



Digital Assessment of Stained Breast Tissue Images for Comprehensive Tumor and Microenvironment Analysis

Shachi Mittal^{1†}, Catalin Stoean^{2†}, Andre Kajdacsy-Balla³ and Rohit Bhargava^{1,4*}

¹ Department of Bioengineering, Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana–Champaign, Urbana, IL, United States, ² Department of Computer Science, University of Craiova, Craiova, Romania, ³ Department of Pathology, University of Illinois at Chicago, Chicago, IL, United States, ⁴ Departments of Mechanical Science and Engineering, Electrical and Computer Engineering, Chemical and Biomolecular Engineering, and Chemistry, Cancer Center at Illinois, University of Illinois at Urbana–Champaign, Urbana, IL, United States

OPEN ACCESS

Edited by:

April Khademi,
Ryerson University, Canada

Reviewed by:

Shihao Shen,
University of California, Los Angeles,
United States

Nitish Kumar Mishra,
University of Nebraska Medical
Center, United States

*Correspondence:

Rohit Bhargava
rbx@illinois.edu

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Bioinformatics and Computational
Biology,
a section of the journal
Frontiers in Bioengineering and
Biotechnology

Received: 13 February 2019

Accepted: 13 September 2019

Published: 01 October 2019

Citation:

Mittal S, Stoean C, Kajdacsy-Balla A
and Bhargava R (2019) Digital
Assessment of Stained Breast Tissue
Images for Comprehensive Tumor and
Microenvironment Analysis.
Front. Bioeng. Biotechnol. 7:246.
doi: 10.3389/fbioe.2019.00246

Current histopathological diagnosis involves human expert interpretation of stained images for diagnosis. This process is prone to inter-observer variability, often leading to low concordance rates amongst pathologists across many types of tissues. Further, since structural features are mostly just defined for epithelial alterations during tumor progression, the use of associated stromal changes is limited. Here we sought to examine whether digital analysis of commonly used hematoxylin and eosin-stained images could provide precise and quantitative metrics of disease from both epithelial and stromal cells. We developed a convolutional neural network approach to identify epithelial breast cells from their microenvironment. Second, we analyzed the microenvironment to further observe different constituent cells using unsupervised clustering. Finally, we categorized breast cancer by the combined effects of stromal and epithelial inertia. Together, the work provides insight and evidence of cancer association for interpretable features from deep learning methods that provide new opportunities for comprehensive analysis of standard pathology images.

Keywords: breast cancer, microenvironment, deep learning, ductal carcinoma *in-situ*, hyperplasia and clustering

INTRODUCTION

The current clinical pipeline for cancer diagnosis involves human expert evaluation of large sections of tissue stained with dyes or a variety of specialized molecular markers (Kiernan, 1999). Increasing cancer incidence rates put an increased burden on pathologists worldwide (Williams et al., 2017), as limited resources and limited growth of trained medical personnel (Robboy et al., 2015) are subjected to greater strains. In addition to the emerging needs for better diagnoses, challenges lie in accurate diagnoses using current methods as well (Robbins et al., 1995; Bahreini et al., 2015; Elmore et al., 2015; Stålhammar et al., 2016). Morphometric features form the bases of decisions today; however, difficulty in recognition of subtle morphologic changes and the process of assigning a discrete grade to a continuum of disease often makes the diagnosis prone to under or over-diagnosis. With the advent of whole slide imaging (Pantanowitz et al., 2011; Ghaznavi et al., 2013), digitized versions of stained slides are available and advanced computational analysis can be readily applied (Gurcan et al., 2009). Digital screening of images, even for simple classifications like

“cancer” vs. “no cancer,” can help reduce the pathologist’s workload (Veta et al., 2014) by triaging and focusing on specific areas as well as alerting them to borderline cases (Dong et al., 2014). Deep convolutional neural networks (CNN) are especially rising in popularity as the method of choice for image processing in medicine, due to their inherent ability to automatically extract features, ranging from those very general to those very specific to the problem under consideration (Ertosun and Rubin, 2015; Litjens et al., 2016; Bychkov et al., 2018; Coudray et al., 2018). Histopathology is an area where the CNN architectures can play an important role due to the intricacy of the decisions and the abundance of data resulting from routine patient screening, digital archiving of data and advances in imaging (Litjens et al., 2016). Depending on the quality of the histopathological images, a CNN architecture offers great flexibility and enables a variety of choices. For example, methods have been proposed to work directly with the available resolution if this is low (Postavaru et al., 2017) or on patches from high resolution images with the multiple decisions being further integrated by different techniques (Komura and Ishikawa, 2018). From hand-crafted architectures (Mishra et al., 2018) to pre-trained networks (Araújo et al., 2017) or residual models (Huang and Chung, 2018), CNNs have gained importance in the last few years in their use for tissue segmentation. One area of active application is in diagnosing and understanding breast cancer (Doyle et al., 2008; Abdel-Zaher and Eldeib, 2016; Ehteshami Bejnordi et al., 2017, 2018; Guo et al., 2019; Harvey et al., 2019; Xie et al., 2019). While progress has been made in mimicking traditional processes involving epithelial transformations (Hu et al., 2005; Finak et al., 2008; Beck et al., 2011; Conklin and Keely, 2012; Mao et al., 2013), another avenue to improved diagnoses is the use of the tumor microenvironment (Finak et al., 2008; Hanahan and Coussens, 2012; Mao et al., 2013; Giussani et al., 2018; Jia et al., 2018), both from conventional microenvironment measures and in using emerging imaging techniques (Kumar et al., 2013). In this study, we analyzed breast tissue microarrays (TMA) and surgical specimens to identify different cell signatures in the tumor and its microenvironment using both unsupervised and supervised strategies. First, a deep learning model was built to separate the epithelial, stromal, and other cellular components of the tissue. This allowed for precise investigation of different cellular distributions and their features in each of these components. Next, we evaluated the epithelial and stromal regions as indicators of cancer. Together, we sought to determine common features that expand the palette of digital markers to characterize breast cancer.

