Published in final edited form as:

Ann Biomed Eng. 2015 August ; 43(8): 1803–1814. doi:10.1007/s10439-014-1238-7.

Automatic classification and quantification of cell adhesion locations on the endothelium

Jie Wei^{1,§}, Bin Cai², Lin Zhang², and Bingmei M. Fu²

¹Dept. of Computer Science, City College of New York, New York, NY 10031

²Dept. of Biomedical Engineering, City College of New York, New York, NY 10031

Abstract

To target tumor hematogenous metastasis and to understand how leukocytes cross the microvessel wall to perform immune functions, it is necessary to elucidate the adhesion location and transmigration pathway of tumor cells and leukocytes on/across the endothelial cells forming the microvessel wall. We developed an algorithm to classify and quantify cell adhesion locations from microphotographs taken from the experiments of tumor cell/leukocyte adhesion in individual microvessels. The first step in is to identify the microvessel by a novel gravity-field dynamic programming procedure. Next, an anisotropic image smoothing suppresses noises without unduly mitigating crucial visual features. After an adaptive thresholding process further tackles uneven lighting conditions during the imaging process, a series of local mathematical morphological operators and eigenanalysis identify tumor cells or leukocytes. Finally, a novel double component labeling procedure categorizes the cell adhesion locations. This algorithm has generated consistently encouraging performances on microphotographs obtained from in vivo experiments for tumor cell and leukocyte adhesion locations on the endothelium forming the microvessel wall. Compared with human experts, this algorithm used 1/500–1/200 of the time without having the errors due to human subjectivity. Our automatic classification and quantification method provides a reliable and cost efficient approach for biomedical image processing.

Key Terms

Biomedical image processing; Biomedical measurement; Classification algorithms; Dynamic programming; Adaptive algorithms; Morphological operations

Introduction

Tumor metastasis through blood and/or lymphatic circulations is widely recognized as the critical component in tumor malignancy [3] [7]. Targeting tumor metastasis is thus highly required in cancer treatment. Tumor cell adhesion to the microvasculature, and extravasation into the stroma of target organs are two steps during tumor hematogenous (through blood circulation) metastasis [11] [13]. However, the preferential locations of tumor cell adhesion

[§]Corresponding author. Correspondence: Jie Wei, Dept. of Computer Science, The City College of New York, 160 Convent Ave., New York, NY 10031, wei@cs.ccny.cuny.edu, Phone: 212-650-5604, Fax: 212-650-6248.

Conflict of Interest: N/A

to and transmigration across the endothelial cells forming the microvessel wall remain unclear. Do tumor cells prefer to adhere to and transmigrate across the cell body or the joints between adjacent endothelial cells? To develop effective therapies targeting at tumor adhesion and transmigration, one needs to first classify and quantify the preferential locations of tumor adhesion to and transmigration across endothelium lining the microvessel wall. Another equally important question concerns the preferential pathway for circulating leukocytes to cross the microvessel wall to perform their normal immune functions in the infected tissues. Previously, human operators classify and quantify cell adhesion locations via eyeballing and tape measuring [4] [2]. However, to obtain significant results, many photomicrographs need to be analyzed with tens to hundreds of adherent cells in each photomicrograph. These adherent cells are of different types and sizes. In addition, other innate difficulties are present, such as uneven illuminations during the imaging process and serious systematic/random noises. All these combined make the human-based data classification and quantification exceedingly difficult, time-consuming, error-prone and thus unsustainable. Computer assistance is hence necessary in dealing with this big and noisy data problem.

To automatically classify and quantify cell adhesion locations on endothelium, the first step is to segment out the region corresponding to the microvessel where the endothelial cells, tumor cells or leukocytes of interest are located. Based on the special properties of microvessel walls and cell borders, a novel gravity-field dynamic programming procedure is formulated. Next, a series of anisotropic smoothing, adaptive thresholding, mathematical morphology operators, eigenanalysis and local topological analysis using a novel double component labeling procedure, are employed to determine cell adhesion locations. Empirical results using this algorithm have consistently shown encouraging performances in classifying and quantifying both tumor cell and leukocyte adhesion locations from the photomicrographs obtained from in vivo experiments.

Materials and Methods

I. GRAVITY-FIELD DYNAMIC PROGRAMMING PROCEDURE FOR MICROVESSEL SEGMENTATION

The tumor cell adhesion image in a single perfused microvessel was obtained using the same method as described in [2]. To identify the locations of adherent tumor cells, silver staining was applied to the individually perfused microvessel to illustrate the junctions of endothelial cells forming the microvessel wall [14].

