Fine Grained Classification of Mammographic Lesions using Pixel N-grams

Kulkarni P. S, MIT WPU, Federation University, Australia, *Member, IEEE*

Abstract— Breast cancer is the most common type of cancer worldwide. Early diagnosis of breast cancer can result in better treatment options increasing the survival chances of a patient. Automated or computer aided detection of breast cancer is applied in order to improve the accuracy and turnover time. However, the accuracy of automated detection systems can still be improved. Most of the efforts in the computer aided detection systems classify the images into cancerous and non-cancerous categories. The aim of this paper is to classify the mammographic lesions into three categories namely circumscribed, speculation and normal. The novel Pixel N-gram features have been used for classification of these lesions. Pixel N-grams are originated from character N-gram concept of text categorization. Classification performance is noted in order to analyse the effect of increasing N and effect of using different classifiers (MLP, SVM and KNN). It was observed that the classification performance increases with increase in N and then starts decreasing again. Moreover, classification performance achieved using MLP classifier was better than the performance using SVM or KNN classifiers.

Keywords—Classification, Mammograms, N-grams, SVM, MLP, KNN

I. INTRODUCTION

Breast cancer is the most common type of cancer worldwide representing nearly a quarter (25%) of all cancers. There is a significant increase in the incidence and cancer-associated morbidity and mortality in Indian subcontinent as described in global and Indian studies [1]. Among Indian females the breast cancer rate as high as 25.8 per 100,000 women and mortality 12.7 per 100,000 women has been reported [2]. The main reasons for high mortality are lack of adequate breast cancer screening, inappropriate diagnosis of disease stage and unavailability of appropriate medical facilities.

Mammography is a reliable technique for early detection of breast cancer. However, interpretation of mammographic images is a demanding job for the radiologists. Radiologists interpretation accuracy is subjected to the perception and interpretation errors. In order to reduce the perception errors automated detection/diagnosis using computerized image processing algorithms have been worked out.

The main signs of abnormalities which suggest possibility of breast cancer are masses and calcifications. Masses and calcifications can be cancerous or non cancerous. Most of the work done in the mammographic classification includes region classification as normal or abnormal. However, the area of fine grained classification of mammographic lesions (circumscribed, speculation, normal) which can be used as diagnostic support tool is not very well explored.

The aim of this paper is to use novel Pixel N-gram features for fine grained classification of mammographic lesions. Pixel N-gram features are inspired from the character N-gram concept of text categorization. Basically, character N-grams are phrases formed by N consecutive characters in a sentence[3]. For example, the 3-grams in the phrase “the fox” are “the, he_, e_f, _fo, fox”. The character N-grams have been found to be very efficient [4]. Pixel N-grams have been demonstrated to be effective for classification of mammograms into normal and abnormal categories[5, 6]

This paper aims at exploring the efficacy of novel Pixel N-grams technique for fine grained classification of mammographic lesions. The paper is organised as follows. Section I introduces the need for mammographic classification, Section II describes related work in the area, methodology is detailed in section III, Experimental results are analyzed in section IV and section V concludes the paper.

II. LITERATURE REVIEW

Automatic detection of breast cancer had been practiced using classification of mammograms into cancerous and non cancerous categories. As mentioned earlier presence of abnormalities such as masses and calcifications indicate the possibility of breast cancer. The masses can be benign or malignant (cancerous). The malignancy of the masses depends on the characteristics such as shape, size and boundary of the mass. The classification of masses according to their size and boundary has not been explored well. This classification of masses into classes such as circumscribed, speculation and normal is really useful for the automated diagnosis.
Various features have been used for classification of mammograms. Mass identification using Haralick’s features can be observed in work of Bovis et al.[7]. Khuzi et al [8] used Grey Level Co-occurrence Matrix (GLCM) features.

