

A model of tumor architecture and spatial interactions with tumor microenvironment in breast carcinoma

Bassem Ben Cheikh¹, Catherine Bor-Angelier², and Daniel Racoceanu^{1,3}

¹Sorbonne Universités, UPMC Univ Paris 06, CNRS, INSERM, Biomedical Imaging Laboratory (LIB), Paris, France

²Unicancer - Auvergne Rhône Alpes, Centre Jean Perrin - Service de Pathologie, Clermont Ferrand, France

³Pontifical Catholic University of Peru, Engineering Department, Electrical and Electronics Section, Medical Imaging Lab, San Miguel, Lima, Peru

ABSTRACT

Breast carcinomas are cancers that arise from the epithelial cells of the breast, which are the cells that line the lobules and the lactiferous ducts. Breast carcinoma is the most common type of breast cancer and can be divided into different subtypes based on architectural features and growth patterns, recognized during a histopathological examination. Tumor microenvironment (TME) is the cellular environment in which tumor cells develop. Being composed of various cell types having different biological roles, TME is recognized as playing an important role in the progression of the disease. The architectural heterogeneity in breast carcinomas and the spatial interactions with TME are, to date, not well understood. Developing a spatial model of tumor architecture and spatial interactions with TME can advance our understanding of tumor heterogeneity. Furthermore, generating histological synthetic datasets can contribute to validating, and comparing analytical methods that are used in digital pathology. In this work, we propose a modeling method that applies to different breast carcinoma subtypes and TME spatial distributions based on mathematical morphology. The model is based on a few morphological parameters that give access to a large spectrum of breast tumor architectures and are able to differentiate in-situ ductal carcinomas (DCIS) and histological subtypes of invasive carcinomas such as ductal (IDC) and lobular carcinoma (ILC). In addition, a part of the parameters of the model controls the spatial distribution of TME relative to the tumor. The validation of the model has been performed by comparing morphological features between real and simulated images.

Keywords: Modeling, Tissue simulation, Breast tumor architecture, Tumor microenvironment, Mathematical Morphology

1. INTRODUCTION

Breast carcinoma is a heterogeneous disease originating from the epithelial cells lining the milk ducts or lobules. In order to categorize the various patterns, pathologists classified breast carcinoma into histopathological subtypes based on architectural features and growth patterns¹. The histological classification of breast carcinoma is extremely important, considering the significant implications of the subtypes in the prognosis and treatment of the disease. The three most common histopathological types, that together represent nearly three-quarters of breast carcinomas, are: invasive ductal carcinomas (IDC), ductal carcinoma in situ (DCIS), and invasive lobular carcinoma (ILC), with an estimated prevalence of 55%, 13%, and 5% of all breast cancers respectively². The sub-classification of invasive breast carcinoma includes carcinoma of no special type (NST) new denomination of IDC and some distinguished subtypes such as lobular carcinoma, Mucinous, Tubular, or Medullary carcinomas

From a different angle, tumor cells inhabits a cellular environment composed of parts of different types, such as immune cells, blood vessels, collagen, fat and many other cell types. The whole of these components is referred to as the tumor microenvironment (TME). The interactions between tumor and its TME are recognized as playing an important role in the progression of the disease. In particular, the inflammatory microenvironment is known to have a great impact on the tumor behavior³. Links have been shown to exist between clinical outcome and immune cell presence, relative abundance, as well as spatial proximity of immune cells to cancer

cells.⁴⁻⁹ . On the other hand, it has been shown that invasion and metastasis of breast tumors are influenced by collagen organization at the tumor-stromal interface^{10,11} . By understanding these interactions, the disease may be treated more effectively. Tumor tissue modeling could help investigate new approaches evaluating these interactions and unraveling the perplexing heterogeneity of the disease.

Few frameworks for synthetic histopathology image generation have been proposed in the literature. A model for the spatial microenvironment of healthy and cancerous colon tissue was proposed in¹² . The model is designed to generate synthetic Haematoxylin and Eosin (H&E) images with parameters that allow control over cancer grades based on crypt sizes, cellularity, cell overlap ratio, and lumen texture. The model was first introduced for simulating healthy colonic crypt microenvironments in immunofluorescence images¹³ . To the best of our knowledge, our model is the first to simulate breast tumor architecture at the tissue level. The model is capable of simulating a large range of tissue types, close enough to reality (according to partner pathologist), with few parameters.