MATERIALS AND METHODS

Data Collection

A tissue microarray (TMA) consisting of 100 patient samples, each 1 mm in diameter was obtained from US Biomax Inc. It spans a wide range of disease state i.e. hyperplasia (20 cases), atypical hyperplasia (20), invasive (20 ductal and 20 lobular) and normal (20). This allows for a generalizable model development as it encapsulates a broad range of tumor and cell heterogeneity. These are formalin fixed and paraffin

embedded and 5 microns thick samples that are typically used in clinics. Multiple consecutive sections of this TMA were used to stain for different molecular markers. In this study, we have utilized the Hematoxylin and Eosin (H&E) stained images. To take into account sample preparation and staining variability, some tissue sections were also obtained from the breast tissue registry at Washington School of Medicine at St. Louis. All the stained slides were scanned using the Hamamatsu whole slide scanner.

Data Preparation for Building the CNN Model

Images of whole H&E (hematoxylin and eosin) slides, each corresponding to a different patient, were acquired. The H&E stains are used to identify the morphological changes associated with cancer development and its subsequent progression (Fischer et al., 2008). The data set that is used for training the CNN architecture does not depend on slides to be manually annotated entirely, which would necessitate intensive effort from human experts. However, a pathologist does annotate some sections from each image that correspond to one of the following classes: epithelium (normal, malignant, DCIS), dense stroma, reactive stroma, and the rest in a class that is further referred as *others*. Each contour is chosen to be specific to its class and its size varies from very small regions having an area of $<1\text{ mm} \times 1\text{ mm}$ up to large ones with areas of about $1.5\text{ mm} \times 1.5\text{ mm}$. There are 222 such contours annotated and each one of them is used to create image tiles that will further comprise the training, validation and test sets of the CNN model. Squared tiles varying in size from 48×48 pixels up to 256×256 pixels are cropped from each annotated section. The sizes of the tiles are randomly generated within the mentioned bounds. Random sizes for the cropped images are considered with the aim of feeding the CNN model with sections that contain multiple views of the tissue, from capturing minor details in the smallest squared image to having an overview in the largest squared tile case. The positions of the cropped images are randomly generated such that they lie entirely inside the contour. All the tiles are resized to the minimal size of 48×48 pixels when they finally enter the CNN model. Next, a filter is applied to remove tiles with minimal tissue. If the background comprises more than 30% of the tile area, it is not passed to the model. The amount of tiles that are cropped from each annotated contour is proportional to the area of the annotation: the larger the initial area, the more patches are selected, but no more than 100 tiles per annotation. The actual amount of tiles is obtained by dividing the area of the annotated region to the area of a tile with a side of 152 (mean between the minimum 48 and the maximum 256); the double of this division result represents the number of tiles selected for the annotated area. This ensures that a large part of the contour is cropped with overlapping tiles. Also, distinct patients are used for training, validation and test sets for robust classification and unbiased estimates. There are 6 patients for training and 5 others for validation, the amount of tiles generated for the training set is 3914 for the epithelial class, 2578 for stroma and 2122 for others. For the validation set, 1758 patches for epithelium, 1114

for stroma and 414 for others are used. The test set is represented by an independent TMA consisting of 100 patients.

Tissue Segmentation Pipeline

The first step of the process for the H&E slides segmentation is the separation between components such as epithelium, stroma and other cellular components. A CNN architecture is employed to achieve this delineation. Next, unsupervised segmentation is applied to both the epithelial and stromal compartments for further investigation in relating to disease. A K-means clustering operation is applied for the pixel values of each component, with the aim of distinguishing between regions in the image that belong to different disease states and the type of cellular moieties present around them. A broad overview of this pipeline is illustrated in **Figure 1**.

CNN Model

The goal of the CNN model is to classify large tissue slides into three classes (epithelium, stroma, and others). Thus, each tile will be associated to one of the learned classes and a global matrix of classes is obtained for the entire slide, where each position in the matrix corresponds to a tile from the image. To achieve this, we are faced with the following competing constraints: while the size of the tile should be large enough to contain meaningful information it should also be as small as possible to make the overall tissue image mask look less rasterized and able to delineate borders between classes. To balance these constraints, we resized the tiles to 48×48 , a size that is considered the minimum acceptable threshold in Keras applications (Chollet, 2017), which are pre-defined deep learning models with pre-trained weights that are known to be useful for prediction as well as feature extraction. Three CNN models, i.e., ResNet50 (He et al., 2015), VGG19 (Simonyan and Zisserman, 2014) and InceptionResNetV2, were implemented to classify the

data using transfer learning and default parameters. The results of the three were competitive with image tiles of 197×197 pixels as input, with ResNet50 reaching a classification accuracy of 98.74%, VGG19 97.18% and InceptionResNetV2 94.93%. However, VGG19 was kept for further parameter tuning since it conveniently allowed tiles as low as 48×48 pixels as input, with the classification accuracy decreasing to 96.14%. Nevertheless, it was considered as a good tradeoff ($\sim 1\%$), since smaller tiles lead to a more refined classification in the larger images.

The general architecture of the model used is illustrated in **Figure 2**. VGG19 has a sequence of several layers that perform convolution (with the given kernel size and number of filters) and max-pooling for down sampling (with the specified filter size and stride). They are followed by a layer that flattens its output volume into a vector to be used as input for the subsequent fully connected layers. A dense layer with a Rectified Linear Unit (ReLU) activation function with 1,024 units continues the suite, followed by a dropout layer with the aim to improve the generalization ability, and consequently decrease overfitting. Finally, a dense layer with a softmax activation function gives the n-dimensional output of the network, given in the form of probabilities of belonging to each class, where n is the number of classes (three in our case). A batch size of 32 images is considered.

