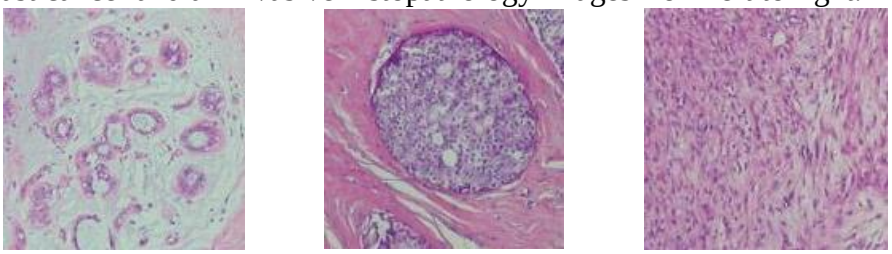
Toward breast cancer histopathology image diagnosis using local color binary patterns

Xingyu Li*, and Konstantinos N. Plataniotis
Multimedia Lab, The Edward S. Rogers Department of Electrical and Computer Engineering
University of Toronto, 10 King's College Road, Toronto, Canada

Introduction. Breast cancer is the most common female cancer in the world. Generally, breast cancer, either "in-situ" or "invasive", has different characteristics from normal breast tissue. In specify, in the normal case, a group of healthy cells are placed along ducts' walls; However in the "in-situ" breast cancer, carcinoma cells fill the ducts, developing further to the "invasive" cancer where malignant cells spread into nearby tissues. In breast cancer diagnosis via histopathology images, Hematoxylin and Eosin (H&E) are usually used to stain biopsy samples for visibility enhancement. In H&E stained images, Hemaxotylin stains nuclei in cells blue/purple, and Eosin stains cytoplasm and connective tissues pink. Hence, color distribution, or color variation, in an image actually reveals the spatial organization of cells. Figures bellow show examples of an H&E stained normal sample, an in-situ breast cancer and an invasive histopathology images from left to right.



In image analysis, image textures describe the spatial distribution and local variation of image intensities or colors. Since color distribution patterns in H&E stained histopathology images are important information source for malignant cell distributions, this paper innovates the use of the powerful texture descriptor, local color binary pattern (LCBP) proposed in [1], in breast cancer diagnosis scenario.

Methods. In an H&E stained breast biopsy image, thanks to the chemical staining, color textures reveal the nuclei spatial structures in an image. Hence, for each H&E stained image, LCBP features are extracted. Let $I=[i_1,i_2,i_3]$ represent an RGB image, where the three components represent the red, green, and blue components respectively. According to [1], LCBP consists of 4 local binary patterns (LBP), among which one is computed from color norm $R=||I||=(i_1^2+i_2^2+i_3^2)^{0.5}$, and the other three LBPs are estimated from ratios of pixel values within color channel pairs $\gamma^{(n,m)}=\arctan[i_n/(i_m+\varepsilon)]$ for $n<m,n=1,2,3$ and $m=2,3$. Consequently, 4 indexed images are generated, which are color norm image R, and color angular images $\gamma^{(1,2)}$, $\gamma^{(1,3)}$, $\gamma^{(2,3)}$. For each of these indexed images, LBP operator with 8-pixel circular neighbourhood is applied, resulting in a 256-length histogram. Then the obtained 4 histogram is concatenated, forming a long feature vector LCBP.

To evaluate the discriminative power of the LCBP pattern in breast cancer histopathology images, the LCBP feature is fed into a classification pipeline, where the feature is first processed by PCA to reduce the feature dimension and then passed to a linear classifier. In this study, we use a breast cancer image set published in 2015 [2] which contains 361 H&E stained breast cancer sample images. Since the image set contains 119 normal tissue images, and 242 in-situ or invasive breast cancer images, we maintain this prevalence in the 10-fold cross-validation experiment. To quantitatively evaluate the performance of LCBP, the classification precision, sensitivity, specificity, and accuracy are summarized based on 10 10-fold cross validation results.

Results. The classification performance on the breast cancer histopathology images is reported as follows: precision=0.8407±0.076, sensitivity=0.7748±0.078, specificity=0.8905±0.071, accuracy=0.8751±0.0382. It is noteworthy that the average classification accuracy outperforms the diagnosis accuracy 0.823 reported in [2].

Conclusions. Motivated by the observation that color distribution conveys the information of nuclei spatial organization in an H&E stained breast histopathology image, we innovate the use of LCBP to describe the nuclei distribution characteristics. Our experiment suggests that LCBP is effective for breast cancer diagnosis. Future work will focus on whether LCBP is discriminative for in-situ and invasive breast cancer classification.

References. [1] Lee et al, "Local color vector binary patterns from multichannel face images for face recognition," IEEE Trans. Image Process., vol. 21, no. 4, pp. 2347 – 2353, Apr. 2012. [2] S.H. Bhandari, "A bad of features approach for malignancy detection in breast histopathology images", IEEE ICIP, 2015.