Our approach for simulating breast tumor tissue images is based on compound mathematical morphology tools, as they have the ability to design objects and spatial interactions according to topological concepts. The model can be divided into two main modules (see figure 1). The first module deals with the architecture of the tumor regardless the TME components and it generates tumor growth patterns, that are defined by aggregates of epithelial cells in a real tissue. The second module concerns the spatial arrangement of TME components relative to the tumor patterns that were generated during the first step. Even though TME is composed of a wide range of biological components, in this study we restrict the second module to model collagen fibers organization and the spatial distribution of immune cells. In order to simulate tissue images that cover significant histological structures, the size of the simulated images is set to 2048×2048 at $0.5\mu m/pixel$ resolution.

The next section provides a description of the proposed approach for modeling breast tumor tissues. Section 3 describes the experiments and the results of this study. Finally, a conclusion summarizes our main contribution and further works.

2. METHOD

In the following, given a binary image I , we denote $\mathcal{D}_b(I)$ the morphological dilation (respectively, $\mathcal{E}_b(I)$ the morphological erosion) of I with a structuring element b . Let $\{x_i\}_{1 \leq i \leq n}$ be the set of the binary objects that are in I , we denote $\delta_B(I)$ the morphological dilation (respectively, $\epsilon_B(I)$ the morphological erosion) of the n binary objects $\{x_i\}$ with the set of structuring elements $B = \{b_i\}_{1 \leq i \leq n}$. In other words, each binary object x_i is dilated (or eroded) individually with the structuring element b_i , then the results are added to a unique binary image. A disk-shaped structuring element (SE) of radius r is denoted Φ_r . And a line-shaped structuring element of length l and orientation θ is denoted by $\Delta_{l,\theta}$. We denote $\mathcal{U}_n(a, b)$ a set of n random values generated from a discrete uniform distribution on $\llbracket a, b \rrbracket$, and $\mathcal{N}_n(\mu, \sigma)$ a set of n random values generated from a normal distribution with mean μ and standard deviation σ . We define the function $\mathcal{P}_{n,d_{min}} : I \mapsto J$ that generates a binary image J with at most n random white pixels, such that the pixels' coordinates are inside the white surfaces of the binary image I and with a minimum distance d_{min} from each other.

2.1 Tumor patterns

The first characteristics that a pathologist inspects during the examination of a breast tissue specimen are the shape and the spatial distribution of the tumor patterns (TP), that are defined by clusters of cancer cells having various spatial arrangements. The histopathologic classification of breast cancer is based mainly on the spatial appearance of these growth patterns, see figure 5(i-1). In this work, we expose our model through four types of breast carcinomas: Ductal Carcinoma In Situ (DCIS), where tumor cells are found inside, and not moved out of, the mammary duct. Two subtypes of invasive ductal carcinomas; 1) medullary carcinoma (MC), where the tumor pushes against the surrounding stroma but stays relatively circumscribed, and 2) tubular carcinoma (TC), which is made up of tube-shaped structures. And Invasive Lobular Carcinoma (ILC), where cancer cells tend to infiltrate as individual rows known as '*Indian files*'.

In order to design tumor patterns, we generate in a blank image n_0 random white pixels with a minimum distance d_0 from each other. These white pixels are the initial binary objects representing the tumor seeds, from