The weights of the initial layers are preserved (frozen) from a pre-trained network as they refer to general feature information obtained from learning over a large database. We searched for a number of trainable end layers that is appropriate for learning H&E images and the optimal number was found to be 12. The optimizer to be used for weight adjustment, when minimizing the difference between the predicted output and the ground truth, was another parameter for which we searched for an appropriate solution. The methods of stochastic gradient descent, RMSprop and Adam with various learning rates and decays were explored. The best solutions were found for Adam with a learning rate of

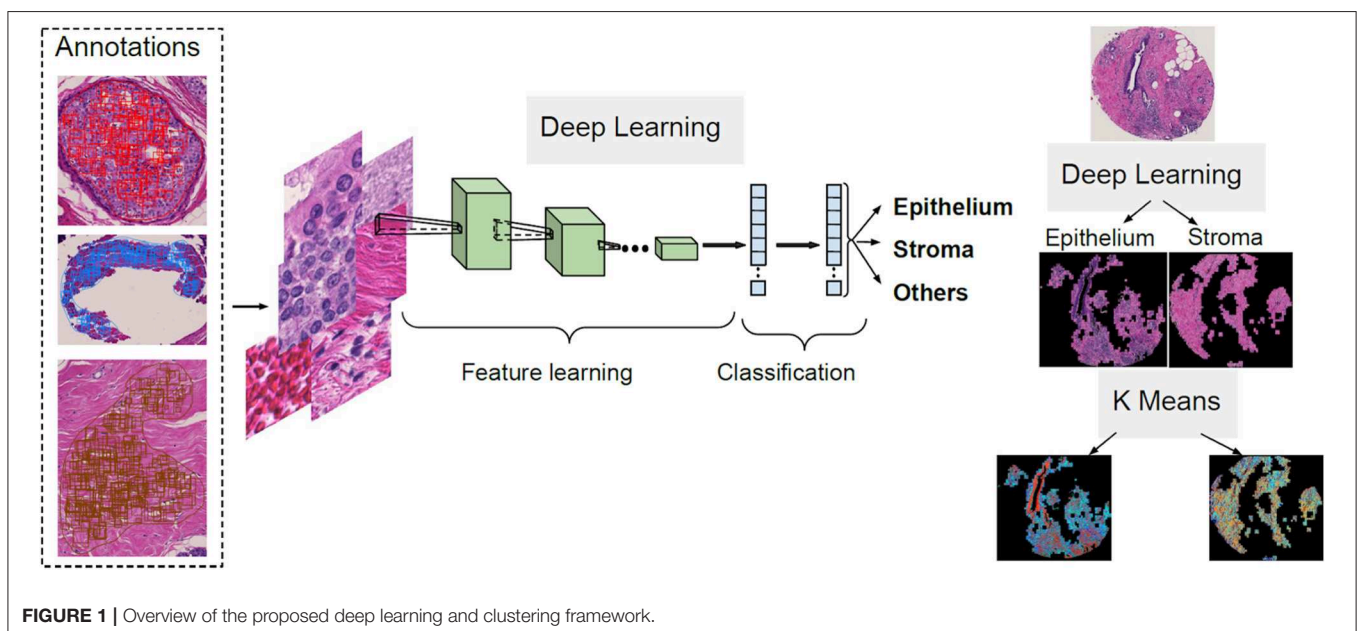
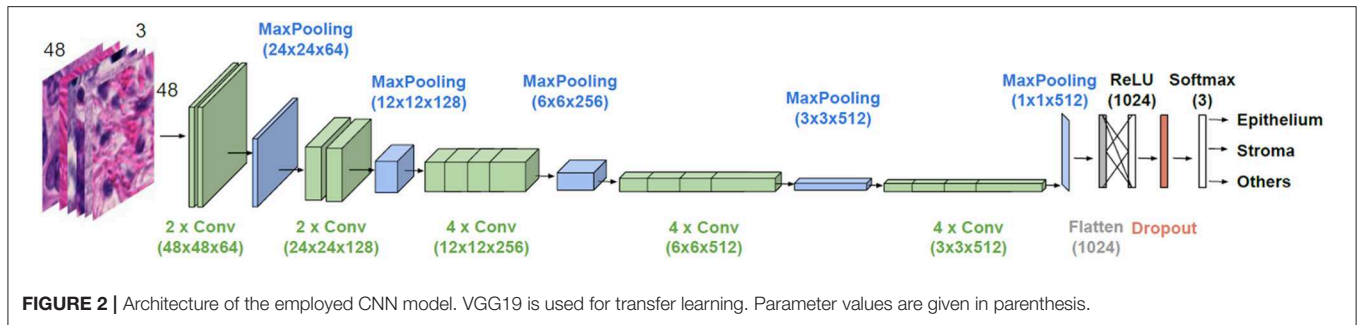


FIGURE 1 | Overview of the proposed deep learning and clustering framework.



1e-07 and without a decay. The dropout rate, giving the percent of units that are blocked, was also tuned and a value of 0.2 was finally used. The decisions took into account the F1 score (reaching in the best configuration 94.42%) on the validation set and also targeted to minimize overfitting.

Based on the information about general tissue architecture, there are rare cases when scarce tiles belonging to one class will have a neighborhood of other classes. For instance, it is uncommon that a few isolated pixels of epithelium will be surrounded by stroma as epithelial cells are present in ducts or lobules. Therefore, a substantial number of pixels from one class would be present next to an appreciable number of pixels from another, especially in the case of epithelium and stroma. So, we apply a majority filter to our classified images to remove classification noise.

