2-SiMDoM: A 2-SIEVE MODEL FOR DETECTION OF MITOSIS IN MULTISPECTRAL BREAST CANCER IMAGERY

*Ardhendu Shekhar Tripathi, *Atin Mathur, *Mohit Daga, *†Manohar Kuse, *†Oscar C. Au

* The LNM Institute of Information Technology, Jaipur, India
*† The Hong Kong University of Science and Technology, Hong Kong
Email: *{ast.lnmiit, mathuratin007, mohitdaga.lnmiit}@gmail.com, *†{mkuse, eeau}@usk.hk

ABSTRACT

In this paper, we propose a 2-Sieve model for the detection of mitosis in breast cancer multispectral images. Multi-resolution wavelet features & Gray Level Entropy Matrix (GLEM) features have been computed for each candidate on all the spectral bands. A novel dimensionality selection algorithm has been introduced and its performance compared with other existing algorithms. Data imbalance and data cleaning have been taken care of using classical data mining techniques. Furthermore, a Second Sieve classification is performed to increase the Positive Predictive Value (PPV) with minimal loss in Sensitivity. A final Sensitivity and PPV of 82.35% & 73.04% respectively was achieved over the testing set using the proposed scheme.

Index Terms— Mitosis Detection, Active Contours, Wavelet Texture, Dimensionality Selection, Data Imbalance

1. INTRODUCTION

Breast cancer is the most frequent cancer found in women and is the second most leading cause of death of women worldwide. More than 8% of women suffer from this disease during their lifetime [1]. Diagnosis of breast cancer at an early stage is critical in improving breast cancer prognosis [2]. CAD (Computer Aided Diagnosis) has resulted in much more standardized grading practices which can be achieved by various diagnostic indicators. WHO has recommended the Nottingham grading scheme for grading of breast cancer biopsies. It is derived from the three morphological features: tubule formation, nuclear pleomorphism and mitotic count [3]. Apart from being tedious and time consuming, manual annotation of mitosis is a subjective task (due to similar looking lymphoid/inflammatory and apoptotic cells). This study focuses on the development of a novel, robust and more accurate scheme for mitotic cell count which is one of the parameters of the Nottingham grading scheme for breast cancer grading.

Earliest attempts in mitotic cell counting were done by Kate *et al.* [4] and Belien *et al.* [5]. Kate *et al.* have reported a classification rate varying from $19-42\%$. The study by Belien *et al.* have reported an accuracy of 81%. In the opinion of the authors of this paper, accuracy cannot be cosidered an apt performance measure as the classification of all the cells as non-mitotic alone will give an accuracy of above 90%. Khan *et al.* [6] have recently proposed a scheme for detection of mitosis cells. A sensitivity of 57% is reported on the test dataset. All the previous studies have been performed on standard histology images. This study attempts to use multi-spectral imagery for detection of mitosis in breast cancer histology.

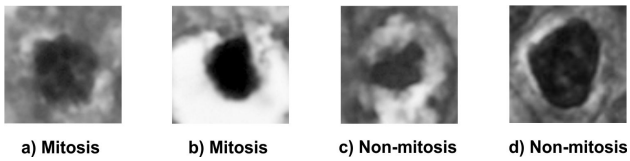

Fig. 1: It is difficult to manually differentiate between Mitosis and Non-Mitosis due to similar looking textures

2. DATA DESCRIPTION

We evaluated our system performance on the MITOS dataset[1]. A total of 48 High Power Fields (HPFs at 40X magnification), each of dimensions $512 \times 512 \mu m^2$ from 5 breast cancer biopsy slides stained by Hematoxylin & Eosin (H & E) were used. For each slide, there was a set of multispectral images captured over 10 visible bands BB00-BB09 (410nm-750nm). For each band, digitization was performed over 17 different focus planes (also referred to as spectral stacks). The multispectral imaging data was acquired by dividing the imaging area into 4 quadrants. Thus, there were in all, 170×4 single channel images for each slide. In the dataset, 33 HPF's were provided for training and the remaining 15 HPF's were provided for testing.

3. SYSTEM DESIGN

The different modules have been discussed in detail in the subsequent subsections. Figure 3 shows an overview of the

[1] http://ipal.cnrs.fr/ICPR2012/ , Mitosis Detection Contest website 2012

Mitosis Detection System.