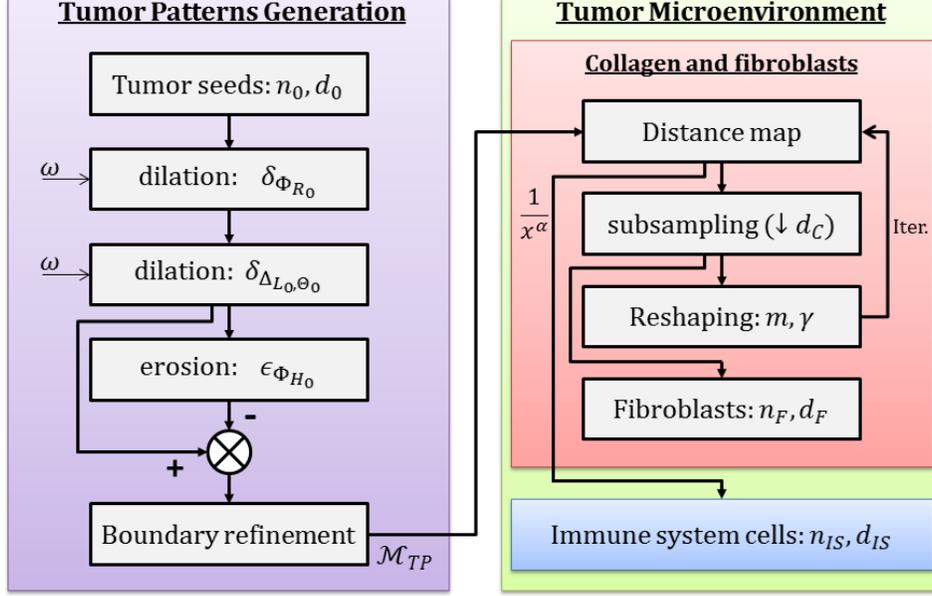


Figure 1: Block diagram of the model. Each step is regulated by parameters that control particular characteristics of the tissue.

which tumor patterns will be created (grown). Each white pixel x_i is dilated by a disk Φ_{r_i} of a random radius r_i , then by a line Δ_{l_i, θ_i} of a random length l_i and a random orientation θ_i . The overlap between the growing seeds during the dilation process is controlled by a Boolean parameter ω . When $\omega = 0$, the overlap is not allowed, the intersection pixels are associated to one of the touching objects, chosen randomly, and the objects are then separated (see examples in figure 5.a). While when $\omega = 1$, the dilation operation allows merging objects (see examples in figure 5.b). In the following, we denote δ_B^ω the morphological dilation with the set of structuring elements B , with overlapping control ω .

In a real tumor tissue, tumor patterns may show empty surfaces, that do not contain tumor cells like lumen of tubules or glands, or necroses. Therefore, holes are created inside the tumor pattern objects by the mean of morphological erosions and subtractions: A morphological erosion is applied to each object with a disk Φ_{h_i} of random radius h_i , and the result is then subtracted from the previous image. It should be noted that if $h_i = 0$, then the erosion is not applied and the tumor pattern x_i is not hollow. Analytically, the binary mask \mathcal{M}_{TP} of the tumor patterns is given by the following operations:

$$\mathcal{M}_I = \delta_{\Delta_{L_0, \Theta_0}}^\omega (\delta_{R_0}^\omega (\mathcal{P}_{n_0, d_0} (I_0))) \quad (1)$$

$$\mathcal{M}_{TP} = \mathcal{M}_I - \epsilon_{\Phi_{H_0}} (\mathcal{M}_I) \quad (2)$$

where I_0 is a blank image. R_0 , L_0 , Θ_0 and H_0 are sets of random variables corresponding to the structuring elements properties. An interpretation of the parameters is given in figure 2.a. Results obtained by different parameter values (in table 1) are shown in figure 5(a-d).

Finally, the boundaries of TPs are refined by generating random white pixels around the objects, followed by alternating sequential filtering in order to make the boundaries "less linear" and more realistic.