Clustering

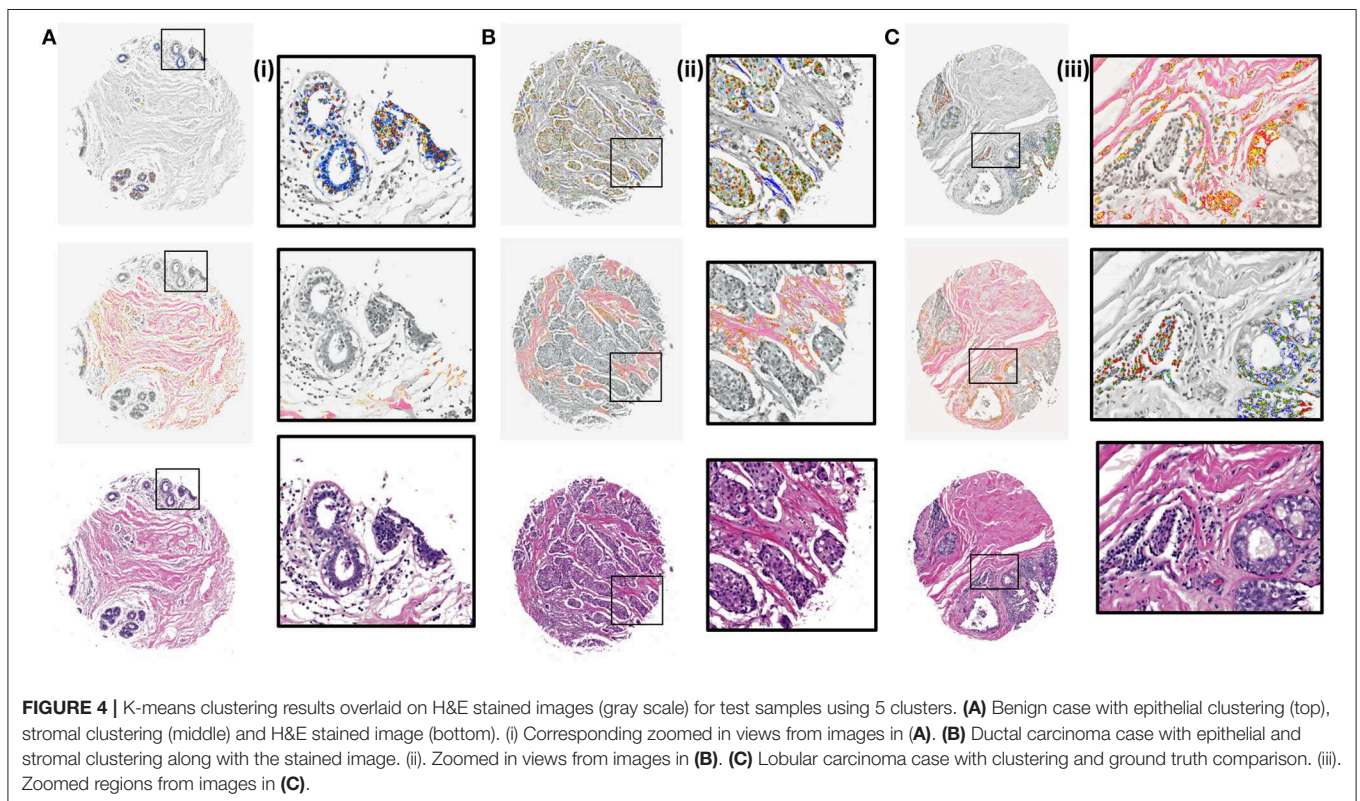
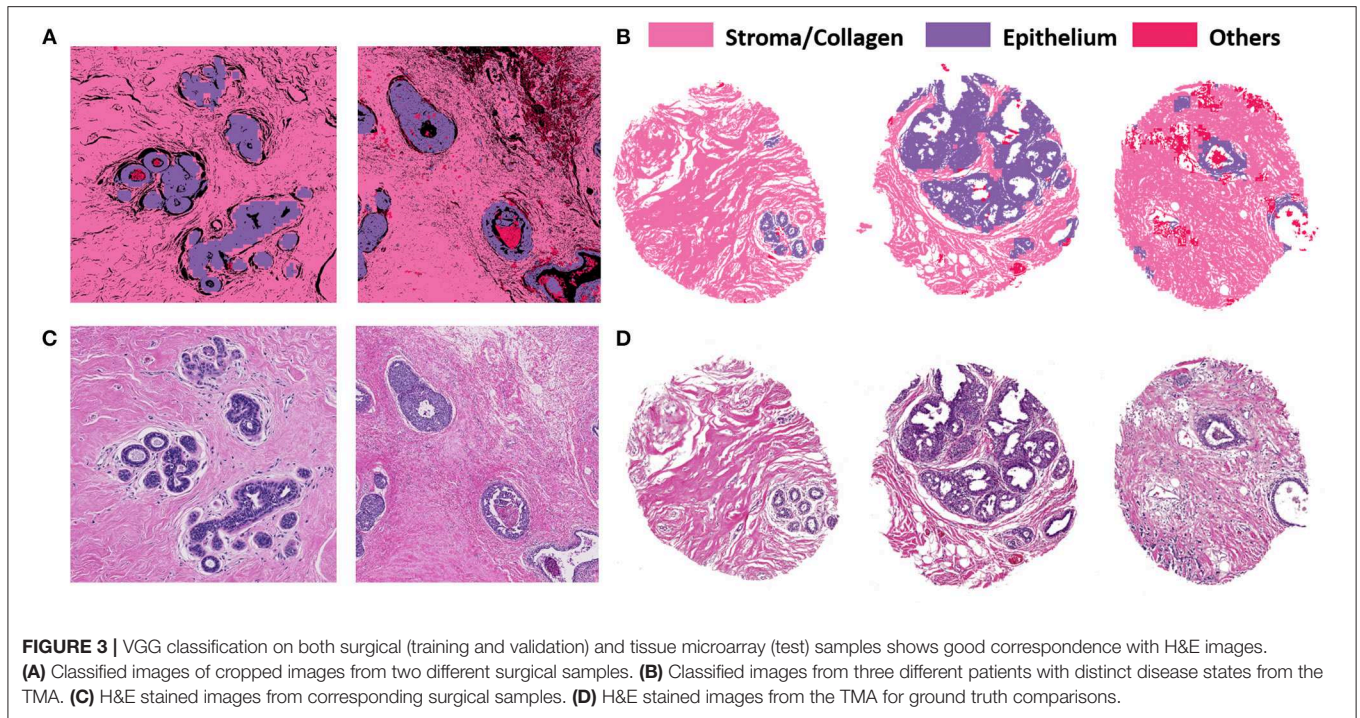
As suggested previously (Beck et al., 2011), it is important to first precisely separate the tissue into the epithelial and stromal compartments for detailed analysis of each of these components for diagnostic or prognostic information. Therefore, we utilize epithelial and stromal regions identified by the deep learning model discussed above. In order to increase the sensitivity of the model, a K-means approach with two clusters is applied to the result obtained to further filter out small regions that were misclassified by the CNN model to only analyze patterns within one cell type at a time. Each detected component is subsequently subject to another K-means clustering algorithm using various number of clusters ranging between 2 and 5. We were interested to measure if there are significant differences between various cluster distributions of each class and how that correlates to distinct disease stages over multiple patients. To achieve this, inertia of both the epithelial and stromal clusters is used. It measures the dispersion of points within a cluster by computing the sum of squared distances for each point to its assigned cluster centroid. A detailed description and code for the CNN model and clustering can be found in the **Supplementary Information**.

