J Pathol Inform

Editor-in-Chief:
Anil V. Parwani , Pittsburgh, PA, USA
Liron Pantanowitz, Pittsburgh, PA, USA

OPEN ACCESS
HTML format

For entire Editorial Board visit : www.jpathinformatics.org/editorialboard.asp

Symposium - Original Research

HyMaP: A hybrid magnitude-phase approach to unsupervised segmentation of tumor areas in breast cancer histology images

Adnan M. Khan[1], Hesham El-Daly[2], Emma Simmons[3], Nasir M. Rajpoot[1,4]

[1]Department of Computer Science, University of Warwick, [2]Addenbrookes Hospital, Cambridge, [3]University Hospital Coventry and Warwickshire, UK, [4]Department of Computer Science & Engineering, Qatar University, Qatar

E-mail: *Nasir M. Rajpoot - nasir.rajpoot@ieee.org
*Corresponding author

Received: 21 January 13 Accepted: 21 January 13 Published: 30 March 13

This article may be cited as:
Khan AM, El-Daly H, Simmons E, Rajpoot NM. HyMaP: A hybrid magnitude-phase approach to unsupervised segmentation of tumor areas in breast cancer histology images. J Pathol Inform 2013;4:S1.

Available FREE in open access from: http://www.jpathinformatics.org/text.asp?2013/4/2/1/109802

Abstract

Background: Segmentation of areas containing tumor cells in standard H&E histopathology images of breast (and several other tissues) is a key task for computer-assisted assessment and grading of histopathology slides. Good segmentation of tumor regions is also vital for automated scoring of immunohistochemical stained slides to restrict the scoring or analysis to areas containing tumor cells only and avoid potentially misleading results from analysis of stromal regions. Furthermore, detection of mitotic cells is critical for calculating key measures such as mitotic index; a key criteria for grading several types of cancers including breast cancer. We show that tumor segmentation can allow detection and quantification of mitotic cells from the standard H&E slides with a high degree of accuracy without need for special stains, in turn making the whole process more cost-effective. **Method:** Based on the tissue morphology, breast histology image contents can be divided into four regions: Tumor, Hypocellular Stroma (HypoCS), Hypercellular Stroma (HyperCS), and tissue fat (Background). Background is removed during the preprocessing stage on the basis of color thresholding, while HypoCS and HyperCS regions are segmented by calculating features using magnitude and phase spectra in the frequency domain, respectively, and performing unsupervised segmentation on these features. **Results:** All images in the database were hand segmented by two expert pathologists. The algorithms considered here are evaluated on three pixel-wise accuracy measures: precision, recall, and F1-Score. The segmentation results obtained by combining HypoCS and HyperCS yield high F1-Score of 0.86 and 0.89 with respect to the ground truth. **Conclusions:** In this paper, we show that segmentation of breast histopathology image into hypocellular stroma and hypercellular stroma can be achieved using magnitude and phase spectra in the frequency domain. The segmentation leads to demarcation of tumor margins leading to improved accuracy of mitotic cell detection.

Key words: Breast cancer grading, histopathology image analysis, magnitude and phase spectra of Gabor filters, mitotic cell detection, texture feature, tumor segmentation

Access this article online

Website: www.jpathinformatics.org

DOI: 10.4103/2153-3539.109802

Quick Response Code:

INTRODUCTION

Grading of breast cancer relies largely on the microscopic examination of tissue slides stained with Hematoxylin and Eosin (H & E). This is a subjective process by its very nature, consequently leading to inter- and even intraobserver variability, potentially affecting the predicted patient prognosis and also the treatment modalities offered. The variability in breast cancer grading may, at least, in part, be responsible for the variability in rates of chemotherapy use between institutions.

Segmentation of areas containing tumor cells in standard H & E histopathology images of the breast (and several other tissues) is a key task for computer-assisted assessment and grading of histopathology slides. Good segmentation of the tumor regions can not only highlight the slide areas consisting of tumor cells, but it is also vital for the automated scoring of immunohistochemical (IHC)-stained slides, to restrict the scoring or analysis to areas containing only tumor cells and avoid potentially misleading results from the analysis of stromal regions. Furthermore, detection of mitotic cells is critical for calculating key measures such as the mitotic index; a key criteria for grading several types of cancers, including breast cancer. We are aware of the existing technologies that are capable of detecting mitotic cells on slides stained with IHC stains (e.g., Ki67, PHH3, etc.).[1] However, as we show in this article, tumor segmentation can allow detection and quantification of mitotic cells from the standard H & E slides with a high degree of accuracy, without the need for special stains, in turn making the whole process more cost-effective.