To classify and quantify cell adhesion locations, human operators first evaluated the area of the microvessel with adherent cells. For a photomicrograph as depicted in Fig. 1, a gray-scale (one-byte, 8 bits, of levels ranging from 0 to 255) one of height 354 pixels and width 709 pixels, if done by a human operator, s/he needs to trace the boundary of the microvessel by eyeballing first, and then manually measure the height h_i 's of the microvessels at several horizontal positions i . The area of this microvessel is approximated by the product of the average h_i and the total width of the photomicrograph. To better estimate the area, more h_i 's are needed, which requires more time and efforts from human experts. Thus the first task of this work is to automate this procedure. Although only the area subtended by the

microvessel is currently required, the retrieved shape of a microvessel, with details in its global and local geometry and topology, will be of potential application in future biomedical studies.

Human operators identify the microvessel boundaries due to the following common properties of the boundary pixels:

1. They run from the photomicrograph's left to the right side and roughly connect continuously with possible noises;
2. They are darker than those non-boundary pixels, thus having smaller intensity values;
3. They are the pixels of smaller gray scales closest to the photomicrograph's two horizontal sides, whereas others are cell boundaries inside the microvessel.

These three properties dictate the design of a new segmentation algorithm to automatically reproduce human operator's process. Property 1 demands that either of the two target boundaries, once identified, is a linked path [5], i.e., adjacent pixels along either path should have minor intensity differences, ranging from the left photomicrograph boundary to the right one. Because "darker" pixels have small intensities close to 0, Property 2 requires that pixels along the ideal path should have small intensity values. Suppose I is a photomicrograph with height m and width n , L is a sequence of n pixel coordinates along a possible path, that is, $L[i] = (x_i, y_i)$ is the image coordinate of the i th pixel along a possible path L , where x and y axis are respectively along the horizontal (from left to right) and vertical (up to down) direction with the origin located at the upper left corner of the image. Then the optimal boundary of length n , denoted by B_n , is defined below

$$B_n = \underset{L}{\operatorname{argmin}} \sum_{i=2}^n \{ |I(L[i]) - I(L[i-1])| * \mu(L[i], L[i-1]) + I(L[i]) \}, \quad (1)$$

where $|I(L[i]) - I(L[i-1])|$ is the absolute intensity difference between two adjacent boundary pixels; $\mu(L[i], L[i-1])$ indicates the geometric distance between two pixel locations. According to Property 1, the following causal geometric distance [5] is needed: $\mu(L[i], L[i-1]) = 1$ if $x_i - x_{i-1} = 1$ and $|y_i - y_{i-1}| = 1, \infty$ otherwise.

The three terms on the right hand side of Eq. (1) reflect the requirements demanded by Properties 1 and 2. The target function dictated by Eq. (1) can be rewritten as the sum of the first $n-1$ points and the last point $L[n]$ on B_n . Next, define B_{n-1}^q as the optimal path of length $n-1$ ended at point q , one of the m points in column $n-1$ of photomicrograph I , with minimized value v_{n-1}^q :

$$\begin{aligned} f(L_{n-1}^q) &= \sum_{i=2}^{n-1} \{ |I(L_{n-1}^q[i]) - I(L_{n-1}^q[i-1])| * \mu(L_{n-1}^q[i], L_{n-1}^q[i-1]) + I(L_{n-1}^q[i]) \}, \\ v_{n-1}^q &= \min_{L_{n-1}^q} f(L_{n-1}^q), \\ B_{n-1}^q &= \underset{L_{n-1}^q}{\operatorname{argmin}} f(L_{n-1}^q), \end{aligned} \quad (2)$$

where L_{n-1}^q is any possible path of length $n-1$ ended at q . After some simple algebraic manipulations, Eq. (1) is changed to the following equivalent form:

$$B_n = \operatorname{argmin}_{B_{n-1}^q \cup p} \{v_{n-1}^q + |I(p) - I(q)| * \mu(p, q) + I(p)\}, \quad (3)$$

where p is the last point, i.e., $L[n]$, that minimizes the sum of energy within the bracket. According to Eq. (3), the optimality of the sub-problem, B_{n-1}^q , (with corresponding minimized energy v_{n-1}^q) determines the global optimality of B_n . One can further formulate another recursive formula to compute B_{n-1}^q and v_{n-1}^q using paths of length $n-2$, etc. The resulting general recursive formula is below:

$$B_{i+1}^r = \operatorname{argmin}_{B_i^s \cup r} \{v_i^s + |I(r) - I(s)| * \mu(r, s) + I(r)\}, \quad (4)$$

where i ranges from 1 to $n-2$. The base cases B_1^r 's and v_1^r 's are trivially defined as the points in the first column and their corresponding intensity values, respectively. If all intermediate values v_i^s 's evaluated along the way are saved in points indexed by i and s , the optimization dictated by Eqs. (3, 4) can be conducted by populating the m by n table with v_i^s 's from column 1 to column n , in a bottom-up manner. A backtracking step in reverse direction on this table recovers the B_{n-1}^q and thus B_n .