RESULTS

Figure 3 illustrates the performance of the developed deep learning strategy for identifying epithelial, stromal and other cellular components (typically consisting of necrosis, red blood

cells, secretions and mucin). VGG19 was the method of choice because it allowed tiles of the smaller considered size, while preserving a high classification accuracy. Panel A shows the classification performance on cropped regions from large surgical samples. Surgical resections represent a challenge in terms of size of the sample and data handling, but do provide unambiguous ground truth for disease that may not be present in limited sections from needle biopsies. High accuracy of tissue segmentation is evident by comparison with ground truth H&E stained images (panel C). We also tested our methods on TMAs, which provide the opportunity for a large number of highly diverse samples but recognize that the potential to examine many components of the tumor and microenvironment in each sample may be limited. **Figures 3B,D** illustrate the model performance on three samples from an independent TMA consisting of 100 patients belonging to different disease states. This array was obtained from a different source to check for model robustness for samples processed and prepared in different institutions or settings. The model performs with similar accuracies on this external validation set. Each sample shown in Panel B (classified images) and Panel D (H&E stained images) belongs to a different patient. The results indicate a good agreement of the histologic units detected by the algorithm (top row) with the ground truth (bottom row). In each case, the prediction was generally good but there were also some small discrepancies as can be seen from the terminal ductal unit in the TMA sample on the left in 3(B) and (D).

Next, K-means images for the components identified by the CNN model were examined. The clustering algorithm is run for 300 iterations, 10 times with random seeds each time and the best solution is kept. It is applied to the RGB values of the masked images based on the CNN results. The clustered images with 5 clusters were overlaid with grayscale H&E stained images to illustrate the class distributions along with a reference to the tissue architecture. Different colors in the overlaid images represent different clusters that are indicative of either different cellular types or subtypes within a cell type. It is evident from zoomed views in **Figure 4** (i), (ii), and (iii) that different signatures within the epithelial (top) and stromal (middle) regions are identified. It can also be noted [**Figure 4** (ii), middle section] that cellular structures like lymphocytes, fibroblasts and plasma cells that are typically present in the stroma get highlighted in the stromal clustering as a different class i.e., a different color. This is important, as sometimes these



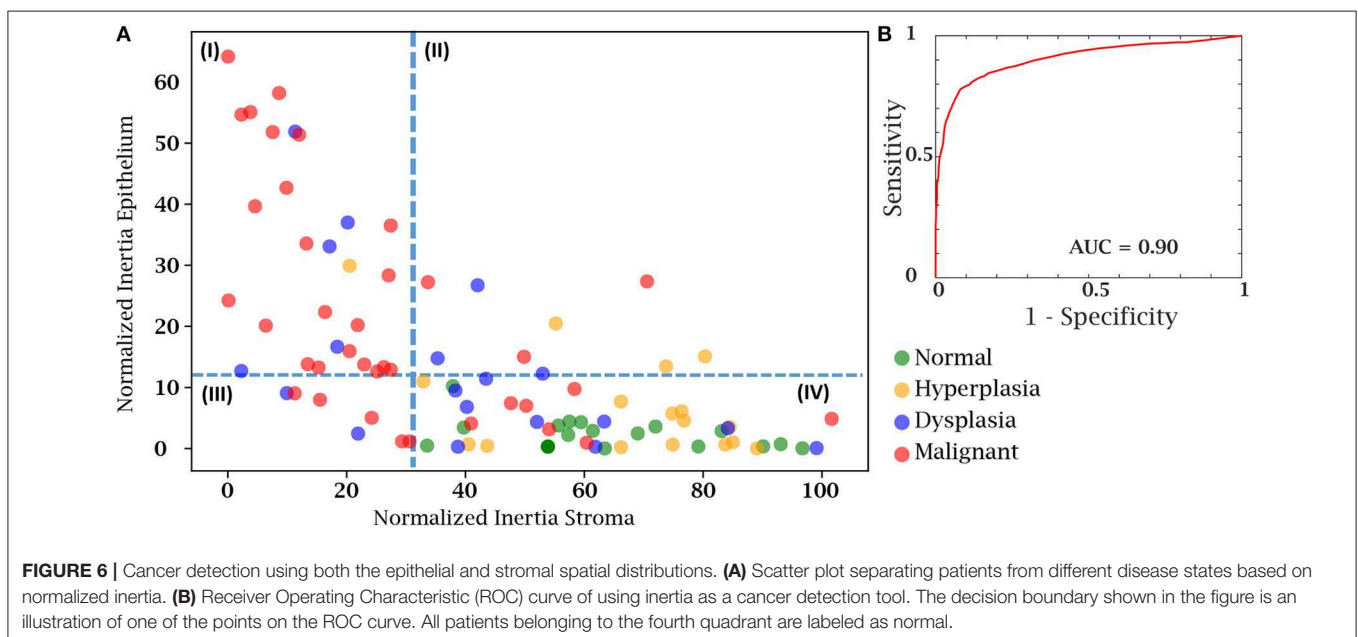
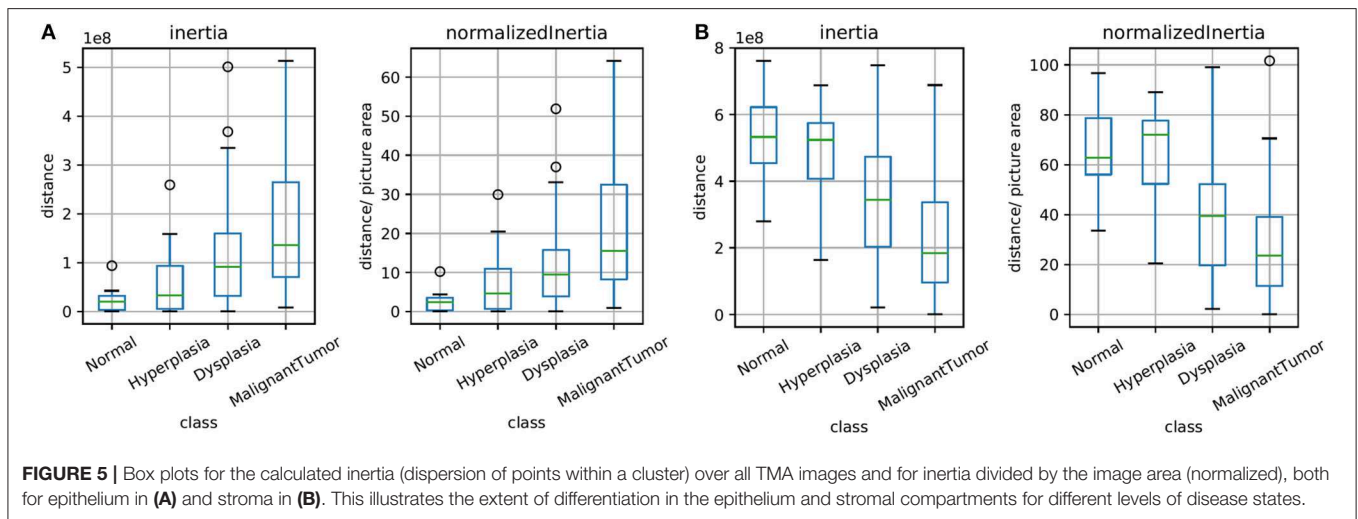
components can be confused with epithelial cells and these cells constitute an important part of the tumor microenvironment. Our method eliminates the confusion between epithelial and