Although some algorithms for segmentation of tumor nuclei, quantitative evaluation of nuclear pleomorphism, detection and grading of lymphocytic infiltration in histology images, and automated malignancy detection have been reported in literature, tumor segmentation in breast histology images has not received much attention.[2-5] Wang, *et al*., has proposed a supervised tumor segmentation approach for tissue microarray spots that exploits tissue architectural and textural features in the Markov Random Field (MRF) based Bayesian estimation framework.[6] However, supervised segmentation of breast cancer histology images containing a highly complex texture often raises questions with regard to an algorithm's ability to avoid overfitting, let alone the issue of training overhead.

Feature-based segmentation approaches often use a filter bank to represent a pixel as a point in a high-dimensional feature space, posing the so-called curse of the dimensionality problem. A dimensionality reduction (DR) technique, giving a low-dimensional representation and preserving relative distances between features from the original feature space, is desirable to solve this problem. Along these lines, Viswanath *et al*., proposed an ensemble embedding framework and applied it to image segmentation and classification.[7] The idea is to generate an ensemble of low dimensional embeddings (using a variety of DR methods, such as graph embedding), evaluate embedding strength to select the most suitable embeddings, and finally generate consensus embedding by exploiting the variance among the ensemble. However, a major limitation of this framework, in the context of histopathology image analysis, is its high storage and computational complexity, mainly due to the very high-dimensional affinity matrix required for graph embeddings.

Random Projections (RPs) have recently emerged as a computationally simple and efficient low-dimensional subspace representation, with a minor drawback: Multiple RPs may produce substantially different projections because of the very nature of the random matrices.[8] Although this may not be a big issue in certain applications (like multimedia compression etc.), it cannot be ignored in applications like segmentation in low-dimensional feature space. Khan *et al*., proposed an ensemble of multiple RPs (which they termed RanPEC, short for Random Projections with Ensemble Clustering) followed by majority voting, to address the issue of variability among multiple RPs.[9] They further showed that ensemble clustering of random projections onto merely five dimensions achieves higher segmentation accuracy than a well-known supervised DR method on breast histology images.[10]

In this article, we propose a fast and totally unsupervised tumor segmentation framework based on dividing the stromal tissue into two types: Hypocellular Stroma (HypoCS) and Hypercellular Stroma (HyperCS). The proposed algorithm employs the magnitude spectrum in the Gabor frequency domain to segment the HypoCS regions and the phase spectrum in the Gabor frequency domain to segment the HyperCS regions. The algorithm has been evaluated on 35 H & E stained breast histology images belonging to five different tissue slides. Instead of evaluating the system using object-based criteria, we have incorporated a much stricter pixel-based quantitative evaluation criterion.

The experimental results show that the proposed system achieves an F1-Score of 0.89 (with respect to the GT markings) for pixel-based segmentation, in H & E images. The main contributions of this article are as follows: (a) We show that the magnitude and phase spectra of the frequency domain are effective in representing the complementary features of the HypoCS regions and HyperCS regions, respectively; (b) we present a fast, unsupervised, and data-independent algorithm for pixel level classification of tumor versus stromal regions (by integrating the HypoCS and HyperCS segmentations) in breast histology images, and (c) we show that segmentation of the stromal regions in breast histology images plays a critical role in mitosis detection, leading to a more accurate calculation of the

mitotic index: One of the three criteria used in the so-called Nottingham breast cancer grading system.[11]

The remainder of this article is organized as follows. Section 2 outlines details of the segmentation algorithm, in particular how the segmentation of HypoCS and HyperCS is performed in a low dimensional feature space. Comparative results and discussion are presented in Section 3. The article concludes with a summary of our results and some directions for future studies.