3.1. Spectral Stack Selection

For each band, out of 17 spectral stacks (as described in section 2), a best quality stack was selected on the basis of maximization of the *image entropy measure*. The best quality stack of each band was used in all the future computations. According to [7], image entropy measure depends on the context probability which is used to track loss in quality. The image entropy measure identifies the best quality stack and is computed as follows:

$$E_{S_i} = \sum_{j=0}^{j=l-1} -p_j.log(p_j) \qquad \forall \text{ stack } S_i \tag{1}$$

where, p_j is the probablity of occurrence of j^{th} intensity. The stack S_k was selected such that, $S_k = \underset{i}{argmax}(E_{S_i})$ where S_k is the *best quality stack* for the k^{th} band ($k \in BB00 - BB09$).

3.2. Cell Segmentation

Accurate segmentation of cells was essential for correct classification of mitosis. The region based active contour model as proposed in [8] was used here for the purpose of cell segmentation.

The active contour model requires seed points as input to initialize the curve which further evolves to render a precise shape of the cell based on the minimization of an energy function. This technique was effective here due to the difference in the average pixel intensity levels inside and outside the cells.

Seed points for the initialization of the active contours were obtained by thresholding of higher contrast image $S_k \in$ BB07 band resulting in a binary mask M which contained the seed points. In order to enhance the contrast, the histogram equalization of this image was performed. Morphological operations were performed to ensure a smooth contour. The base image I for active contour model was obtained by convolution of $S_k \in$ BB06 band with a 3×3 gaussian mask. Thus, active contours were made more sensitive by incorporating Gaussian smoothing. By maintaining a smooth contour, this model becomes more robust towards noise [9].

3.3. Feature Extraction

Texture is an important aspect which provides significant information for image classification. The wavelet transform provides a unified framework for the multiresolution decomposition of images. It allows texture to be examined in a number of resolutions whilst maintaining spatial resolution [10].

The texture features were evaluated from a multi-level wavelet decomposition upto 3 levels from Daubechies wavelet family. These were selected because the analysis with daubechies is orthogonal, and both scaling and wavelet function are compactly supported [11].

5 Gray-Level Co-occurence Matrix (GLCM) features (cluster shade, dissimilarity, contrast, difference entropy and information measure of correlation) [12] were computed in the wavelet domain (4 components- LL, LH, HH, HL). Three level decomposition was chosen based on the minimum entropy decomposition algorithm [13]. This was done for all the 10 bands (BB00 - BB09) resulting into 600 features.

Apart from this, 9 Gray-Level Entropy Matrix (GLEM) features [14] were also computed on all the 10 bands. These features provide an estimate of the degree of homogeneity & other entropy measures.

Thus, there were a total of 690 features for each cell. These features were then linearly mapped in the range [0 1].

Algorithm 1 SDSNE Algorithm

$inputFeatureSpace \leftarrow [F_1F_2.....F_{totalFeatures}]$, where $F_1, F_2, ...F_{totalFeatures}$ are the individual features;
$reducedFeatureSpace \leftarrow \emptyset$;
$errors \leftarrow numberOfMitosis$;
for $i = 1 \rightarrow totalFeatures$ **do**
1. Concatenate F_i of $inputFeatureSpace$ to $reducedFeatureSpace$.
2. Calculate the 7 Nearest Neighbour for each mitosis instance (in the feature space) in the training set. If the majority of the Nearest neighbours are of non-mitosis class then declare it as $error$. Count such $errors$. It is stored in $newErrors$ for feature space defined by $reducedFeatureSpace$.
3.
if $newErrors \geq errors$ **then**, discard this feature and increment i.
else
Keep this feature in $reducedFeatureSpace$. Also, $errors$ = $newErrors$; Now, check backwards by removing features one by one and checking the $newErrors$ i.e, On removing a feature,
if $newErrors \geq errors$ **then** keep it.
else
Discard it from the $reducedFeatureSpace$ & $error$ = $newError$;
end if
end if
end for

3.4. Supervised Dimensionality Selection based on Neighborhood Examination

In this paper, we present a novel algorithm for dimensionality selection i.e. Supervised Dimensionality Selection based on Neighborhood Examination (SDSNE). This algorithm selects the set of features which minimizes the error rate i.e. the number of misclassified mitosis on the basis of the neighborhood

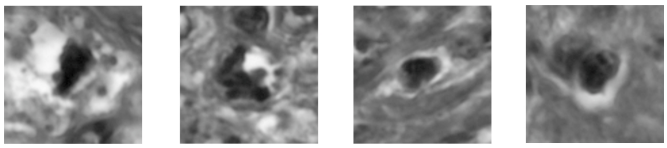

Fig. 2: 100x100 window around the bounding box of candidate mitosis (detected by Sieve 1). First 2 (from left) are mitotic and the last 2 are non-mitotic

majority rule in the feature space. For the pseudo code of the algorithm see Algorithm 1. It was able to reduce the dimensionality of the input feature space from 690 to 43. For the qualitative and quantitative evaluation of this algorithm refer to section 4.