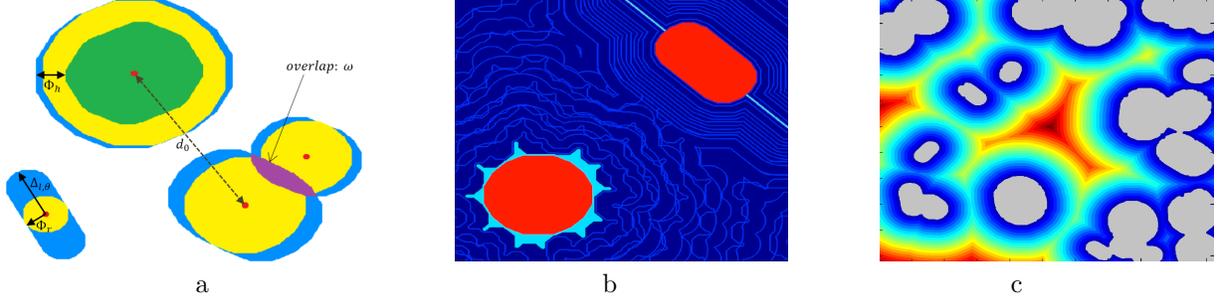


Figure 2: (a) Interpretation of the morphological parameters that regulate the architecture of the tumor growth patterns. (b) Example of two different configurations of collagen fiber curves: TP objects that are output of the first module are shown in red, TP reshaping is shown in cyan, and the generated collagen fiber curves are highlighted in light blue. (c) Example of a distance map that is used for the configuration of spatial distribution of immune cells: TP objects that are output of the first module are shown in grey, the distance map values increase from the coldest-color to the hottest-color layers.

2.2 Collagen fibers

The second aspect we are interested in, is the density and the orientation of collagen fibers relative to the tumor growth patterns. The motivation comes from studies that have shown that the invasion of breast tumors is influenced by collagen organization at the tumor-stromal interface. In fact, the authors of^{10,11}, have defined three Tumor-Associated-Collagen-Signatures (TACS): (i) TACS-1, which is a limited collagen density localized around small tumor foci, (ii) TACS-2 characterized by stretched collagen fibers tangentially oriented along a smooth tumor boundary, and (iii) TACS-3 characterized by collagen fibers aligned perpendicularly to an irregular tumor boundary and oriented in the direction of cell invasion. Hence, we incorporate in our approach parameters adapted to these exposed configurations. In order to introduce in our model an aspect of interactions between tumor patterns (TP) and collagen fibers (CF), the curves of the latter are generated from the reshaping of the objects of the first and their associated distance maps. The reshaping is made in two different fashions:

i) To generate waved lines tangentially oriented along with the boundary of a binary object, we first generate in the image $\mathcal{M}_{\mathcal{TP}}$, $m \times N$ random white pixels around the objects at a distance γ from the TP boundaries, where N is the total number of pixels at the distance γ . Then, we apply a morphological closing $\mathcal{E} \circ \mathcal{D}_{\Phi_\gamma}$. This operation produces deformations and reshaping of the TP objects. The lines corresponding to the collagen fibers are obtained from the subsampling $\downarrow d_C$ of the distance map of the reshaped objects. The parameter d_C controls the number and the distance between fibers, and thus the density. While the parameters m and γ control the stretchiness of the fibers. The whole process is repeated iteratively, such that at each iteration i , $\mathcal{M}_{\mathcal{TP}} \leftarrow \mathcal{D}_{\Phi_{i*\gamma}}(\mathcal{M}_{\mathcal{TP}})$, until the whole image becomes white.

ii) To generate aligned collagen fibers oriented in the direction of the cell invasion, we use the same technique but we apply a different reshaping method. Instead of generating random segments radial to the object boundary (see figure 2.b), we generate a straight line passing through the object center and oriented in the same direction as the object. An examples of two collagen fiber configurations is also given in figure 2.b.

Fibroblasts coordinates were chosen randomly from pixels of the generated curves: $P_F = \mathcal{P}_{n_F, d_F}(C_i)$, where C_i is the collagen curve obtained at iteration i .