The Proposed Algorithm

Description of the Dataset

We evaluated our segmentation framework on the MITOS dataset. The dataset consists of 35 HPF (High Power Field) images taken from five different breast cancer biopsy slides, stained with H & E, and scanned at ×40 magnification, using an Aperio ScanScope slide scanner. Each HPF has a digital resolution of 2084 × 2084 pixels.

The Segmentation Algorithm

On the basis of the tissue morphology, the breast histology image contents can be divided into four regions [Figure 1]: Tumor, Hypocellular Stroma (HypoCS), Hypercellular Stroma (HyperCS), and tissue fat and/or retractions/artifacts (Background). The background is removed during the pre-processing stage, on the basis of color thresholding, while the HypoCS and HyperCS regions are segmented by calculating features using the magnitude and phase spectra, respectively, in the frequency domain and performing a RanPEC segmentation on these features.[9] The algorithm pipeline can be subdivided into three stages: (1) Pre-processing, to normalize the staining artifacts and remove the tissue fat, artifacts, and background; (2) Segmentation of HypoCS and HyperCS regions; (3) Post-processing to combine the result of background removal in (1) and segmentation in (2). A block diagram of the proposed tumor segmentation framework is shown in Figure 2. Algorithm 1 outlines the algorithmic details of the pipeline.

Pre-processing

Stain color constancy is one of biggest challenges of the digitized images of H & E–stained tissue slides. Several factors such as thickness of the tissue section, dye concentration, stain timings, and stain reactivity, result in variable stain color intensity and contrast. We have evaluated various stain normalization methods, but have found the Magee *et al*., method to be the most effective in terms of dealing with tissues containing large amount of retractions/staining artifacts.[12] The second stage of the pre-processing pipeline is to estimate the background. First the stain-normalized (color) tissue image is transformed from the RGB space into the YCbCr space. The luminance channel is then thresholded using an empirically determined, fixed, global threshold. The rough binary mask resulting from this thresholding is finally refined via morphological operations, in order to fill up the small gaps. Finally, the stain-normalized and background-free image is converted into the CIE's La*b* color space and anisotropic diffusion is applied to its b* channel, to remove the inherent camera noise while preserving the edges.[13]

Hypocellular Stromal Features

A traditional approach to texture segmentation is inspired by the multi-channel filtering theory. The idea is to characterize an image by a bank of filters, to generate a set of features that are capable of discriminating texture patterns belonging to different categories. A two-dimensional Gabor function consists of a sinusoidal plane wave of a certain frequency and orientation, modulated by a two-dimensional Gaussian. A Gabor filter in the spatial domain is given by the following equation:[14]

Algorithm 1: Hybrid magnitude-phase (HyMaP) - based tumor segmentation

1: Input: I ← RGB breast histology image

2: Output: T, a binary image where pixels belonging to the tumor regions have a value of 1 and all other pixels have a value of 0

3: Pre-processing

$I_{norm} = StainNormalize(I)$

$B = EstimateBackground(I_{norm})$

$\tilde{I}^{\beta}_{norm} = AnisotropicDiffusion(I^{\beta}_{norm})$

where $\tilde{I}^{\beta}_{norm}$ is the b*channel from the La*b*color space[12] of I_{norm}.

4: HypoCS Segmentation

$G_{gabor} = \{GaborFilter(\tilde{I}^{\beta}_{norm}, \theta, f) \ | \theta \in \{0,\ldots,\pi\} \ and \ f \in F\}$

where F is the set of frequencies as defined at the end of Section 2.2.

$G_{ener} = TextureEnergy(|G_{gabor}|, \mu, \sigma)$

where μ and σ are parameters of a Gaussian window used to compute the texture energy.

$H_0 = RanPECSegmentation(G_{ener})$, as described in Section 2.2.

5: HyperCS Segmentation

$G_{gabor} = \{GaborFilter(\tilde{I}^{\beta}_{norm}, \theta, f) \ | \theta \in \{0,\ldots,\pi\} \ and \ f \in F\}$

where F is as defined in Section 2.2.

$G_{pg} = GradientFeature(G_{gabor}, N)$, according to equation (7), where an $N \times N$ ($N=15$) window is used to compute the local phase gradients.

$H_r = RanPECSegmentation(G_{pg})$, as described in Section 2.2.