3.5. Handling Imbalanced Dataset

There was a high degree of imbalance in the dataset, mitotic instances bieng very few in number as compared to the non-mitotic instances. Mitotic classification was just like finding a needle in a heap of hay stack.

In an imbalanced data classification, the class boundary learned by the standard machine learning algorithms is biased towards the majority class resulting in a high false negative rate [15]. It was of utmost importance to balance the class distribution in the training dataset before training a classifier. The imbalanced dataset was dealt by 1) Oversampling of mitotic instances and 2) Data Cleaning.

1) Oversampling of mitotic instances : Oversampling was done for creating additional training instances of mitotic cells (minority class). The segmented cells were rotated at various angles (45 deg, 90 deg etc.) to perturb the training data as proposed by [16]. The data was further oversampled by applying Synthetic Minority Oversampling Technique (SMOTE) [17] in the "reduced feature space". SMOTE provides more related minority class samples to learn from, thus allowing a classifier more coverage of the minority class due to broader decision regions.

2) Data Cleaning : Removal of class label noise as well as borderline examples (that have a higher probability of being classified incorrectly) is vital and must be done prior to the classification stage. Instances participating in tomek links [18] were eliminated from the dataset. Tomek links consist of points that are each others closest neighbors (in the reduced feature space), but do not share the same class label.

3.6. 2-Sieve Model for Classification of Mitotic Cells

3.6.1. First Level Sieve

Support Vector Machines (SVM) with a nonlinear radial basis kernel was trained on 33 HPF's and the model was used to grade the remaining 15 HPF's for mitotic occurences. A sensitivity and a Positive Predictive Value (PPV) of 85.29% and 59.58% were achieved after this basic classification step.

3.6.2. Second Level Sieve

At this level of classification, efforts were made to take advantage of the textural differences around a mitotic and a non-mitotic cell (see figure 4). This was done to reduce the number of False Positives (FP). The instances filtered at the first level as mitotic were considered for classification at this stage. A set of textural features were extracted from a window of 100×100 around the bounding box of each of the segmented cells. These were 48 Phase Gradient features (4 scales, 4 orientations and 3 bands) [19] and 48 Gabor features (4 frequencies, 4 orientations and 3 bands) [20]. The Gabor features were computed by convolving the 100×100 window with the respective Gabor filters and then finding the local energy at each pixel using the Gaussian filter. The 3 bands that were used correspond to the wavelengths 665-710nm (red band-BB00), 480-540nm (green band-BB05) and 430-490nm (blue band-BB03). These bands were used as these are the primary colors for human color vision. For each of these 96 features 4 statistical measures i.e. 1) mean, 2) skewness, 3) kurtosis and 4) standard deviation were calculated hence providing a 384 dimensional feature vector for each instance.

The training set for the second sieve was prepared by including all the mitotic instances present in the 33 HPF's which were used for training at the first level and randomly selecting an equal number of non-mitotic instances from the same. An ensemble of Random Projections [21] and SVM (linear kernel) with a majority rule was used to predict the mitotic instances from the candidate mitosis. Random Projections are computationally less expensive & using an ensemble of SVM classifiers addresses the issue of variability in the results of classifying low dimensional feature data generated by random projections.

A final sensitivity and PPV of 82.35% and 73.04% were obtained respectively after the second level filtering of the cells classified as mitosis in 3.6.1.

4. EXPERIMENTAL RESULTS

We evaluated the performance of our proposed algorithm on various measures: 1) Segmentation accuracy, 2) Quantitative and Qualitative evaluation of SDSNE, 3) Performance Comparison with the participants of the ICPR 2012 Mitosis Detection Contest and 4) The ability of our algorithm to handle extreme levels of imbalance in the dataset.

The segmentation accuracy was evaluated in terms of the localisation quality of the detected mitosis. The mean and the standard deviation of the distance between the detected mitosis and the ground truth centroids were found out to be

