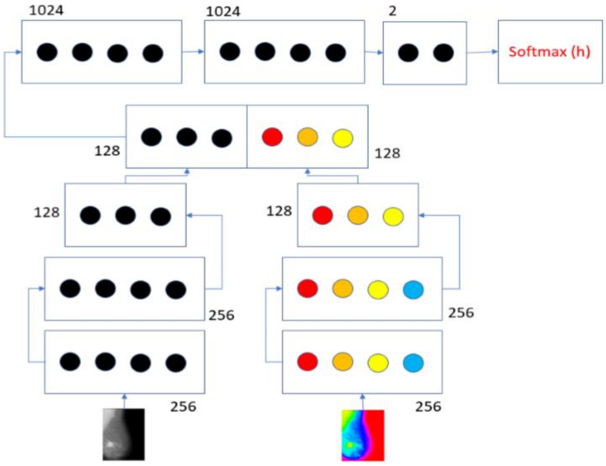

Fig. 5: The structure of DNN. It consists of grayscale feature, Spectrum feature and joint feature representation layers and softmax classifier.

2.3. Results and Discussion

Accuracy, precision, recall and AUC are used as performance measurements. Two single DNN and one SVM models are developed to analyze the performance of the proposed DNN models by comparing it with the three in term of accuracy, precision, recall and AUC. Two of these models are a DNN model using one feature-level which is different from our proposed DNN models that employs DNN using feature-level data integration approach. The single feature-level DNN models use either grayscale feature or Spectrum feature while the proposed DNN use both grayscale and Spectrum features.

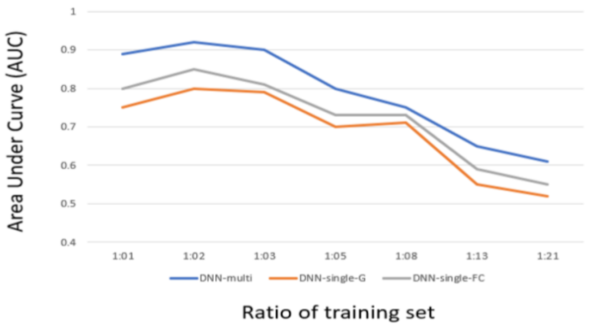

Fig. 6: Performance evaluation results according to the ratio of training set. The ratios are 1:1, 1:2, 1:3, 1:5, 1:8, 1:13 and 1:21

At first, we evaluate the performance of the two single DNN models by extracting features from grayscale and Spectrum by VGG16 in ratio data split. The split balance data is 1:1 (normal: malignant) However, for comprehensiveness, four other ratios are chosen: 1:2, 1:3, 1:5, 1:8. 1:13 and 1:21. These ratios follow Fibonacci series. This step requires VGG16 fed with mammogram images to extract features from the images. The features extracted by VGG16 are fed to DNN single feature and DNN multi-modal feature across 5 random train/test splits for each ratio of training set. Figure 6 shows the performance of the three DNN models based on AUC metric. Overall, in Figure 6, the performance of each model decreases when the dataset become more imbalance. The peak performance is when the ratio between normal and malignant is 1:2, respectively. This is true to both DNN multimodal and DNN single-model cases. The second observation that can be discerned from Figure 6 is the performance between DNN grayscale-feature and DNN false-color feature. Evidently, the performance of DNN Spectrum feature is better than DNN grayscale feature. The last significant information given by Figure 6 is the performance between DNN multi-modal feature and DNN single feature. DNN multi-modal feature shows better performance than DNN grayscale and DNN Spectrum features. The only explanation is that the DNN multi-modal model combines features from both grayscale and Spectrum color. As can be seen from Figure 6, Spectrum feature boosts DNN better than grayscale feature and it makes sense since the combination of these features will have additive effect on prediction power of the classifier.

Eventually the performance of DNN multi-modal is analyzed by comparing it with SVM. This SVM is fed with multi-modal grayscale/false-color features, grayscale feature only and false-color feature only. SVM is a general machine learning framework. The SVM is trained while varying the parameter c to obtain the optimal value. The performance of DNN and SVM is measured by calculating the accuracy, precision, recall and area under curve (AUC). Table 1 shows the empirical results of the performance evaluation between DNN and SVM. Table 1 suggests that SVM single feature is dominated by Spectrum feature in term of accuracy, precision, recall and AUC. While SVM multimodal is better than SVM single feature in term of accuracy, precision, recall and AUC. The performance of DNN multimodal is even better than all approaches in term of precision, accuracy, recall and AUC.