stromal cells by this two-tiered approach and provides a means to discover less abundant cells in the microenvironment. The challenge remains to harness these signatures of different

components obtained by unsupervised discovery for diagnosis. While this work focuses on providing a viable digital pathway, careful curation and labeling of microenvironment components in many surgical samples will be needed to further refine this approach.

The unsupervised clusters are indicative of potential changes in the microenvironment. Here, we sought to explore whether the clusters have any bearing on diagnoses of disease. **Figure 5** illustrates the spread of clusters (inertia) in the epithelial and stromal compartments for different disease states. Inertia is the within cluster sum of squares indicative of intra cluster coherence. For the epithelial distribution in **Figure 5A**, the average inertia for the malignant class is the highest and the normal class is the lowest. It is interesting that the intraductal proliferations (Ellis, 2010) (hyperplasia and dysplasia) also follow

the same trend with the high risk group (dysplasia) having a higher mean inertia and the low risk group (hyperplasia) having a lower inertia. This can help in increasing confidence of the diagnostic decision especially since the histology criteria can be subjective and in some cases not clearly defined. Our hypothesis is that disruption of normal physiologic structure of the tissue results in a difference in inertia, either by decreased order in consistent structures (e.g., changes in epithelial morphology) or imposition of homogeneity on well-differentiated functional units of tissue (e.g., tissue transitioning from clear functional units in a well-defined differentiated pattern to becoming poorly differentiated in space). This hypothesis can be used as a basis for designing digital features that indicate disease. Though a single feature is not unambiguous for any given sample, the trends are useful in adding value to automated methods,



provide a comforting validation of designing digital markers using hypotheses of organization and provide potential for further refinement.

Finally, we sought to use both the epithelial and stromal inertia together to examine whether they provided diagnostic value. From the total of 100 cases in our test case, 8 patient cases either had insufficient epithelial or stromal yield. These were discarded as the epithelium and stromal inertia couldn't be compared. Therefore, we confirmed our hypothesis using an independent set of 92 patient cases belonging to different disease states. We compared the normalized inertia of the epithelial and stromal compartments. By considering a cut-off line for each of the epithelial and stromal figures of merit, we obtain two intersecting lines that divide the entire distribution of inertia into four quadrants. It is evident from **Figure 6A** that the normal and hyperplasia cases are mostly in quadrant IV and malignant cases are distributed in all the other quadrants with a majority lying in the first quadrant. The dotted lines can also be used as a threshold to separate cancer from the normal cases. A variety of 2D thresholds can be used and the receiver operating characteristic curve in **Figure 6B** shows the sensitivity and specificity profile of using inertia as a measure of detecting malignancy. For estimating the ROC curve, the malignant cases are labeled as the cancer class and all other categories i.e., hyperplasia, dysplasia, and normal are combined together into the non-cancer class. The area under the curve is a measure of the model performance and the closer it is to 1, the better the model. The high AUC values suggest that a combinatory epithelial and stromal approach to extract spatial features (inertia in this case) is a good indicator of disease and its progression. Though, again, not a perfect discriminator, the result provides a useful means to utilize the microenvironment and ensure that the results are consistent with underlying diagnoses. Most importantly, the work paves the way for further assessment of complex cellular features in both epithelial cells as well as different stromal cells. Distribution of these cells and other spatial measures of the tumor can provide a further boost in accuracy to the method developed here. An advantage of the proposed method is that it is simple to understand and easy to interpret. Unlike a typical deep learning approach in which images are the input and a decision is the output, this tiered approach provides an insight to understand and interpret some of the vast information encoded in tissue on disease states.

CONCLUSIONS

This paper presents a deep learning scheme for digital tissue segmentation using stained images which can further be used to extract sophisticated spatial features for cancer diagnosis in an automated, fast and accurate manner. Our approach relied on a tiered understanding of tissue structure as deep learning methods were applied. First, we segmented tissue into epithelial and stromal compartments. Second, we examined the

heterogeneity in the stroma using clustering, which resulted in separation of tissue microenvironment components like immune cells and cancer associated fibroblasts. Finally, we sought to use a simple and interpretable characterization of the tissue in terms of inertia to recognize disease. The work establishes the concept of multidimensional inertia to be used as an indicator of disease and its outcome, without needing any additional resources or disruption to the current workflow. The analytics tool can be integrated with whole slide scanners for interactive and real time sample analysis to aid human experts. This study focused on understanding the power of simultaneously using the epithelial and stromal signatures to separate different diseased states. This needs to be further tested in supervised classification schemes along with other tissue predictors. With further refinement of this approach, more spatial signatures and detailed understanding of the tumor and the microenvironment can lead to increased insight into digital characterization of breast cancer using conventionally stained images.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

This study used de-identified human subject specimens. All protocols were performed according to the approved project by the University of Illinois at Urbana-Champaign Institutional Review Board i.e., IRB 06684.