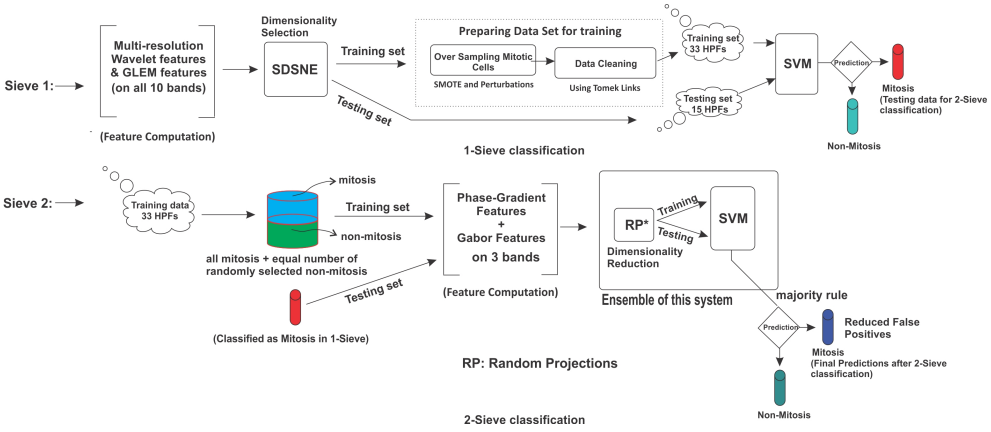

Fig. 3: 2-Sieve Model for Classification of Mitotic cells

0.87 pixels and 0.45 pixels respectively. Hence, justifying the use of active contours for cell segmentation.

Our dimensionality selection technique outperformed the existing dimensionality reduction algorithms in this scenario. The potency of our dimensionality selection scheme can be visualised using the representative Chernoff faces [22] for each class under consideration i.e. mitotic and non-mitotic before and after applying SDSNE (see figure 2). A comparative study of the performance of various dimensionality reduction and selection algorithms has been shown in the table 1.

Dimensionality Reduction Technique	PPV	Sensitivity
Multilayer Encoders [23]	65.37%	75.24%
Diffusion Maps [23]	31.92%	67.81%
Maximum Variance Unfolding [23]	44.07%	64.27%
Local Linear Embedding [23]	58.21%	61.20%
Sequential Forward Floating Selection [24]	62.79%	70.35%
Sequential Backward Floating Selection [24]	52.98%	71.26%
SDSNE	**73.04%**	**82.35%**

Table 1: A comparative study: Dimensionality Reduction Algorithms

A 22.59% increase in the overall PPV was achieved at the cost of 3.45% decrease in the sensitivity after applying the second level sieve to the filtered mitotic candidates that were obtained after the first level sieve.

The only studies the authors of this paper are aware in the direction of mitosis detection in multi-spectral images are the participants of the ICPR-2012 Mitosis Detection contest [2]. 2-SiMDoM outclasses the performance of all the ICPR 2012 Mitosis Detection Contest participants on the same dataset (see table 2).

An experiment was carried out to test the confidence limit of our proposed scheme. Testing data was prepared such that it contained the HPF's having less than or equal to two mitotic instances. An overall sensitivity and PPV of 81.13% and 74.97% was obtained respectively. The result obtained validates the fact that the proposed scheme performs well even in the case when the minority class occurences are extremely few in number.

[2]http://ipal.cnrs.fr/ICPR2012/?q=node/12 , ICPR 2012 Mitosis Detection Contest Results

ICPR Mitosis Detection Contest Participants	PPV	Sensitivity
NEC	73.85%	48.98%
ALBERTA	41.23%	72.45%
LNMIIT	37.69%	50.00%
OKAN-IRISA-LIAMA	37.93%	33.67%
2-SiMDoM	**73.04%**	**82.35%**

Table 2: Performance of the ICPR 2012 Mitosis Detection Contest Participants on the Multispectral Images

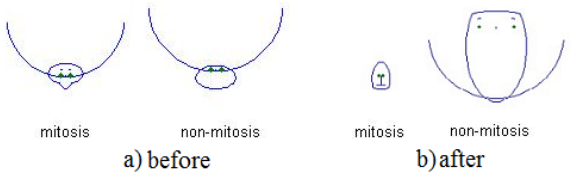

Fig. 4: Chernoff faces of Mitosis and Non-mitosis a) before applying SDSNE, and b) after applying SDSNE

5. CONCLUSION AND FUTURE WORK

A 2-Sieve model for the detection of mitotic cells in multispectral breast cancer imagery has been proposed in this study. We evaluated the performance of the proposed detection algorithm in terms of sensitivity and PPV over a set of 48 HPFs. It showed a remarkable increase in the sensitivity and the PPV over previous studies.

With collaborative studies and association with medical community, the other two parameters of Nottingham Grading Scheme i.e. tubule formation and nuclear pleomorphism can also be automated for effective and fast diagnosis of breast cancer.

6. REFERENCES

[1] H. Zhi, B. Ou, B.M. Luo, X. Feng, Y.L. Wen, and H.Y. Yang, "Comparison of ultrasound elastography, mammography, and sonography in the diagnosis of solid breast lesions," *Journal of ultrasound in medicine*, vol. 26, no. 6, pp. 807–815, 2007.