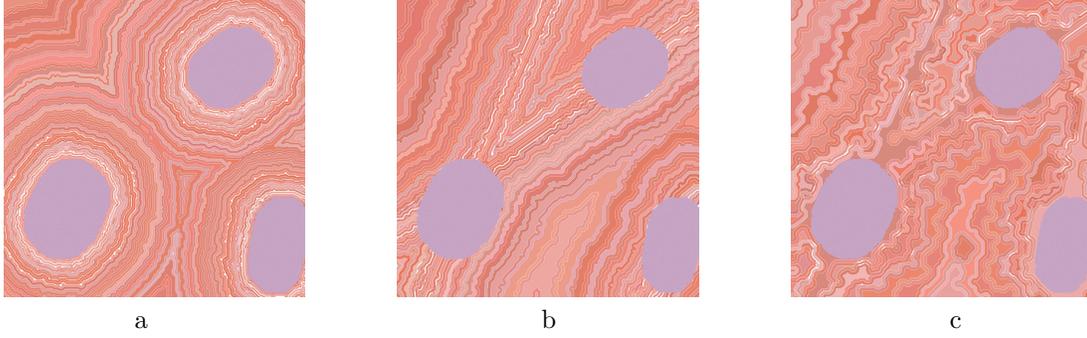


Figure 3: Different configurations of collagen fibers relative to the tumor patterns, obtained by varying the parameters of the model. (a) Collagen fibers are stretched lines, tangentially oriented with the TP boundaries: $d_C = 15$, $m = 0.02$, $\gamma = 20$. (b) Collagen fibers are stretched lines, aligned and oriented in the direction of TPs: $d_C = 20$, $m = 0.01$, $\gamma = 10$. (c) Collagen fibers are unstretched, but oriented in the direction of TPs: $d_C = 20$, $m = 0.05$, $\gamma = 50$

2.3 Immune system cells

It has become obvious that the inflammatory microenvironment (iTME) has a great impact on the tumor progression and behavior. Many studies have shown links between the clinical outcome and the immune cell presence, the relative abundance, as well as the spatial proximity to cancer cells. To that end, we introduce in our model parameters that control the density and the proximity of immune cells to the tumor growth patterns. Let k be the number of layers in the discrete distance map of the binary image $\mathcal{M}_{\mathcal{T}\mathcal{P}}$ that contains the tumor patterns as defined in equation (2), we calculate k coefficients $\lambda_1, \lambda_2, \dots, \lambda_k$, such that $\lambda_i = \frac{i^{-\alpha}}{\sum_{i=1}^k i^{-\alpha}}$, $\forall i \in \llbracket 1, k \rrbracket$.

The coefficient λ_i is the ponderation coefficient associated to the number of immune cells to generate in the layer i of the distance map (an example is shown in figure 2.c). In each layer i we generate $\lambda_i \times n_L$ random immune cells with a minimum distance d_L from each other, where n_L is the desired number of immune cells. If $\alpha > 0$ then $\lambda_i < \lambda_{i-1}$, $\forall i \in \llbracket 1, k \rrbracket$, which means that immune cells are mostly generated in the first layers of the distance map, and thus, collected around the tumor patterns (see example in figure 4.c). If $\alpha < 0$ then $\lambda_i > \lambda_{i-1}$, and consequently immune cells are mostly generated in the last layers of the distance map, away from the tumor-stromal interface (see example in figure 4.a). And finally, if $\alpha = 0$ then $\lambda_i = \lambda_j = \frac{1}{k}$, $\forall i, j \in \llbracket 1, k \rrbracket$ and the spatial distribution is uniform (see example in figure 4.b).

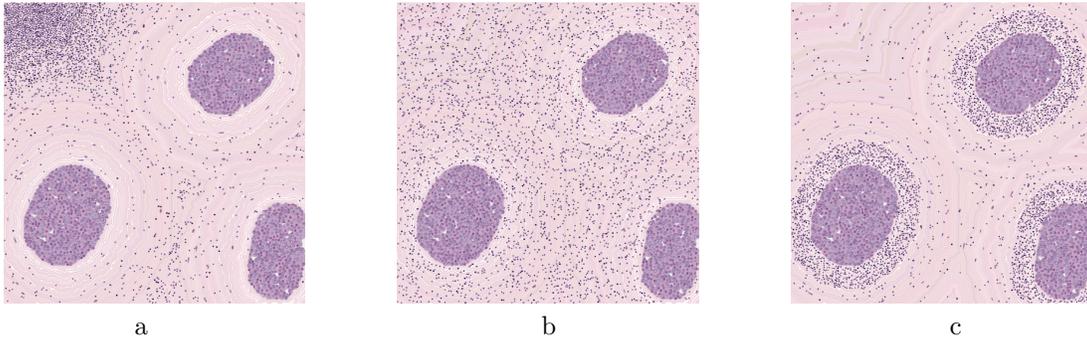


Figure 4: Different spatial distributions of the immune system cells relative to the tumor patterns: In all the configurations, the total number of immune cells is $n_L = 3000$ and the minimum relative distance is $d_L = 10$ pixels. The parameter α is set to -1 , 0 and 3 respectively in (a), (b) and (c).