Property 3 of the boundary pixels states that the microvessel boundaries should present themselves near the upper and lower horizontal boundaries of the photomicrograph. To encode this property, the objective function in the DP procedure should reward paths that are close to two horizontal boundaries, and penalize otherwise. For example, to search for the lower boundary, instead of using only the gray-scale intensities $I(L[i])$ as the self-energy of each path pixel, the lower horizontal side of the photomicrograph is treated as the horizon of the Earth. Consequently each pixel within the photomicrograph has a “weight” term $W(L[i])$ induced by this “gravity-field”:

$$W(L[i]) = c * y(L[i]), \quad (5)$$

where c is a constant simulating the gravity constant. The value of c is highly dependent on the imaging processing and noise levels in the microphotograph, which is selected after an offline training (Section III.F) or on-line readjustments on the photomicrograph. $y(L[i])$ is the y component of $L[i]$, i.e., the height or vertical distance from the lower side of the photomicrograph. Eq. (5) mimics the gravity energy $m * g * h$ in mechanics, where $m=1$, $g=c$. The resulting optimality function is rewritten as,

$$B_n = \operatorname{argmin}_{B_{n-1}^q \cup p} \{v_{n-1}^q + |I(p) - I(q)| * \mu(p, q) + I(p) + W(p)\}. \quad (6)$$

The general recursive equation Eq. (4) can be similarly rewritten. This additional weight due to the induced gravity field drags the smooth shortest path from the center of the microvessel in the middle panel of Fig. 1 toward the authentic boundary. The same method was used to detect the upper boundary except that the photomicrograph was flipped upside down before

performing the DP procedure based on Eq. (6). By automatically reproducing human operator's process, this gravity-field DP procedure identified the two boundaries of the microvessel, as depicted on the right panel of Fig. 1. The upper and lower boundaries are colored by green and blue for ease of visualization. In contrast, without inducing the gravity field, the "boundary" detected by the DP procedure dictated by Eq. (4), shown in Fig. 1(b), is not correct. A straightforward image pixel counting procedure evaluates the precise area in-between these two boundaries, which yields significantly better estimates of microvessel areas than a human operator's measurements. The formulation and development of this algorithm goes along the same line as the content-aware image resizing in computational photography [1] where DP is the main workhorse in search of the global minimum [8].

II. THE CELL ADHESION LOCATION CLASSIFICATION AND QUANTIFICATION ALGORITHM

After automatically singling out the microvessel from the photomicrograph, we classified the adherent tumor cells/leukocytes locations as on one cell, at the joints of two, three and four endothelial cells forming the microvessel wall. In order to account for those tumor cells/leukocytes adhering at the microvessel boundaries, the microvessel region was extended by about 5-pixel beyond each boundary.

A. PDE-based anisotropic image de-noising—The first step in any visual object detection procedure is image cleaning or de-noising [6], otherwise many false positives and negatives of endothelial borders and tumor cells/leukocytes will induce false results. To clean up photomicrographs without unduly compromising strong edges and compact regions of endothelial cell borders and tumors/leukocytes, a de-noising step other than conventional Gaussian smoothing is necessitated. The anisotropic smoothing can satisfy the needs. This approach suppress smoothing operations along the normal direction of edges and object boundaries. The corresponding controlling PDE is given below:

$$I_t = \text{div}(\alpha(|\nabla I|)\nabla I), \quad (7)$$

where div is the divergence operator; ∇I is the gradient of image I ; and $\alpha()$ is a decreasing function, of which a typical choice is defined in the following form:

$$\alpha(z) = \frac{1}{1+k^2 z^2}, \quad (8)$$

where k is a controlling constant, ordinarily ranging from 0.01 to 0.1, used to decide the magnitude of smoothing. In a region of weak high frequency energies, $|\nabla I|$ is small and makes Eq. (7) a Gaussian diffusion. In contrast, in regions with large $|\nabla I|$, i.e., those close to endothelial cell borders and/or tumor cells/leukocytes, $\alpha(|\nabla I|) \approx 0$, thus no smoothing is conducted. Therefore the selective, or anisotropic, smoothing is achieved. In consequence, the valuable endothelial cell borders and/or tumor cells/leukocytes boundaries are preserved after this anisotropic smoothing step.

B. Adaptive threshold to tackle uneven illuminations—To classify and quantify tumor cells/leukocytes and their adhesion locations, the gray-scale image needs to be

converted to a binary one to apply mathematical morphological operations. In the photomicrographs measured in terms of micrometers, the illuminations received by different regions are uneven. In Fig. 2, the mid-left region of the photomicrograph is brighter than other regions, especially the right half. The binary (or logical) image generated by the original Otsu's threshold is demonstrated on the middle panel. The weaker illumination on the right side of the photomicrograph makes many background regions falsely classified as foreground ones (dark ones). This makes it extremely difficult to classify and quantify border geometry/topology of tumor cells/leukocytes and endothelial cells. The adaptive thresholding approach works by making a more humble assumption in determining the threshold: the illumination is assumed constant only in a small window. A pixel p is labeled as foreground only if its value is larger than the statistics γ of a local window w_p centered at p , in this work γ is chosen to be the mean value of w_p minus the window size l of w_p [5]. The binary image produced by the adaptive thresholding procedure is shown on the right panel of Fig. 2, where the varying illuminations present in the original photomicrograph are effectively removed.