Using the contrast, energy and homogeneity accuracy of 84% was observed. Descrete wavelet transform (DWT) and Grey Level Co-occurrence Matrix (GLCM) features have been used by Beura et al.[9]. The accuracies of 98.0% and 94.2% were observed for normal/abnormal and benign/malignant classification respectively. Another benign/malignant classification using texture features can be seen in the work of Rocha et al. [10]. The best result achieved using SVM classifier was 88.31%. Mass detection using Gabor filter bank is observed by Hussian et al.[11]. Shape, edge sharpness and texture features are used by Mu et al.[12].

The classifiers used also makes effect on the classification performance. Classification of mammographic images has been achieved using various classifiers such as multilayer perceptron (MLP), support vector machine (SVM), k-nearest neighbor (KNN). The classification accuracy is found to be better by use of ensemble technique which incorporates combination of classifiers.[13]

The global features such as texture, shape and boundary fail to describe the global features of the lesion. The Bag of visual words is a concept originated from text categorization context where the image is represented with the help of number of occurrences of various visual words present in it. The visual word is also known as Texton. Mammographic classification using Texton approach is successfully implemented by Li et al. [14]. Another similar concept inspired from text categorization is Pixel N-grams[5]. The pixel N-grams have been successfully applied for classification of mammograms into normal and abnormal categories[6]. Also, the pixel N-grams have been very effective for texture categorization[15]. It has also been demonstrated that the Pixel N-grams are size, location and resolution invariant for shape classification [16]. Further, Pixel N-grams are shown to be computationally efficient than co-occurrence matrix features[17].

In this paper the pixel N-gram features are used for fine grained classification of mammograms into circumscribed, speculation and normal classes.

III. Materials and Methods

We conducted couple of experiments to test the efficacy of Pixel N-grams for fine grained mammographic lesion classification. First experiment was designed to find out the optimum value of N. Optimum value of N is the value of N for which we get the maximum classification performance. The second experiment was designed to figure out the best classifier for mammographic lesion classification.

Pixel N-gram feature extraction program was developed in Matlab 7.9 Software. Weka 3.6 data mining software was used for all the classification experiments. Generalization was estimated using 10 fold cross validation technique. The machine used for the experiments is i5-4370S CPU @2.90GHz PC with windows 7 (64 bit) operating system.

A. Dataset

We have used the Mini-MIAS database provided by Mammographic Image Analysis Society as benchmark [18]. The mammograms have been reduced to 200 micron pixel edge and clipped or padded so that every image is 1024 × 1024 pixels. Abnormality area or region of interest (ROI) is specified with the help of x, y image coordinates of centre of abnormality and the radius of a circle enclosing the abnormality. Information about the type of abnormality such as calcification, mass (circumscribed, speculated or ill-defined), architectural distortion or asymmetry is specified in the database. Type of Background tissue such as Fatty, Fatty-glandular, Dense-glandular has also been mentioned for each mammogram. Images of normal mammograms are also present in the database.

B. ROI Extraction

Region of Interest of size 140 × 140 pixels around the centre of the abnormality is cropped. For normal mammograms a square of size 140 × 140 pixels was cropped from the centre of the mammogram. The sample Regions of Interests (ROI) extracted from the miniMIAS dataset can be seen in the Figure 1.

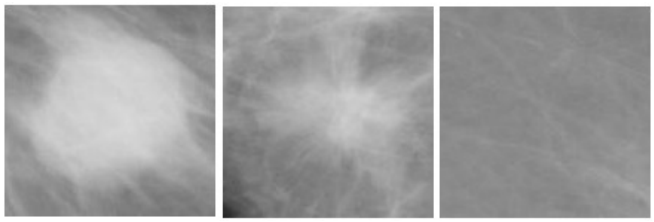

Figure 1. Sample ROIs from miniMIAS

C. Grey Scale Reduction

It is highly desirable to have low computational cost for mammographic classification algorithm in order to increase the efficiency of the radiologists. By reducing the images in grey scales the number of N-grams to be computed are reduced. Thus the computational cost can certainly be reduced.

Also, it is clear that the possible number of N-grams varies with the grey scale reduction. However, it has been observed that not all the possible N-grams are present in the given corpus which further reduces the dimension of the feature vector.

The cropped ROIs are grey scale reduced using 8 grey levels as this is the optimum grey level concluded by earlier experiments on miniMIAS dataset [5].