2.4 Refinement:

Although this study is dedicated to the architectural modeling of breast tumors and tumor microenvironment spatial configurations, we are also interested in the esthetic aspect of the simulated images. Therefore, we have added a third module to draw the simulated tissue components based on ingredients extracted from real images of H&E-stained breast tumor tissue. Patches of nuclei of different cell types, i.e. epithelial cell nuclei, lymphocyte nuclei and fibroblasts, were extracted, randomly rotated, and then pasted into the simulated image at their corresponding locations. In particular, fibroblast patches are oriented in the same direction as their corresponding collagen fibers. Patches containing collagen and tumor patterns were also extracted. Then, in each simulation, pixels are randomly selected from the patches, and their colors are set to the corresponding component in the simulated image. Tiny white surfaces were also randomly generated to simulate the tearing that commonly appears in real tissue images. And finally, gaussian noise and median filter are applied consecutively. Figure 5.f shows an example of simulated medullary carcinoma, where a thin mucinous layer was added around the tumor patterns, using mathematical morphology operations. The approach may provide other options such as artifacts or necrosis.

3. EXPERIMENTS AND RESULTS

In order to assess the effectiveness of our model, we compare our simulation results to real data based on morphological features. The number of tumor patterns (TP), their area, their major axis and minor axis lengths, their solidity and the coefficient of variance (CoV) of their orientations were calculated from the four types of breast carcinoma and compared to the modeled structures. A comparison between results of 80 images (20 of each type) and 32 images from real data is shown in table 2. The calculated feature values are relatively close between the simulated data and the real data within the same group type. Qualitative results, obtained with parameters in table 1, are shown in figure 5.

Table 1. Parameters of the results that are shown in figure 5

	n_0	d_0	R_0	L_0	Θ_0	H_0	ω
DCIS	$\mathcal{N}_1(7, 2)$	800	$\mathcal{N}_{n_0}(400, 80)$	$\mathcal{N}_{n_0}(300, 50)$	$\mathcal{U}_{n_0}(0, 180)$	$\mathcal{N}_{n_0}(0, 0)$	0
MC	$\mathcal{N}_1(40, 8)$	100	$\mathcal{N}_{n_0}(80, 25)$	$\mathcal{N}_{n_0}(70, 15)$	$\mathcal{U}_{n_0}(0, 180)$	$\mathcal{N}_{n_0}(0, 0)$	1
TC	$\mathcal{N}_1(85, 5)$	40	$\mathcal{N}_{n_0}(40, 7)$	$\mathcal{N}_{n_0}(27, 1)$	$\mathcal{U}_{n_0}(0, 180)$	$\mathcal{N}_{n_0}(14, 1)$	1
ILC	$\mathcal{N}_1(100, 5)$	40	$\mathcal{N}_{n_0}(20, 4)$	$\mathcal{N}_{n_0}24(0, 20)$	$\mathcal{N}_{n_0}(\mathcal{U}_1(0, 180), 10)$	$\mathcal{N}_{n_0}(0, 0)$	1

Table 2. mean and standard deviation (in brackets) of the evaluation parameters.