C. Adherent tumor cell/leukocyte detection—Most mathematical morphological operators such as component labeling, image erosion, dilation, closing, area opening and watershed [5] are the workhorses in detecting adherent tumor cells/leukocytes from the binary (logical) image produced by the preceding adaptive thresholding step.

The right panel of Fig. 2 shows that two outstanding heuristics can distinguish adherent tumor cells/leukocytes from other foreground regions (in dark), which survive the adaptive thresholding:

1. Unlike noises that happen to survive the adaptive thresholding step by chance, adherent tumor cells/leukocytes have areas significantly larger than 0.
2. Unlike endothelial borders that are ordinarily long and thin, the tumor cells/leukocytes are more like disks.

To materialize the first heuristic, an image closing operator using a disk-shaped structural element [10] is performed to remove foreground regions with the area less than a prescribed threshold δ_1 . To realize the second heuristic, an image erosion operator using a disk-shaped structural element is conducted to remove foreground pixels whose local neighbors are usually smaller than this element.

Besides the desirable clear individual tumor cells/leukocytes, two other possible regions may pass the preceding two steps: a) two or more connected tumor cells/leukocytes, or b) noisy regions or cell borders of thick width that cannot be removed entirely by the erosion procedure. The water-shed procedure [5] can handle large patches for case a) transform the binary blobs to topological reliefs, and then compute the water-shed lines, which separate boundaries of the formerly connected tumor cells/leukocytes. Heuristic 2 can address the trouble posed by case b) the sought-after tumors/leukocytes are approximately like disks whereas those thick noisy patches and thick cell borders have extremely elongated shapes. This can be identified by eigenanalysis: first conduct singular value decomposition (SVD) [9] for each patch $\mathbf{M}$:

$$\mathbf{M} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T, \quad (9)$$

where $\mathbf{U}$ and $\mathbf{V}^T$ are the left and right matrices of eigenvectors, respectively; and the diagonal matrix $\mathbf{\Sigma}$ has two eigenvalues λ_1 and λ_2 , which are the magnitudes of the major and minor axes of the covering ellipse. Heuristic 2 demands that for authentic tumors or leukocytes the ratio of λ_1 over λ_2 should be close to 1. Conversely for an elongated shape, this ratio will be far larger than 1. In the experiments, the threshold is set at 2: a blob is declared a tumor cell or leukocyte if $\lambda_1/\lambda_2 < 2$, removed as non-tumor cell/leukocyte otherwise.

D. Tumor cell/leukocyte adhesion location classification by local analysis—

The last step is to designate the adhesion location of tumor cells or leukocytes: on one cell, or at the joints of two, three or four endothelial cells forming the vessel wall. A novel *double component labeling procedure* is formulated toward this end. From the binary image obtained from Subsection III.B, this procedure analyzes each tumor cell/leukocyte identified by the preceding step to decide the number of endothelial cell borders, if any, that touch a tumor/leukocyte. For each tumor cell/leukocyte T , adhesion location is analyzed on a square window W_T centered by the centroid of T of side $5 \cdot \lambda_1$, λ_1 being the estimated major of T from the SVD procedure Eq. (9).

First, a component labeling procedure [5] is conducted for all pixels inside W_T . The component labeling procedure merges connected pixels into one region based on a neighborhood system, where the 8-neighborhood system is chosen in the experiments. Next, all foreground pixels in components different from T are removed because they do not touch T at all. Keep in mind most foreground pixels connected to tumor cells/leukocytes are endothelial cell borders except some noises, thus they are the basis to further decide the adhesion location of tumor cell/leukocyte T .

To decide the number of endothelial cell borders that touch T , the region corresponding to tumor cell/leukocyte T is removed from the image resulted from the preceding procedure, which likely separates the originally connected components. To determine the number of endothelial cell borders, a second application of the component labeling procedure is needed. The number of connected components from this round of labels finally decides the number of endothelial cells where the tumor/leukocyte T is located. To avoid the negative impacts of noises, before counting the final number of endothelial cell borders, all isolated components with the size less than 5 are removed by the image closing operator. T 's locations are declared according to the following rule:

1. T is designated as *on one endothelial cell* if the number of components is zero or one, that is, after removing T , all components are of negligible size (< 5).
2. T is labeled as *at the joints of two endothelial cells* if the number of connected components is two, that is, T cuts through one endothelial cell border thus T 's removal yields two connected components.