[2] H.D. Cheng, X. Cai, X. Chen, L. Hu, and X. Lou, "Computer-aided detection and classification of microcalcifications in mammograms: a survey," *Pattern recognition*, vol. 36, no. 12, pp. 2967–2991, 2003.

[3] H. Pereira, SE Pinder, DM Sibbering, MH Galea, CW Elston, RW Blamey, JFR Robertson, and IO Ellis, "Pathological prognostic factors in breast cancer. iv: Should you be a typer or a grader? a comparative study of two histological prognostic features in operable breast carcinoma," *Histopathology*, vol. 27, no. 3, pp. 219–226, 1995.

[4] TK Ten Kate, JAM Belien, AWM Smeulders, and JPA Baak, "Method for counting mitoses by image processing in feulgen stained breast cancer sections," *Cytometry*, vol. 14, no. 3, pp. 241–250, 1993.

[5] JAM Belien, JPA Baak, PJ Van Diest, and AHM Van Ginkel, "Counting mitoses by image processing in feulgen stained breast cancer sections: the influence of resolution," *Cytometry*, vol. 28, no. 2, pp. 135–140, 1998.

[6] A Khan, H El-Daly, and NM Rajpoot, "A gamma-gaussian mixture model for detection on mitotic cells in breast histology images.," *International Conference on Pattern Recognition (ICPR)*, 2012.

[7] I. Avcibas, M. Nasir, and B. Sankur, "Steganalysis based on image quality metrics," in *Multimedia Signal Processing, 2001 IEEE Fourth Workshop on*. IEEE, 2001, pp. 517–522.

[8] T.F. Chan and L.A. Vese, "Active contours without edges," *Image Processing, IEEE Transactions on*, vol. 10, no. 2, pp. 266–277, 2001.

[9] D.L. Pham, C. Xu, and J.L. Prince, "Current methods in medical image segmentation 1," *Annual review of biomedical engineering*, vol. 2, no. 1, pp. 315–337, 2000.

[10] S. Mallat, *A wavelet tour of signal processing*, Academic press, 1999.

[11] I. Daubechies et al., *Ten lectures on wavelets*, vol. 61, SIAM, 1992.

[12] R. Nithya and B. Santhi, "Mammogram classification using maximum difference feature selection method," *Journal of Theoretical and Applied Information Technology*, vol. 33, pp. 197–204, 2011.

[13] R.R. Coifman and M.V. Wickerhauser, "Entropy-based algorithms for best basis selection," *Information Theory, IEEE Transactions on*, vol. 38, no. 2, pp. 713–718, 1992.

[14] K. Yogesan, T. Jørgensen, F. Albregtsen, KJ Tveter, and HE Danielsen, "Entropy-based texture analysis of chromatin structure in advanced prostate cancer," *Cytometry*, vol. 24, no. 3, pp. 268–276, 1998.

[15] S. Kotsiantis, D. Kanellopoulos, and P. Pintelas, "Handling imbalanced datasets: A review," *GESTS International Transactions on Computer Science and Engineering*, vol. 30, no. 1, pp. 25–36, 2006.

[16] T.M. Ha and H. Bunke, "Off-line, handwritten numeral recognition by perturbation method," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 19, no. 5, pp. 535–539, 1997.

[17] N.V. Chawla, K.W. Bowyer, L.O. Hall, and W.P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," *Arxiv preprint arXiv:1106.1813*, 2011.

[18] I. Tomek, "Two modifications of cnn," *IEEE Trans. Syst. Man Cybern.*, vol. 6, pp. 769–772, 1976.

[19] K. Murtaza, S. Khan, and N. Rajpoot, "Villagefinder: Segmentation of nucleated villages in satellite imagery," *British Mission Vision Corporation*, 2009.

[20] I. Fogel and D. Sagi, "Gabor filters as texture discriminator," *Biological cybernetics*, vol. 61, no. 2, pp. 103–113, 1989.

[21] E. Bingham and H. Mannila, "Random projection in dimensionality reduction: applications to image and text data," in *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 2001, pp. 245–250.

[22] H. Chernoff, "The use of faces to represent points in k-dimensional space graphically," *Journal of the American Statistical Association*, vol. 68, no. 342, pp. 361–368, 1973.

[23] LJP Van der Maaten, EO Postma, and HJ Van Den Herik, "Dimensionality reduction: A comparative review," *Journal of Machine Learning Research*, vol. 10, pp. 1–41, 2009.

[24] P. Pudil, J. Novovičová, and J. Kittler, "Floating search methods in feature selection," *Pattern recognition letters*, vol. 15, no. 11, pp. 1119–1125, 1994.