D. Pixel N-gram Computation

Image representation is very important for image classification. Pixel N-grams representation of image is used for mammographic classification here. Pixel N-grams are nothing but sequence of N consecutive Grey level pixels in an image. A sliding window of size N is used so that we are looking at phrase formed by N consecutive pixel intensities at a time. The count of how many times the visual phrase is repeated in an image would be used as N-gram features.

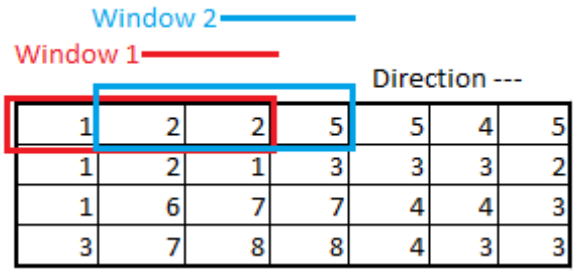

Figure 2. Pixel N-gram feature computation

Figure 2 shows the computation of three-gram features in horizontal direction. Likewise, N can be varied and different directions can be taken into consideration. For detailed explanation of computation of Pixel N-gram features please refer to our previous work [5]. Here pixel N-grams in horizontal,vertical and diagonal direction are computed.

E. Classification

These features are normalized using min-max normalization technique. The normalized pixel N-gram features are then given as input to the classifiers for classification of the ROI into three classes namely circumscribed, speculation and normal.

IV. Experimental Results and Discussion

A. Finding Optimum Value of N

The value of N is expected to have a significant effect on the classification performance. It can be hypothesized that the classification performance would increase as N is increased up to certain level but most probably start decreasing with further increase in N. This is because with increase in value of N more complete representation of image is possible.

Further, with increase in N the sequences are hard to find therefore resulting in features more specific to a particular image. This will make it harder for the classifier to generalize. Moreover, the computational cost is increased with the increase in N. Thus a balance has to be achieved between increasing N for achieving the complete image representation (and therefore better classification performance) and decreasing N in order to reduce the dimensionality of the feature vector (improve classifier generalisation, avoid increase in computational cost). It is therefore necessary to obtain an optimum value of N.

For finding the optimum value of N, classification performance is analyzed with change in value of N. Therefore values of N=1,2,3,4,5 were considered. The N-gram features are computed and normalized using min-max normalization. The normalized features are fed to the MLP classifier.

The classification performance was compared using various parameters such as Fscore, sensitivity, specificity and Receiver Operating System curve area. The result of the experiment is noted in Table 1. The effect on N on classification is shown in graphical format in Figure 3

Table 1. Effect of N on Classification Performance

	1-gram	2-gram	3-gram	4-gram	5-gram
Fscore	72.5	80.2	86.4	85.2	79.6
Sensitivity	73.0	79.9	84.6	82.2	80.2
Specificity	88.2	88.0	88.9	88.3	88.5
ROC Area	80.3	83.4	85.8	84.1	81.6

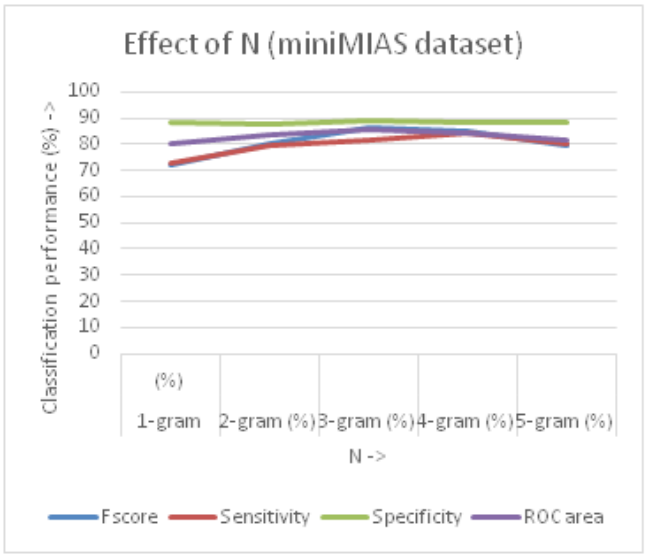

Figure 3 Classification Performance (N=1,2,3,4,5)