		Nbr. of TP	Area (10^4)	Major axis	Minor axis	Solidity	CoV of Orient.
DCIS	Real	5.14(1.95)	25.5(21)	660(280)	432(195)	0.97(0.03)	0.63(0.43)
	Synthetic	5.67(2.34)	27.6(17)	713(243)	464(126)	0.98(0.02)	0.43(0.12)
MC	Real	22.6(10.7)	5.89(4.3)	375(263)	186(184)	0.89(0.10)	0.61(0.14)
	Synthetic	17.7(12.9)	7.34(5.6)	412(287)	211(191)	0.78(0.23)	0.63(0.22)
TC	Real	40.6(7.63)	2.74(2.3)	238(114)	137(51.5)	0.96(0.04)	0.71(0.11)
	Synthetic	63.8(10.2)	1.97(3.2)	217(129)	120(62.7)	0.94(0.10)	0.83(0.13)
ILC	Real	95.7(23.5)	0.76(0.5)	198(166)	46.7(27.9)	0.89(0.07)	0.32(0.15)
	Synthetic	89.3(35.1)	0.98(0.3)	230(183)	39.2(19.1)	0.84(0.12)	0.29(0.09)

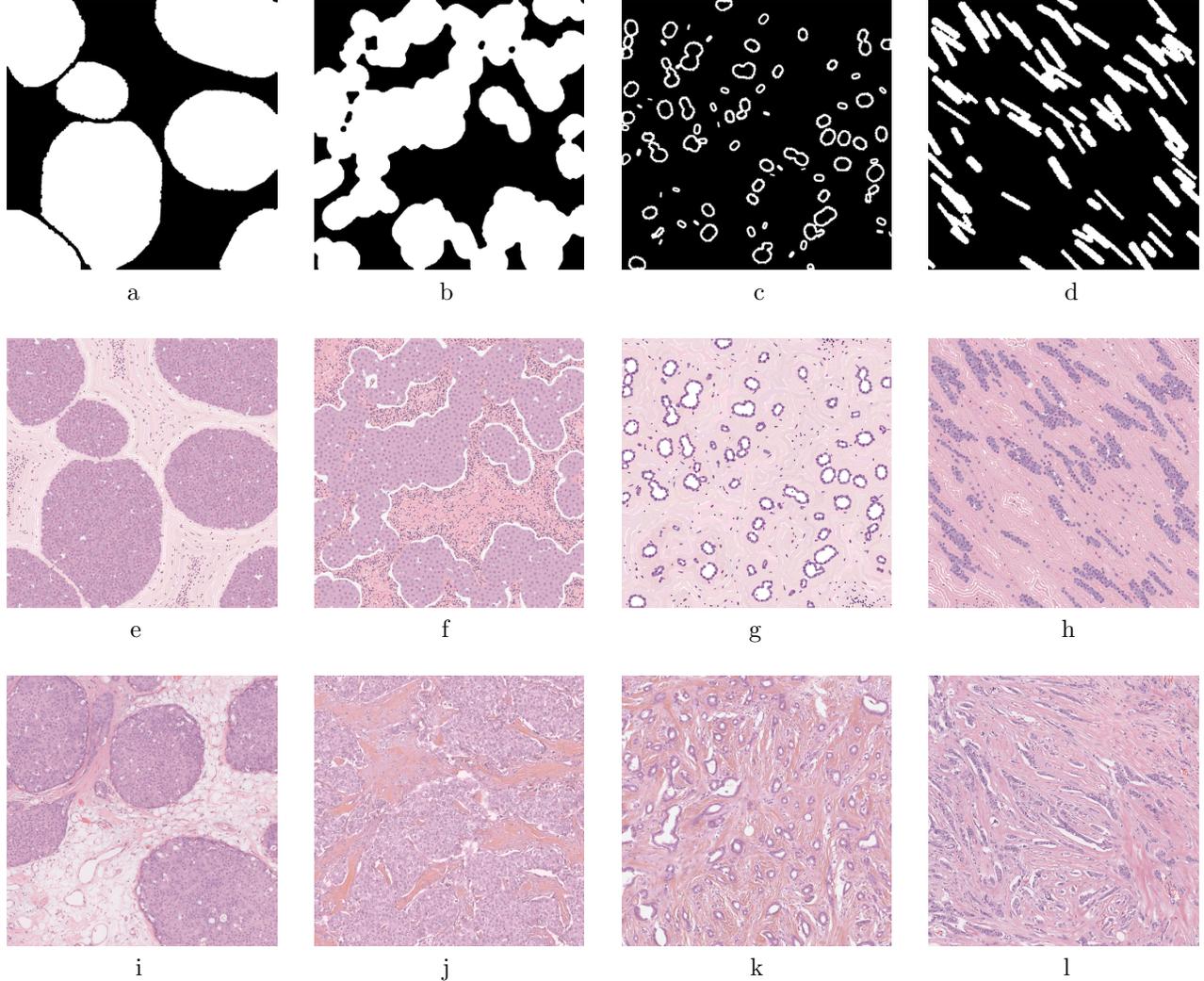


Figure 5: (a-d) tumor pattern masks, $\mathcal{M}_{\mathcal{TP}}$. (e-f) simulated images, associated parameters are given in table 1: (e) DCIS, (f) MC shown immune cells around the TPs, (g) TC, and (h) ILC. (i-l) real tissue images.

4. CONCLUSION

In this work, we presented a model for simulating different breast carcinoma architectures and various tumor microenvironment spatial configurations based on compound mathematical morphology. The validation of the model has been performed by comparing morphological features between real and simulated images. The model gives promising results that are closely related to reality based on few parameters. In the future, we will explore the parameters of the model to simulate other subtypes of breast cancer. We will use the model to study cancer from organ, such as colorectal cancer, and consider more microenvironmental components; such as adipose tissue and blood vessels.

ACKNOWLEDGMENTS

This study has been done with the support of the European Institute of Innovation & Technology (EIT) Health - PAPHOS project and the French Foundation for Medical Research (FRM) - convention DBS20131128436.

REFERENCES

- [1] Malhotra, G. K., Zhao, X., Band, H., and Band, V., [*Histological, molecular and functional subtypes of breast cancers*], Cancer Biology & Therapy (2010).