6: Post-processing

$T = \overline{H_o} \ \& \ \overline{H_r} \ \& \ \overline{B}$

where $\overline{H_o}, \overline{H_r}$, and $\overline{B}$ are morphologically post-processed versions of H_0, H_r and B.

7: Return T.

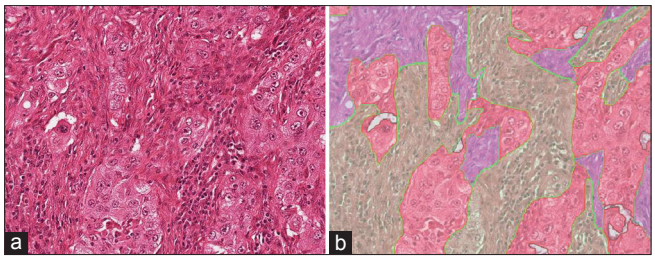

Figure 1: A sample H & E–stained breast cancer histology image: (a) Original image, and (b) Overlaid image, with four types of contents shown in different colors. The tumor areas are shown in Red, HypoCS in Purple, and HyperCS in Green. Areas containing background or fat tissue are shown in white with black outline. Note the difference in morphology of the Hypo- and Hypercellular stromal regions

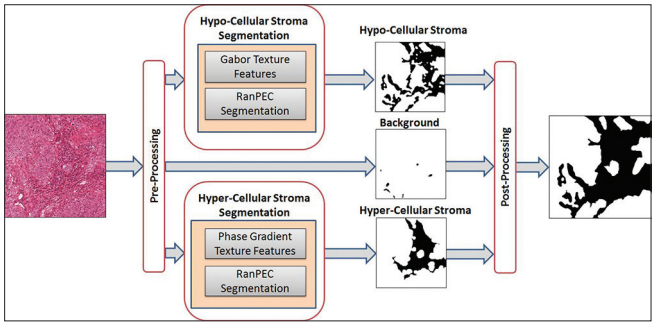

Figure 2: Overview of the proposed algorithm: HyMaP

$$G_{\theta,f}(x,y) = g_{\sigma}(x,y)\exp(j2\pi f(x\cos\theta + y\sin\theta)) \quad (1)$$

where $g_{\sigma}(x,y)$ is a Gaussian kernel with a bandwidth of σ. The parameters f and θ represent frequency and orientation of the 2D Gabor filter, where θ varies between 0 and π in regular intervals, $f \in F$, and F denotes a set of possible frequencies, and is defined as follows.

$$F_L(i) = 0.25 - 2^{i-0.5}/N_C \quad | \quad 0 < F_L(i) < 0.25 \quad (2)$$

$$F_H(i) = 0.25 + 2^{i-0.5}/N_C \quad | \quad 0.25 < F_H(i) < 0.5 \quad (3)$$

Where $i = 0,1,\ldots,\log_2(N_{col}/8)$, N_{col} is the width of the image in terms of the nearest power of 2. We then define the set F of possible frequencies as follows,

$$F = F_L \cup F_H \quad (4)$$

For an image with 512 columns, for example, a total of 84 Gabor filters can be used (six orientations and 14 frequencies). The hypocellular stromal features are then computed by convolving the Gabor filters $G_{\theta,f}(\cdot)$ with $\tilde{I}^{\beta}_{norm}$ (obtained from step 3 of Algorithm 1), and computing local energy on the results of the convolution.

Hypercellular Stromal Features

Phase information could be used as an important cue in modeling the textural properties of a region. Murtaza *et al.*, used local frequency estimates in the Gabor domain over a range of scales and orientations to yield a signature, which was shown to efficiently characterize the texture of a village in satellite images.[16] We chose the phase spectrum to represent the attributes of the HyperCS regions in a breast histology image, due to the recently established efficacy of the phase in textures exhibiting randomness.