B. Comparison of Classifiers

The objective of this experiment was to compare different classifiers in order to find out which one works best for mammographic lesion classification. The best value of grey scale reduction (8 grey levels) and optimum value of N obtained from the experiment described above was used for reducing the ROIs in grey scale and computing N-gram features. The N-gram features were normalized using the min-max normalisation. The normalised N-gram features were fed as inputs to the classifier.

Various classifiers can be used for classification of images. Three most commonly used classifiers (MLP, SVM, KNN)

were used for classification of mammographic lesions into circumscribed, speculation and normal categories. Classification accuracy by using different classifiers is noted in Table 2.

Table 2. Effect of Classifiers

	Classification Accuracy(%) - miniMIAS			
Classifier	Cicumscribed	Speculation	Normal	Overall
KNN	60.0	56.0	70.0	70.0
SVM	42.0	61.0	82.0	71.0
MLP	72.7	68.3	90.4	82.0

It can be observed from the Figure 3 that the classification performance increases with increase in N until certain value of N (3) and then starts decreasing. As the value of N is increased, the Pixel N-gram representation of image becomes more and more complete. The complete representation of an image is obtained if the spatial relationship of every pixel with every other pixel is modeled while generating features. However, with increase in N the dimensionality of the feature vector is increased producing the risk of over fitting which is normally referred to as 'Curse of dimensionality problem' [19]. Due to this the classification performance could be degraded. Additionally, as the longer sequences are hardly observed, with increase in N the feature vector becomes too specific to a particular image making it hard for the classifier to generalize well. Moreover, the computational cost is increased with the increase in N.

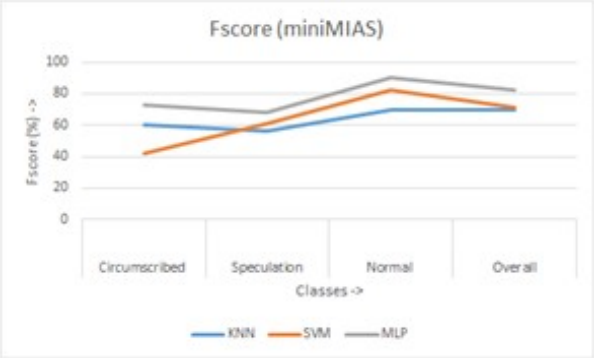

Figure 4. Classification Accuracy (MLP, SVM, KNN)

Figure 4 shows the graphical representation of classification accuracy using different classifiers. It can be seen that the highest classification accuracy is obtained using MLP classifier as compared to SVM and KNN classifiers for all the classes. Using SVM, the classification accuracy is lowest for circumscribed class whereas the KNN gives lowest classification accuracy for normal class.

V. Conclusion and Future Work

In this work, fine grained classification (circumscribed, speculation, normal) of mammographic lesions using Pixel N-gram features was attempted. The ROI images were grey scale reduced to 8 grey levels in order to reduce the computational complexity. Classification performance is dependent upon the value of N. Experiments were carried out to see the effect of N on the classification performance.

The classification performance (Fscore, Sensitivity, Specificity, Receiver Operating Characteristic (ROC) curve) was found to be highest when N=3. Therefore, the optimum value of N is considered 3 and is used for the further experiments.

Another experiment was carried out to compare the performance of different classifiers (MLP, SVM and KNN). It was observed that the MLP Classifier performed significantly better than SVM and KNN classifiers. The highest classification accuracy acquired with 3-gram features and MLP classifier was 82%. Additionally, Pixel N-grams are computationally less expensive and thus are good candidate for mammographic classification.

Future work involves improving classification performance by trying out various normalization techniques and using various mammographic datasets in order to analyse the use of Pixel N-grams efficacy for fine grained classification of mammograms.