3. T is labeled as *at the joints of three and four endothelial cells*, respectively, if the number of components is three and four, correspondingly, given that in total three/four endothelial cell borders are identified after removing T .

In Fig. 3, the step-by-step results generated by the double component labeling procedure are illustrated. Panel (a) is the local window W_T , the central compact blob is tumor cell/leukocyte T detected by the approach detailed in Subsection III.C. The first application of the component labeling procedure removes foreground regions (dark ones) that are disconnected from T , hence the top large component and the small one in the lower left corner disappear, resulting in the single component in panel (b). Panel (c) presents the picture after removing T . The second application of the component labeling procedure produces three large components for regions on panel (d). The small component in the upper right part is removed as noises. According to the adhesion location labeling rule, tumor cell/leukocyte T is thus designated at the joints of three cells, which is correct for this specific adherent tumor cell/leukocyte.

E. The proposed cell adhesion location classification and quantification

algorithm—The proposed algorithm for a given input photomicrograph I of height m and width n is summarized below.

Algorithm cell adhesion location classification¹

- Step 1** Call the gravity-field DP procedure (Section I) for input photomicrograph I based on Eqs. (5, 6) to generate microvessel mask $\mathbf{M}$, and $\mathbf{J}=\mathbf{M}' * \mathbf{I}$, where $\mathbf{M}'$ is an expanded $\mathbf{M}$ with an additional narrow band added on lower and upper ends of $\mathbf{M}$.
- Step 2** Conduct anisotropic smoothing (Subsection II.A) for $\mathbf{J}$ using Eqs. (7, 8) to get a smoothed image $\mathbf{K}$.
- Step 3** Perform adaptive thresholding (Subsection II.B) on $\mathbf{K}$ to get a binary image $\mathbf{L}$.
- Step 4** Detect tumor cells/leukocytes from $\mathbf{L}$ using mathematical morphological operators and SVD procedure as detailed in Subsection II.C.
- Step 5** Designate tumor cell/leukocyte adhesion locations after the double component labeling local analysis procedure as developed in Subsection II.D.
- Step 6** Compute the microvessel area from mask $\mathbf{M}$ in Step 1, and evaluate the cell adhesion location classification statistics based on locations assigned in Step 5.

In Step 1, the gravity-field DP procedure is in $O(m*n)$. The anisotropic smoothing step again is in linear order with the number $m*n$ of pixels. Adaptive thresholding in Step 3 applies for each pixel in its local constant-size windows, hence also in the order of $O(m*n)$. Steps 4, 5 and 6 analyze local neighborhoods of constant sizes for each tumor cell/leukocyte,

¹All steps needed to implement the proposed algorithm are given here. Most functions can be found in MATLAB's image processing toolbox, such as Otsu's thresholding function `graythresh`, morphological functions `imclose`, `imopen` and `watershed`.

thus again invoking at most $O(m*n)$ computations. The endothelial cell borders and adherent tumor cells at the not well-focused layer of the vessel have smaller intensity (or degree of darkness) than those at the well-focused layer, our adaptive thresholding step and the ensuing image opening and size filtering can effectively reduce the false counting problem due to this cause.

Therefore, the proposed algorithm in total has an $O(m*n)$ time complexity, namely, a $k*m*n$ time is needed, k being a certain constant. However, because of the several iterations of Step 2 and many local windows in Step 3, k —heavily depending on the number of detected adherent tumor cells/ leukocytes—is of considerable magnitude. The proposed algorithm thus cannot deliver real-time performances. To process a photomicrograph of dimension 354 by 709 with more than 200 tumor cells or leukocytes, the algorithm implemented in MATLAB spends 10.3 seconds on a Dell Precision M6600 laptop.

F. Parameter learning from sample microphotographs—In this algorithm, the classification performances rely on the optimal choices of several crucial parameters: the gravity constant c in Eq. (5), controlling constant k in Eq. (9), window size l of w_p in Section II.B, and area threshold δ_1 in II.C. These parameters cannot be fixed for all microphotographs or selected entirely in a heuristic manner. Conversely, they should be tuned or adapted under different imaging settings such as different illuminations, sample sizes and image resolutions. This desirable adaptivity was achieved in this work by an offline parameter searching procedure for sample microphotographs using ground truths, a typical hyper-parameter optimization process in machine learning: from hundreds of microphotographs taken under the same imaging conditions, a typical one (or two if the number of similar microphotographs is more than 100, the second serves as the cross validation data) is chosen randomly as the training data from this set of microphotographs. Human operators evaluated the ground truths used in the ensuing parameter learning procedure. Two different approaches were employed based on the different natures of the parameters in question.