AUTHOR CONTRIBUTIONS

SM and RB designed the research. SM and AK-B collected and annotated the data. CS and SM performed the computational analysis. SM, CS, and RB wrote the manuscript.

FUNDING

The work was supported by the National Institutes of Health through award R01EB009745 (RB).

ACKNOWLEDGMENTS

CS acknowledges the support of the Fulbright program for his research at the Beckman Institute.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fbioe.2019.00246/full#supplementary-material>

REFERENCES

- Abdel-Zaher, A. M., and Eldeib, A. M. (2016). Breast cancer classification using deep belief networks. *Expert Syst. Appl.* 46, 139–144. doi: 10.1016/j.eswa.2015.10.015
- Araújo, T., Aresta, G., Castro, E., Rouco, J., Aguiar, P., Eloy, C., et al. (2017). Classification of breast cancer histology images using Convolutional Neural Networks. *PLoS ONE* 12:e0177544. doi: 10.1371/journal.pone.0177544
- Bahreini, F., Soltanian, A. R., and Mehdipour, P. (2015). A meta-analysis on concordance between immunohistochemistry (IHC) and fluorescence *in situ* hybridization (FISH) to detect HER2 gene overexpression in breast cancer. *Breast Cancer* 22, 615–625. doi: 10.1007/s12282-014-0528-0
- Beck, A. H., Sangoi, A. R., Leung, S., Marinelli, R. J., Nielsen, T. O., van de Vijver, M. J., et al. (2011). Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Sci. Transl. Med.* 3:108ra113. doi: 10.1126/scitranslmed.3002564
- Bychkov, D., Linder, N., Turkki, R., Nordling, S., Kovanen, P. E., Verrill, C., et al. (2018). Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci. Rep.* 8, 1–11. doi: 10.1038/s41598-018-21758-3
- Chollet, F. (2017). “Xception: deep learning with depthwise separable convolutions,” in *Conference on Computer Vision and Pattern Recognition (CVPR) (IEEE)*, 1800–1807. Available online at: <https://ieeexplore.ieee.org/abstract/document/8099678>
- Conklin, M. W., and Keely, P. J. (2012). Why the stroma matters in breast cancer: insights into breast cancer patient outcomes through the examination of stromal biomarkers. *Cell Adh. Migr.* 6, 249–260. doi: 10.4161/cam.20567
- Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyo, D., et al. (2018). Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat. Med.* 24, 1559–1567. doi: 10.1038/s41591-018-0177-5
- Dong, F., Irshad, H., Oh, E.-Y., Lerwill, M. F., Brachtel, E. F., Jones, N. C., et al. (2014). Computational pathology to discriminate benign from malignant intraductal proliferations of the breast. *PLoS ONE* 9:e114885. doi: 10.1371/journal.pone.0114885
- Doyle, S., Agner, S., Madabhushi, A., Feldman, M., and Tomaszewski, J. (2008). “Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features,” in *5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro Proceedings, ISBI*, 496–499. Available online at: <https://ieeexplore.ieee.org/abstract/document/4541041>
- Ehteshami Bejnordi, B., Mullooly, M., Pfeiffer, R. M., Fan, S., Vacek, P. M., Weaver, D. L., et al. (2018). Using deep convolutional neural networks to identify and classify tumor-associated stroma in diagnostic breast biopsies. *Mod. Pathol.* 31, 1–11. doi: 10.1038/s41379-018-0073-z
- Ehteshami Bejnordi, B., Veta, M., Johannes van Diest, P., van Ginneken, B., Karssemeijer, N., Litjens, G., et al. (2017). Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA* 318, 2199–2210. doi: 10.1001/jama.2017.14585
- Ellis, I. O. (2010). Intraductal proliferative lesions of the breast: morphology, associated risk and molecular biology. *Mod. Pathol.* 23, S1–S7. doi: 10.1038/modpathol.2010.56
- Elmore, J. G., Longton, G. M., Carney, P. A., Geller, B. M., Onega, T., Tosteson, A. N., et al. (2015). Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA* 313, 1122–1132. doi: 10.1001/jama.2015.1405
- Ertosun, M. G., and Rubin, D. L. (2015). Automated grading of gliomas using deep learning in digital pathology images: a modular approach with ensemble of convolutional neural networks. *AMIA Annu. Symp. Proc.* 2015, 1899–1908.
- Finak, G., Bertos, N., Pepin, F., Sadekova, S., Souleimanova, M., Zhao, H., et al. (2008). Stromal gene expression predicts clinical outcome in breast cancer. *Nat. Med.* 14, 518–527. doi: 10.1038/nm1764
- Fischer, A. H., Jacobson, K. A., Rose, J., and Zeller, R. (2008). Hematoxylin and eosin staining of tissue and cell sections. *CSH Protoc.* 2008:pdb.prot4986. doi: 10.1101/pdb.prot4986
- Ghaznavi, F., Evans, A., Madabhushi, A., and Feldman, M. (2013). Digital imaging in pathology: whole-slide imaging and beyond. *Annu. Rev. Pathol. Mech. Dis.* 8, 331–359. doi: 10.1146/annurev-pathol-011811-120902
- Giussani, M., Landoni, E., Merlino, G., Turdo, F., Veneroni, S., Paolini, B., et al. (2018). Extracellular matrix proteins as diagnostic markers of breast carcinoma. *J. Cell. Physiol.* 233, 6280–6290. doi: 10.1002/jcp.26513
- Guo, Y., Shang, X., and Li, Z. (2019). Identification of cancer subtypes by integrating multiple types of transcriptomics data with deep learning in breast cancer. *Neurocomputing* 324, 20–30. doi: 10.1016/j.neucom.2018.03.072
- Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M., and Yener, B. (2009). Histopathological image analysis: a review. *IEEE Rev. Biomed. Eng.* 2, 147–171. doi: 10.1109/RBME.2009.2034865
- Hanahan, D., and Coussens, L. M. (2012). Accessories to the crime: functions of cells recruited to the tumor microenvironment. *Cancer Cell* 21, 309–322. doi: 10.1016/j.ccr.2012.02.022
- Harvey, H., Heindl, A., Khara, G., Korkinof, D., O’Neill, M., Yearsley, J., et al. (2019). “Deep learning in breast cancer screening,” in *Artificial Intelligence in Medical Imaging*, eds E. R. Ranschaert, S. Morozov, and P. R. Algra (Cham: Springer), 187–215. doi: 10.1007/978-3-319-94878-2_14
- He, K., Zhang, X., Ren, S., and Sun, J. (2015). *Deep Residual Learning for Image Recognition*. Available online at: <http://arxiv.org/abs/1512.03385> (accessed December 7, 2018).
- Hu, M., Yao, J., Cai, L., Bachman, K. E., Van Den Brù Le, F., Velculescu, V., et al. (2005). Distinct epigenetic changes in the stromal cells of breast cancers. *Nat. Genet.* 37, 899–905. doi: 10.1038/ng1596
- Huang, Y., and Chung, A. C.-S. (2018). “Improving high resolution histology image classification with deep spatial fusion network,” in *Computational Pathology and Ophthalmic Medical Image Analysis*, eds D. Stoyanov, Z. Taylor, F. Ciompi, Y. Xu, A. Martel, L. Maier-Hein, N. Rajpoot, J. van der Laak, M. Veta, S. McKenna, D. Snead, E. Trucco, M. K. Garvin, X. J. Chen, and H. Bogunovic (Cham: Springer), 19–26. doi: 10.1007/978-3-030-00949-6_3
- Jia, Z., Lee, C., Zi, X., McClelland, M., and Mercola, D. (2018). “Tumor microenvironment: prospects for diagnosis and prognosis of prostate cancer based on changes in tumor-adjacent stroma,” in *Precision Molecular Pathology of Prostate Cancer*, eds B. Robinson, J. Mosquera, J. Ro, and M. Divatia (Cham: Springer), 259–275. doi: 10.1007/978-3-319-64096-9_16
- Kiernan, J. A. (1999). Histological and histochemical methods: theory and practice. *Shock* 12:479. doi: 10.1097/00024382-199912000-00012
- Komura, D., and Ishikawa, S. (2018). Machine learning methods for histopathological image analysis. *Comput. Struct. Biotechnol. J.* 16, 34–42. doi: 10.1016/j.csbj.2018.01.001
- Kumar, S., Desmedt, C., Larsimont, D., Sotiriou, C., and Goormaghtigh, E. (2013). Change in the microenvironment of breast cancer studied by FTIR imaging. *Analyst* 138, 4058–4065. doi: 10.1039/c3an00241a
- Litjens, G., Sánchez, C. I., Timofeeva, N., Hermsen, M., Nagtegaal, I., Kovacs, L., et al. (2016). Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci. Rep.* 6, 1–11. doi: 10.1038/srep26286
- Mao, Y., Keller, E. T., Garfield, D. H., Shen, K., and Wang, J. (2013). Stromal cells in tumor microenvironment and breast cancer. *Cancer Metastasis Rev.* 32, 303–315. doi: 10.1007/s10555-012-9415-3
- Mishra, R., Daescu, O., Leavey, P., Rakheja, D., and Sengupta, A. (2018). Convolutional neural network for histopathological analysis of osteosarcoma. *J. Comput. Biol.* 25, 313–325. doi: 10.1089/cmb.2017.0153
- Pantanowitz, L., Valenstein, P. N., Evans, A. J., Kaplan, K. J., Pfeifer, J. D., Wilbur, D. C., et al. (2011). Review of the current state of whole slide imaging in pathology. *J. Pathol. Inform.* 2:36. doi: 10.4103/2153-3539.83746
- Postavaru, S., Stoean, R., Stoean, C., and Caparros, G. J. (2017). “Adaptation of deep convolutional neural networks for cancer grading from histopathological images,” in *International Work-Conference on Artificial Neural Networks* (Cham: Springer), 38–49.
- Robbins, P., Pinder, S., de Klerk, N., Dawkins, H., Harvey, J., Sterrett, G., et al. (1995). Histological grading of breast carcinomas: a study of interobserver agreement. *Hum. Pathol.* 26, 873–879. doi: 10.1016/0046-8177(95)90010-1
- Robboy, S. J., Gupta, S., Crawford, J. M., Cohen, M. B., Karcher, D. S., Leonard, D. G. B., et al. (2015). The pathologist workforce in the United States: II. An interactive modeling tool for analyzing future qualitative and quantitative