- [2] Ehemann, C. R., Shaw, K. M., Ryerson, A. B., Miller, J. W., and White, A. A. M. C., [*The Changing Incidence of In situ and Invasive Ductal and Lobular Breast Carcinomas: United States, 1999-2004*], Cancer Epidemiol Biomarkers Prev (2009).
- [3] de Visser KE and LM, C., [*The inflammatory tumor microenvironment and its impact on cancer development*], Contrib Microbiol (2006).
- [4] R., A. H. and al, [*Computational pathology of pre-treatment biopsies identifies lymphocyte density as a predictor of response to neoadjuvant chemotherapy in breast cancer*], Breast Cancer Research (2016).
- [5] Yuan, Y., [*Modelling the spatial heterogeneity and molecular correlates of lymphocytic infiltration in triplenegative breast cancer*], R Soc Interface (2016).
- [6] Chang, A. Y. and al, [*Spatial organization of dendritic cells within tumor draining lymph nodes impacts clinical outcome in breast cancer patients*], Journal of Translational Medicine (2013).
- [7] Setiadi, A. F. and al, [*Quantitative, Architectural Analysis of Immune Cell Subsets in Tumor-Draining Lymph Nodes from Breast Cancer Patients and Healthy Lymph Nodes*], J. Chem. Phys. (2010).
- [8] Sidra, N., Andreas, H., Konrad, K., and Yinyin, Y., [*Beyond immune density: critical role of spatial heterogeneity in estrogen receptor-negative breast cancer*], Modern Pathology (2015).
- [9] Kruger, M. and al, [*Combat or surveillance? Evaluation of the heterogeneous inflammatory breast cancer microenvironment*], J. Pathol (2013).
- [10] Conklin, M. W. and al, [*Aligned Collagen Is a Prognostic Signature for Survival in Human Breast Carcinoma*], American Journal of Pathology (2011).
- [11] PP, P. and al, [*Collagen reorganization at the tumor-stromal interface facilitates local invasion*], BMC Med (2006).
- [12] Kovacheva, V. N., Snead, D., and Rajpoot, N. M., [*A model of the spatial tumour heterogeneity in colorectal adenocarcinoma tissue*], BMC Bioinformatics (2016).
- [13] Kovacheva, V. N., Snead, D., and Rajpoot, N. M., [*A model of the spatial microenvironment of the colonic crypt*], ISBI (2015).