1. For the parameters unrelated to others, the linear or sequential parameter searching is applied to make the optimal choice. As an illustration, to find the best c , call the gravity-based DP procedure by looping for c ranging from a small value, e.g., 0.1, to a large value, e.g., 2.5, with step size 0.1. The c that induces the smallest error from the two known boundaries in the ground truth is the selected one. If the possible ranges are large (cases we have not encountered), a hierarchical search can be invoked: first conduct a loop for c goes from 0 to 100 (for example) by step size 10; if the smallest error occurs at 30, then another loop sets c from 25 to 35 with step size 1. Continuing in this manner, the c value for any possible precision can be selected. Since each call of the algorithm costs about 10 seconds, it takes several minutes to find the optimal value for one parameter.
2. If it is suspected that some parameters are correlated, then a grid search is conducted, where the grid is formed by the Cartesian product of all correlated parameters. The optimal choice is the point in this parameter space that yields the

best performance judged by the ground truth. The grid search takes longer time than the linear search.

After this offline parameter learning procedure for sample microphotographs (for the photomicrograph shown in Fig.4, $c=0.6$, $k=\frac{1}{30}$, $l=7$, and $\delta_l=30$, all using approach 1), the learned parameters then plug in the classification algorithm conducted on other microphotographs with similar imaging conditions.

Results

A. Classification and quantification of tumor cell adhesion location

A.1 Determination of the ground-truth—To evaluate the performance of the proposed algorithm, human experts need to first generate the ground truth. To do this, three human experts were asked to independently evaluate one microphotograph. Each expert evaluated the same microphotograph 3 times at different days without cross-reference to the previous results. Each expert gave his/her own average classification results. The ground truth was the average of these 3 individual averages after a discussion among human experts. We believe that the ground truth determined in this way may considerably reduce the human bias. Table I demonstrates the tumor cell adhesion location classification results produced by these 3 experts for the same microphotograph shown in Fig. 4, where the designations of true positive, false positive and false negative are based on the ground truth.

Table I indicates the interpersonal variation due to different human operators. Even for the same operator, the classifications reported in different days are different. Table II tabulates the quantification results by one expert (Expert 1) reported in 3 independent evaluations. Table II shows the intrapersonal variation, which is less than the interpersonal variation. However, both intra- and inter-personal variations are inevitable due to human subjectivity.

To gauge the goodness of a classifier, three classifier quality measures are popularly used: a) Precision: the probability of classified instances that are correct. b) Recall: the probability of correct instances that are identified by the classifier. c) F_1 score: the harmonic mean of precision and recall.

The average values for the corresponding precision, recall and F_1 scores for the 3 evaluators are shown in Table III. In the Expert 1 column, the individual values for the 3 evaluations are also listed in the parenthesis.

A.2. Classification and quantification results by the algorithm—Fig. 4 illustrates the cell adhesion location classification results by the algorithm, the microvessel (outside regions are colored in cyan for better visualization) generated by the gravity-field DP procedure that closely follows the ground truth measured by human experts. The area due to the gravity-field DP procedure is 99.03% of the ground truth, which only misclassified a small cell in the middle right of the upper boundary. In contrast, the ratio received from crude tape measuring by averaging 3 human operators' evaluations are 105.5%, 98.8% and 95.6%, respectively.

For the photomicrograph shown in Fig. 4, the algorithm detects that there are 1 tumor cell on one endothelial cell, 151, 48, 25 tumor cells at the joints of 2, 3 and 4 endothelial cells, respectively. These statistical quantification results closely approximate those produced by human operators. Table IV demonstrates the results compared with the ground truth. The corresponding precision, recall and F_1 score are 0.92, 0.85 and 0.89, respectively, which are indistinguishable from those attained by the 3 human experts, as listed in Table III.

From Tables I—IV, the quantification and classification results delivered by the proposed algorithm are slightly worse than those by the best human expert, they are better than, or at least comparable to those by one human expert (Expert 2). Tables I, II and III demonstrate that the intra- and inter- personal variations are inevitable when human operators perform the classification for the same task. However, those errors due to human subjectivity can be completely avoided if we use the proposed algorithm.

In addition, the entire classification procedure takes the laptop (specified in the end of Section III) merely 9.16 seconds. In such a short time, human operators cannot single out these 238 tumor cells from a noisy and unevenly lighted photomicrograph, letting alone categorize their adhesion locations. The same classification process for Fig. 4 took the three human experts about one and half hours on average.

A.3. Classification and quantification results for multiple test

microphotographs—To further examine the performance of the proposed algorithm, five experiments were conducted on five microvessels that were taken under different imaging conditions. For each microvessel, 4 photomicrographs were collected [4] [2], generating a total of 20 photomicrographs. Fig. 5 summarizes the corresponding tumor cell adhesion location classification and quantification results for these 20 photomicrographs.