- staffing demands for services. *Arch. Pathol. Lab. Med.* 139, 1413–1430. doi: 10.5858/arpa.2014-0559-OA
- Simonyan, K., and Zisserman, A. (2014). *Very Deep Convolutional Networks for Large-Scale Image Recognition*. Available online at: <http://arxiv.org/abs/1409.1556> (accessed December 7, 2018).
- Stålhammar, G., Fuentes Martinez, N., Lippert, M., Tobin, N. P., Mølholm, I., Kis, L., et al. (2016). Digital image analysis outperforms manual biomarker assessment in breast cancer. *Mod. Pathol.* 29, 318–329. doi: 10.1038/modpathol.2016.34
- Veta, M., Pluim, J. P. W., van Diest, P. J., and Viergever, M. A. (2014). Breast cancer histopathology image analysis: a review. *IEEE Trans. Biomed. Eng.* 61, 1400–1411. doi: 10.1109/TBME.2014.2303852
- Williams, B. J., Bottoms, D., and Treanor, D. (2017). Future-proofing pathology: the case for clinical adoption of digital pathology. *J. Clin. Pathol.* 70, 1010–1018. doi: 10.1136/jclinpath-2017-204644
- Xie, J., Liu, R., Luttrell, I. V., J., and Zhang, C. (2019). Deep learning based analysis of histopathological images of breast cancer. *Front. Genet.* 10:80. doi: 10.3389/fgene.2019.00080

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Mittal, Stoean, Kajdacsy-Balla and Bhargava. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.