Let $v_i(x,y)$ denote the i^{th} Gabor channel for the stain normalized and smoothened version of an input image $I(x,y)$, where $i = ,1,2,\ldots,N_g$, $N_g = N_{\theta} \times |F|$ and N_{θ} denotes the number of orientations. We can represent it as follows,

$$v_i(x,y) = |v_i(x,y)|\exp(j\phi_i(x,y)) \quad (5)$$

where $|\cdot|$ denotes the magnitude operator and $\phi_i(x,y)$ denotes the local phase. The gradient of the local phase and its magnitude can then be computed as below,

$$\phi_i'(x,y) = \left[\frac{|v_i(x,y)|'}{|v_i(x,y)|} - \frac{v_i'(x,y)}{v_i(x,y)}\right] \quad (6)$$

and

$$|\phi_i'(x,y)| = \sqrt{\frac{d\phi_i^2}{dx} + \frac{d\phi_i^2}{dy}} \quad (7)$$

The phase gradient features are computed using (7) for each of the Gabor filter responses, over a window of size $N \times N$(where N=15).

RanPEC Segmentation

RanPEC is a fast, unsupervised, and data-independent framework for dimensionality reduction and clustering of high-dimensional data points.[9] The main idea of RanPEC is to project high-dimensional feature vectors onto a relatively small number of orthogonal random vectors belonging to a unit ball and perform ensemble clustering in the reduced-dimensional feature space. By getting an ensemble of projections for each feature vector and then picking a cluster for a pixel by the majority voting selection criterion, ensures stability of the results among different runs. Experimental results in[9] suggest that promising classification accuracy can be achieved by random projections when using fast matrix operations in an unsupervised manner.

RESULTS

All 35 images in the database were hand segmented by two expert pathologists. We generate all experimental results on three criteria: (1) Considering the first pathologist's markings (P-1) as the ground truth (GT); (2) considering the second pathologist's markings (P-2) as GT; (3) fusing P-1 and P-2 using the logical OR rule (i.e., a pixel is considered to be tumorous if any one of the two pathologists marked the pixel as tumorous), and considering the fused image as GT. Some of the HPF images contain large tumor regions with small islands of stroma here and there; however, a majority of HPF images contain a fair share of hypo- and hypercellular stroma (approximately 33%, on an average). The average

degree of disagreement between the two pathologists on GT images is 11.55 ± 5.37%.

All 35 images in the dataset are pre-processed in a similar manner, with stain normalization carried out as described in Section 2.2, and the background removal performed to remove fat and other artifacts caused by staining and fixation. This provides robustness in the subsequent steps of the pipeline. In order to segment HypoCS, a total of 84 Gabor textural features (14 scales and six orientations) are calculated for each pixel of the input image *I*. In order to generate HyperCS, the PG features are calculated on 10 orientations and three scales. The gradient features are computed in a window of size 15 × 15. The RanPEC segmentation framework is applied on both Gabor and PG features, independently, yielding HypoCS and HyperCS segmentations, respectively. The RanPEC[9] framework requires two parameters: r the dimensionality of lower dimensional space, and n_c the number of runs of the ensemble. As recommended in,[9] we used $r = 5$ and $n_c = 20$ in our experiments. We have compared our proposed algorithm (HyMaP) with RanPEC using the same experimental setup as suggested in.[9]

The algorithms considered here are evaluated on three pixel-wise accuracy measures: Precision, recall, and F1-Score. The F1-Score is a measure that combines precision and recall in a statistically more meaningful way. Let TP denote the number of true positives, FP the number of false positives, TN the number of true negatives, and FN the number of false negatives; the precision is defined as TP/(TP + FP), recall is defined as TP/(TP + FN), and the F1-Score is defined as 2 (precision recall)/(precision + recall).

Figure 3 Provides an illustration of the efficiency of HypoCS segmentation [Figure 3b] and HyperCS segmentation [Figure 3c] in capturing the complementary stromal subtypes. Figure 4 provides an illustration of the proposed tumor segmentation algorithm on two different HPF images. The

Table 1: Quantitative results of tumor segmentation accuracy indicators (precision, recall, and F1-Score) for 35 BC histopathology images using three feature spaces (1) Unreduced ($n = 114$), (2) RanPEC ($n = 10$, as in [9]), and (3) HyMaP ($n = 10$), where n denotes the dimensionality of the feature space