Fig. 5a compares the area ratios computed by the proposed algorithm and those by 3 human experts. The human-based measures, the right three, have much larger fluctuations compared to that by the algorithm, which is tightly clustered around 99%.

In Fig. 5b, boxplots for precisions, recalls and F_1 scores, respectively, summarize the tumor cell adhesion location classification results produced by the algorithm and 3 experts. The classification results delivered by the algorithm is comparable with human experts—well within the subjective error margin due to human subjectivity. Hence for all these 20 cases the proposed algorithm consistently attains desirable results. The time consumed by each of the 20 classifications and quantifications ranges from 3.6 to 11.0 seconds. While for human experts, the average time ranges from 0.5 hour to close to 2 hours.

B. Classification and quantification of leukocyte adhesion location

The proposed algorithm was also applied to the photomicrograph showing endothelial cell boundaries and adherent leukocytes from a segment of a single perfused venular microvessel in the rat mesentery [14]. The figure in [14] was first converted to the gray-scale format. Because many leukocytes in their original micrographs are transparent, to fit in the algorithm, a Hough transform based circle searching algorithm locates and fills in the apparently small adherent leukocytes. The left panel of Fig. 6 shows the resultant gray-scale

image. For this photomicrograph of size 574 by 998 pixels, by the same laptop, the time consumed is 8.9 seconds, which is slightly less than the preceding tumor adhesion case. This is because in this photomicrograph, fewer adherent leukocytes are present, namely, 71 vs. 238. Specifically, there are 0 adherent leukocytes on one endothelial cell (red), 33 adherent leukocytes at joints of two cells (green), 14 and 24 adherent leukocytes at joints of 3 (blue) and 4 cells (yellow), respectively. The ground truth for this microphotograph is derived by the same method described in IV.A.1. The average processing time taken by the 3 human experts on this microphotograph is around 35 minutes.

The following observations from Fig. 6 are in order:

1. The microvessel boundaries detected by the gravity-field DP procedure partially reproduce the ground truth provided by [14], but by a small margin. The upper boundary is perfect, whereas in the lower boundary two segments around the mid region are missed. Because of some relatively larger magnitudes in both designated boundary segments, the gravity-field DP procedure takes the routes one cell upper to obtain the shortest distance between the two sides. The area computed by the proposed procedure is 97.24% of the ground truth. In contrast the region by human tape measuring is 104.8%.
2. The adhesion location classification results are qualitatively reasonable by eyeballing. Classifying large transparent leukocytes present in the original photomicrograph lies beyond the range of this algorithm because current features confuse boundaries of the large leukocytes with authentic endothelial cell borders thus causing many false positives.

Table V shows the quantitative location classification results generated by the proposed algorithm (the first number in each cell) and the three experts (the trailing three numbers in each cell). The corresponding precision, recall and F_1 scores are shown in Table VI.

These results are slightly worse than those in Table IV for Fig. 4, especially the recall value. This indicates that the proposed algorithm missed relatively more true leukocytes for the reason given above. The quality as shown in Table VI of the proposed algorithm is slightly worse than human experts but still comparable.

The above results demonstrate that the proposed algorithm achieved encouraging performances in adhesion location classification and quantification for both tumor cells and leukocytes from different experiments.

To determine whether tumor cells prefer to adhere at the endothelial junctions rather than randomly adhere on the vessel surface, using the quantification approach as devised in [14] and [2], we first measured the length of endothelial junctions per unit area for the microvessel shown in Fig. 4, the averaged value is $107.9 \pm 14.4 \text{ mm/mm}^2$. This junction length is then expanded by half of the contact length (diameter) of an adherent cell, $3.1 \mu\text{m}$, the band area is thus $33.4\% \pm 4.5\%$ of the total area. If the cells just randomly adhered, the percentage of the cells adhered to junctions should be close to this value. However, the percentage of our classification results is 99.5% on the junctions as reported in Table 1. For the 20 microphotographs tested in Fig. 5, the average percentage of the band area is $36.2\% \pm$

3.7%, while the minimal percentage of cells adhered to junctions is 93.8%. For the photomicrograph showing endothelial cell boundaries and adherent leukocytes in Fig. 6, the computed band area is $37.9\% \pm 2.4\%$, whereas the percentage of leukocytes adhered to junctions is $>90\%$, similar to that originally reported in [14]. The statistical supports for the preferential adherence of tumor cells and leukocytes to endothelial junctions from all our test cases are overwhelming and thus significantly distinguishable from a random distribution.