Table 1: Results of performance evaluation SVM and DNN

Model	Accuracy (%)	Precision (%)	Recal (%)	AUC
DNN multi-modal with HARIRAYA features	85	78	88	0.9
SVM gray-scale multi-modal	80	75	85	0.8
SVM gray-scale feature	69	80	20	0.6
SVM HARIRAYA feature	75	82	50	0.71

3. Conclusion

In essence, this study suggests that the proposed HARIRAYA features that combine Spectrum and grayscale show better performance than grayscale alone which increases predicting power of the classifier. This is logical because Spectrum mammogram consists of a dozen of colors that give more information than grayscale. Moreover, the pre-trained VGG16 architecture is able to extract the rich features from both grayscale and Spectrum mammogram. Unifying these features into multi-modal representation boosted the performance of classifier even better in term of accuracy, precision, recall and AUC.

Acknowledgement

This research is funded by Fundamental Research Grant Scheme.

References

[1] L. Hadjiiski *et al.*, "Breast masses: computer-aided diagnosis with serial mammograms.," *Radiology*, vol. 240, no. 2, pp. 343–356, 2006.

[2] T. Balakumaran, I. L. A. Vennila, and C. G. Shankar, "Detection of Microcalcification in Mammograms Using Wavelet Transform and Fuzzy Shell Clustering," Feb. 2010.

[3] J. Dheeba, N. Albert Singh, and S. Tamil Selvi, "Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach," *J. Biomed. Inform.*, vol. 49, pp. 45–52, 2014.

[4] T. Bhangale, U. B. Desai, and U. Sharma, "An unsupervised scheme for detection of microcalcifications on mammograms," *Proc. 2000 Int. Conf. Image Process. (Cat. No.00CH37101)*, vol. 1, pp. 184–187, 2000.

[5] R. N. Strickland and Hee Il Hahn, "Wavelet transforms for detecting microcalcifications in mammograms," *IEEE Trans. Med. Imaging*, vol. 15, no. 2, pp. 218–229, Apr. 1996.

[6] T. C. Wang and N. B. Karayiannis, "Detection of microcalcifications in digital mammograms using wavelets.," *IEEE Trans. Med. Imaging*, vol. 17, no. 4, pp. 498–509, 1998.

[7] L. M. Bruce and R. R. Adhami, "Classifying mammographic mass shapes using the wavelet transform modulus-maxima method," *IEEE Trans. Med. Imaging*, vol. 18, no. 12, pp. 1170–1177, 1999.

[8] H. Yoshida, K. Doi, and R. Nishikawa, "Automated detection of clustered microcalcifications in digital mammograms using wavelet transform techniques," *SPIE Image Process.*, 1994.

[9] M. M. Eltoukhy, I. Faye, and B. B. Samir, "Breast cancer diagnosis in digital mammogram using multiscale curvelet transform," *Comput. Med. Imaging Graph.*, vol. 34, no. 4, pp. 269–276, 2010.

[10] C. L., V. Ponomaryov, and J. Sanchez R., "Detection of Microcalcifications in Digital Mammograms Images Using Wavelet Transform," in *Electronics, Robotics and Automotive Mechanics Conference (CERMA'06)*, 2006, vol. 2, pp. 58–61.

[11] J. Manzano-Lizcano and C. Sánchez-Ávila, "A microcalcification detection system for digital mammography using the contourlet transform," in *on Computational & …*, 2004.

[12] J. F. D. Addison, S. Wermter, and G. Z. Arevian, "A Comparison of Feature Extraction and Selection Techniques," *Proc. Int. Conf. Artif. Neural Networks*, pp. 212–215, 2003.

[13] M. Sharkas, M. Al-Sharkawy, and D. A. Ragab, "Detection of Microcalcifications in Mammograms Using Support Vector Machine," in *2011 UKSim 5th European Symposium on Computer Modeling and Simulation*, 2011, pp. 179–184.

[14] Y. Zheng, "Methodologies for Cross-Domain Data Fusion: An Overview," *IEEE Trans. Big Data*, vol. 1, no. 1, pp. 16–34, 2015.

[15] J. Ngiam, A. Khosla, M. Kim, J. Nam, H. Lee, and A. Y. Ng, "Multimodal Deep Learning," *Proc. 28th Int. Conf. Mach. Learn.*, pp. 689–696, 2011.

[16] N. Srivastava and R. Salakhutdinov, "Multimodal Learning with Deep Boltzmann Machines," in *Advances in neural information processing systems (NIPS)*, 2012, pp. 2222–2230.