Ground truth	Algorithm	Precision	Recall	F1-score
Pathologist-1	Unreduced	0.89±0.05	0.86±0.07	0.87
	RanPEC	0.85±0.06	0.86±0.05	0.85
	HyMaP	0.88±0.03	0.88±0.05	0.88
Pathologist-2	Unreduced	0.90±0.08	0.86±0.09	0.88
	RanPEC	0.86±0.07	0.85±0.07	0.85
	HyMaP	0.9±0.06	0.88±0.07	0.89
Fused	Unreduced	0.93±0.06	0.84±0.08	0.88
	RanPEC	0.88±0.07	0.83±0.06	0.85
	HyMaP	0.93±0.04	0.86±0.06	0.89

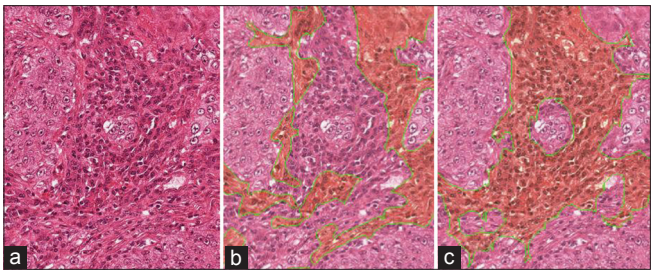

Figure 3: Illustration of complementary segmentations obtained by hypo- and hypercellular stroma segmentation: (a) Original images; (b) Results of HypoCS shown in slightly darker contrast, outlined in green color; (c) Results of HyperCS shown in slightly darker contrast, outlined in green color

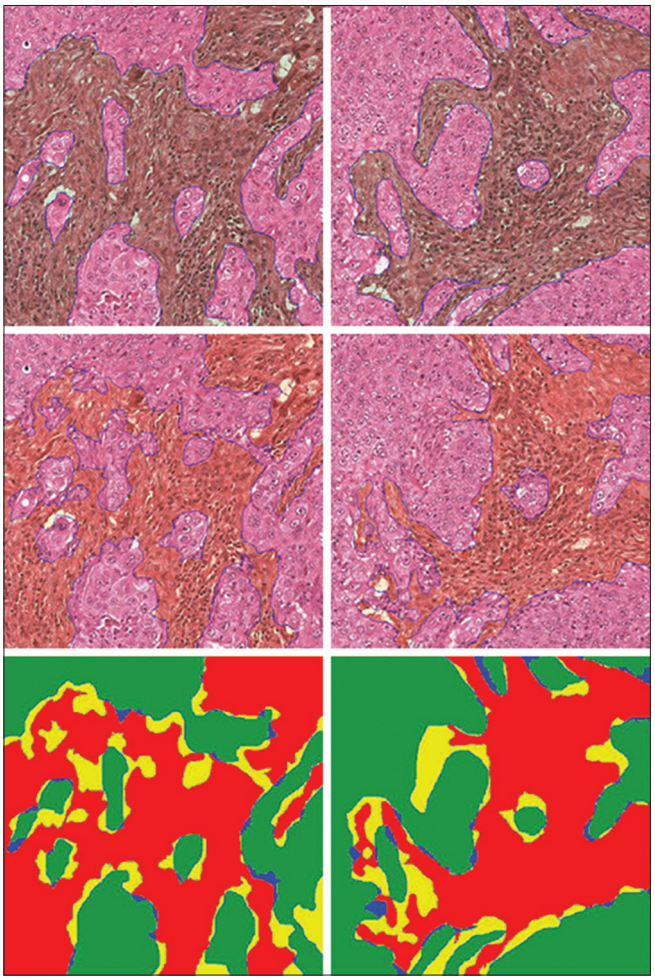

Figure 4: Visual results of tumor segmentation in two sample images: First row: Original images with fused ground truth (GT), marked non-tumor areas shown in a slightly darker contrast with blue boundaries; Second row: Results of combining HypoCS and HyperCS using the proposed framework (F1-Score = 0.86 and 0.89, respectively);Third row:Visual illustration of segmentation accuracy, green = TP, red = TN, yellow = FN, and blue = FP