References

[1] P. L. Porter, "Global trends in breast cancer incidence and mortality," Salud publica de Mexico, vol. 51, pp. s141-s146, 2009.
[2] S. Malvia, S. A. Bagadi, U. S. Dubey, and S. Saxena, "Epidemiology of breast cancer in Indian women," Asia Pacific Journal of Clinical Oncology, 2017.
[3] P. Náther, "N-gram based Text Categorization," Lomonosov Moscow State Univ, 2005.
[4] I. Kanaris, K. Kanaris, I. Houvardas, and E. Stamatatos, "Words versus character n-grams for anti-spam filtering," International Journal on Artificial Intelligence Tools, vol. 16, no. 06, pp. 1047-1067, 2007.
[5] P. Kulkarni, A. Stranieri, S. Kulkarni, J. Ugon, and M. Mittal, "Visual character n-grams for classification and retrieval of radiological images," The International Journal of Multimedia & Its Applications, vol. 6, no. 2, p. 35, 2014.
[6] P. Kulkarni, A. Stranieri, S. Kulkarni, J. Ugon, and M. Mittal, "Hybrid Technique Based On Ngram And Neural Networks For Classification Of Mammographic Images," in *Second I*nternational Conference on Signal, Image Processing and Pattern Recognition, 2014, pp. 297-306.
[7] K. Bovis and S. Singh, "Detection of masses in mammograms using texture features," in Proceedings of 15th International Conference on Pattern Recognition, 2000, vol. 2, pp. 267-270: IEEE.
[8] A. M. Khuzi, R. Besar, W. W. Zaki, and N. Ahmad, "Identification of masses in digital mammogram using gray level co-occurrence matrices," Biomedical imaging and intervention journal, vol. 5, no. 3, 2009.

[9] S. Beura, B. Majhi, and R. Dash, "Mammogram classification using two dimensional discrete wavelet transform and gray-level co-occurrence matrix for detection of breast cancer," *Neurocomputing,* vol. 154, pp. 1-14, 2015.
[10] S. V. da Rocha, G. B. Junior, A. C. Silva, A. C. de Paiva, and M. Gattass, "Texture analysis of masses malignant in mammograms images using a combined approach of diversity index and local binary patterns distribution," Expert Systems with Applications, vol. 66, pp. 7-19, 2016.
[11] M. Hussain, S. Khan, G. Muhammad, M. Berbar, and G. Bebis, "Mass detection in digital mammograms using gabor filter bank," 2012.
[12] T. Mu, A. K. Nandi, and R. M. Rangayyan, "Classification of breast masses using selected shape, edge-sharpness, and texture features with linear and kernel-based classifiers," Journal of Digital Imaging, vol. 21, no. 2, pp. 153-169, 2008.
[13] Y. Zhang, N. Tomuro, J. Furst, and D. S. Raicu, "Building an ensemble system for diagnosing masses in mammograms," International Journal of Computer Assisted Radiology and Surgery, vol. 7, no. 2, pp. 323-329, 2012.
[14] Y. Li, H. Chen, G. K. Rohde, C. Yao, and L. Cheng, "Texton analysis for mass classification in mammograms," Pattern Recognition Letters, vol. 52, pp. 87-93, 2015.
[15] P. Kulkami, A. Stranieri, and J. Ugon, "Texture image classification using pixel N-grams," in IEEE International Conference on Signal and Image Processing (ICSIP), 2016, pp. 137-141: IEEE.
[16] P. Kulkarni, A. Stranieri, and J. Ugon, Pixel N-grams: Size, Location and Resolution Invariance for Shape Classification International Journal of Science Engineering and Management 1(8), 38-44
[17] P. Kulkarni, Analysis and Comparison of Co-occurrence Matrix and Pixel N-gram Features for Mammographic Images. International Conference on Communication and Computing(2015), Banglore, India, 7-14
[18] J. Suckling *et al.*, "The mammographic image analysis society digital mammogram database," in Exerpta Medica. International Congress Series, 1994, vol. 1069, pp. 375-378.
[19] I. Bankman, Handbook of medical image processing and analysis. academic press, 2008.