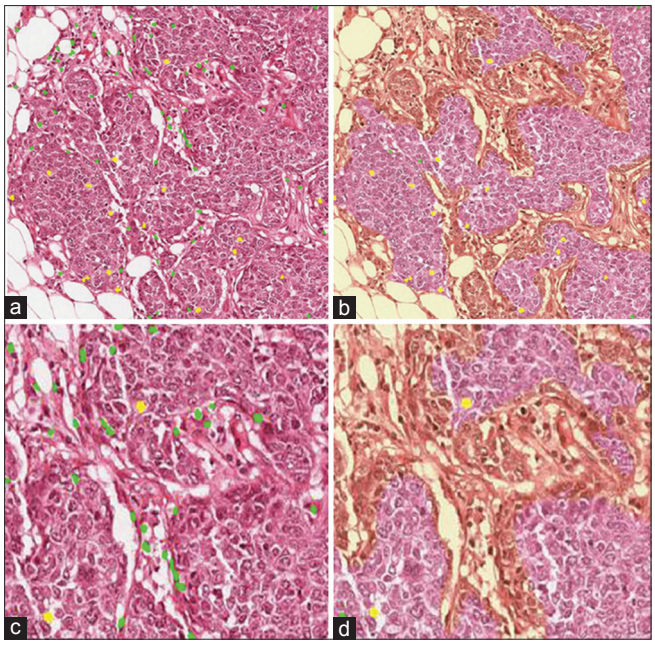

Figure 5: Visual results of Mitotic cell (MC) detection in a sample image: (a) Results of MC detection without tumor segmentation (TP = 17, FP = 82) using.[17] All the true positive MCs are shown in yellow color, while all the false positives are shown in green color. (b) Results of MC detection with tumor segmentation (TP = 17, FP = 4). All the true positive MCs are shown in yellow color, while all the false positives are shown in green color. (c) Zoomed-in version of a portion of (a) for better visibility. (d) Zoomed-in version of a portion of (b) for better visibility

segmentation results obtained by combining HypoCS and HyperCS yield high F1-Scores of 0.86 and 0.89, with respect to the fused GT. Considering the degree of disagreement between the two pathologists (i.e., 11.5 ± 5.37%), the results can be termed as highly accurate. Table 1 shows the segmentation accuracies (in terms of precision, recall, and F1- Score) of the unreduced and reduced feature spaces resulting from automated tumor segmentation. Note that the F1-Scores obtained from HyMaP (0.88, 0.89, and 0.89) are higher when compared with those from the unreduced feature space (0.87, 0.88, and 0.88) and RanPEC (0.85, 0.85, and 0.85). Table 1 also reveals that the reduced textural feature space achieves F1-Scores of 0.88, 0.89, and 0.89, suggesting, in turn, that DR removes the redundant features and preserves the distances between high dimensional feature spaces, thereby improving segmentation accuracy.

Further to the accuracy of segmentation, we present an application of tumor segmentation to Mitotic Cells (MCs) detection in tumor areas. MC detection is critical for calculating key measures such as the mitotic index: One of the three criteria used in the Nottingham grading system to grade breast cancer histology slides.[15] Khan *et al*., proposed a statistical approach that models the pixel intensities in mitotic and non-mitotic regions by a mixture of the Gamma-Gaussian mixture model.[17] Figure 5 visually illustrates how tumor segmentation can improve mitotic cell detection accuracy in breast histology images. Using the algorithm reported in,[17] Figure 5a shows the results of MC detection without tumor segmentation and Figure 5b shows the results of MC detection with tumor segmentation. Figure 5c and Figure 5d show the zoomed-in version of Figure 5a and Figure 5b. Note that the number of false positives increase significantly (from 4 to 82) when tumor segmentation is not performed for mitotic cells detection.

CONCLUSION

In this article, we presented an algorithm for segmentation of tumor areas in breast histology images based on segmentation of the image into hypocellular stroma and hypercellular stroma regions, using the magnitude and phase spectra in the Gabor domain. The complementary nature of the segmentation of two stromal subtypes was shown, resulting in high segmentation accuracy for the tumor areas. It was further demonstrated that the specificity of mitotic cell detection can be significantly enhanced when detection is restricted to the tumor areas. We anticipate further applications of our method to accurate, tumor-localized quantification/scoring of IHC stained slides and its validation on large-scale datasets.

ACKNOWLEDGMENTS

The authors would like to thank the organizers of the ICPR 2012 contest for mitosis detection in breast cancer. The images used in this article are part of the MITOS dataset, a dataset setup for the ANR French project MICO. The authors would also like to thank Dr. Derek Magee for sharing the executable for his algorithm for stain normalization. We would also like to thank Dr. Asha Rupani for her help in marking the tumor ground truth on the breast histopathology images. The first author gratefully acknowledges the financial support provided by the Warwick Postgraduate Research Scholarship scheme and the Department of Computer Science at the University of Warwick.

REFERENCES

1. Roullier V, Lézoray O, Ta VT, Elmoataz A. Multi-resolution graph-based analysis of histopathological whole slide images: Application to mitotic cell extraction and visualization. Comput Med Imaging Graph 2011;35:603-15.
2. Jeong HJ, Kim TY, Hwang HG, Choi HJ. Comparison of thresholding methods for breast tumor cell segmentation. In: Proceedings of 7th International Workshop on Enterprise networking and Computing in Healthcare Industry, 2005. HEALTHCOM 2005. IEEE; 392-5. Available from: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=1500489. [Last accessed on 2012 Oct 23].
3. Dalle JR, Leow WK, Racoceanu D, Tutac AE, Putti TC. Automatic breast cancer grading of histopathological images. Conf Proc IEEE Eng Med Biol Soc 2008;2008:3052-5.
4. Basavanhally AN, Ganesan S, Agner S, Monaco JP, Feldman MD,

Tomaszewski JE, *et al*. Computerized image-based detection and grading of lymphocytic infiltration in HER2+breast cancer histopathology. IEEE Trans Biomed Eng 2010;57:642-53.
5. Chekkoury A, Khurd P, Ni J, Bahlmann C, Kamen A, Patel A, *et al*. Automated malignancy detection in breast histopathological images. In: Pelc NJ, Haynor DR, van Ginneken B, editors. SPIE Medical Imaging; 2012:831515–831515–13. Available from: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1285467. [Last accessed on 2012 Oct 23].
6. Wang CW, Fennell D, Paul I, Savage K, Hamilton P. Robust automated tumor segmentation on histological and immunohistochemical tissue images. In: Serrano-Gotarredona T, editor. PloS one. 2011;6:e15818.
7. Viswanath S, Madabhushi A. Consensus embedding: theory, algorithms and application to segmentation and classification of biomedical data. BMC Bioinform 2012;13):26.
8. Dasgupta S. Experiments with random projection. In: Proceedings of the Sixteenth conference on Uncertainty in artificial intelligence; 2000. p. 143-51.
9. Khan AM, El-Daly H, Rajpoot N. RanPEC: Random Projections with Ensemble Clustering for Segmentation of Tumor Areas in Breast Histology Images. In: Medical Image Understanding and Analysis (MIUA). Swansea, UK; 2012:17-23. Available from: http://miua2012.swansea.ac.uk/uploads/Site/Programme/CS01.pdf. [Last accessed on 2012 Oct 23].
10. Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell 2005;27:1226-38.
11. Galea MH, Blamey RW, Elston CE, Ellis IO. The nottingham prognostic index in primary breast cancer. Breast Cancer Res Treat 1992;22:207-19.
12. Magee D, Treanor D, Chomphuwiset P, Quirke P. Context Aware Colour Classification in Digital Microscopy. In: Proceedings of Medical Image Understanding and Analysis. London, UK; 2011. Available from: http://www.biomedical-image-analysis.co.uk/images/stories/magee-posters2-29.pdf. [Last accessed on 2012 Oct 23].
13. Pietro Perona JM. Scale-space and edge detection using anisotropic diffusion. Available from: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.108.2553. [Last accessed on 2012 Oct 23].
14. Daugman JG. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J Opt Soc Am A 1985;2:1160.
15. Schwarz MW, Cowan WB, Beatty JC. An experimental comparison of RGB, YIQ, LAB, HSV, and opponent color models. ACM Trans Graph 1987;6:123-58.
16. Murtaza K, Khan S, Rajpoot N. VillageFinder: Segmentation of Nucleated Villages in Satellite Imagery. British Mission Vision Conference 2009; 83:1-11.
17. Khan AM, El-daly H, Rajpoot NM. A Gamma-Gaussian Mixture Model for Detection of Mitotic Cells in Breast Cancer Histopathology Images. J Pathol Inform (Accepted 2013).