Discussions

In this paper we developed a new classification and quantification algorithm of cell adhesion locations on the endothelium. As demonstrated, within seconds, a laptop can satisfactorily perform the adhesion location classification and quantification from the noisy photomicrographs with tens to hundreds adherent tumor cells or leukocytes. Since the proposed algorithm always delivers the same classification and quantification results for the same task, it can avoid subjective errors that are inevitable by human experts. Our automatic classification and quantification method provides a reliable and cost and time efficient approach for biomedical image processing. It can also be adapted to other types of biomedical images for a broader application, especially for images containing more complicated features. In this paper, only still microphotographs were employed, equipped with algorithms developed in this paper and our previous work on video processing and analysis [9], in the near future we will further process and analyze the time-lapsed images of cell adhesion, crawling, or transmigration dynamics in the batch mode to gain deeper insights into the dynamic, kinetic and evolutionary properties of endothelial cells and tumor cells/leukocytes during tumor metastasis or leukocyte movement. So far more applications have been found by use of the algorithmic ideas developed here, such as lung cancer segmentation and motion tracking [12] for 4-dimensional computer tomography (4DCT), a sequence of time-lapsed 3DCT of human lungs within a breathing period, more work will be reported in the near future.

In addition, our classification and quantification algorithm can be applied to quantify the cell alignment under flow and calculate the aspect ratio of the original and deformed cell shape, to classify different cell types in brain images including neurons and various glial cells, and to identify the microvascular pattern in different tissues as well as under physiological and pathological conditions for the same tissue, to name a few.

Acknowledgments

This work was supported in part by NIH U54CA132378-06, NIH U54CA137788-01, NIH CA153325-01, NSF CBET-0754158 and City College of New York Seed grant 2013.

References

1. Avidan S, Shamir A. Seam carving for content-aware image resizing. *ACM Trans. on Graphics*. 2007; 26(3) No. 10.
2. Cai B, Fan J, Zeng M, Zhang L, Fu BM. Adhesion of malignant mammary tumor cell MDA-MB-231 to microvessel wall increases microvascular permeability via degradation of endothelial surface glycocalyx. *J. of Appl. Physiol*. 2012; 113(7):1141–1153. [PubMed: 22858626]

3. Chambers AF, Groom AC, MacDonald IC. Dissemination and growth of cancer cells in metastatic sites. *Nat Rev Cancer*. 2002; 2(8):563–572. [PubMed: 12154349]
4. Fan J, Cai B, Zeng M, Hao Y, Giancotti F, Fu BM. Integrin beta-4 signaling promotes mammary tumor cell adhesion to brain microvascular endothelium by inducing ErbB2-mediated secretion of VEGF. *Annals of Biomedical Engineering*. 2011; 39(8):2223–2241. [PubMed: 21556948]
5. Gonzalez, RC.; Woods, RE.; Eddins, SL. Digital image processing using MATLAB. 2nd edition. Gatesmark Publishing; 2009.
6. Szeliski, R. Computer Vision: Algorithms and Applications. Springer; 2011.
7. Talmadge JE, Fidler IJ. AACR centennial series: the biology of cancer metastasis: historical perspective. *Cancer Res*. 2010; 70(14):5649–5669. [PubMed: 20610625]
8. Wei J. Markov edit distance. *IEEE Trans. on Pat. Ana. Mach. Intel*. 2004; 26(3):311–321.
9. Wei J. Shape indexing and recognition based on regional analysis. *IEEE Trans. on Multimedia*. 2007; 9(5):1049–1061.
10. Wei J. Small moving object detection from video sequences. *Int. J. of Image and Graphics*. 2013; 14(3)
11. Weiss L. Comments on hematogenous metastatic patterns in humans as revealed by autopsy. *Clin. Exp. Metastasis*. 1992; 10(3):191–199. [PubMed: 1582089]
12. Wei J, Yuan A, Li G. An automatic toolkit for efficient and robust analysis of 4D respiratory motion. AAPM 2014 (Best in Physics award). 2014
13. Yan WW, Cai B, Liu Y, Fu BM. Effects of wall shear stress and its gradient on tumor cell adhesion in curved microvessels. *Biomech. Model Mechanobiol*. 2012; 11(5):641–653. [PubMed: 21818636]
14. Zeng M, Zhang H, Lowell C, He PN. Tumor necrosis factor-alpha-induced leukocyte adhesion and microvessel permeability. *Am. J. of Physiol*. 2002; 283(6):H2420–H2430.

Fig. 1. Microvessel boundary segmentation using DP procedure with or without the introduction of gravity field. (a) original photomicrograph; (b) shortest path according to original dynamic programming procedure; (c) two detected boundaries by the proposed gravity-field dynamic programming method.

Fig. 2. Image thresholding. (a) Original photomicrograph; (b) binary image according to Otsu's method; (c) binary image according to adaptive thresholding method

Fig. 3. Illustration of the double component labeling procedure for cell adhesion location classification. (a) initial binary image; (b) results after the first component labeling procedure; (c) results after removing detected cell T; (d) results after the second application of component labeling algorithm.

Fig. 4.
Tumor cell adhesion location classification: red—on one cell, green—at joints of two cells,
blue—at the joints of three cells, yellow—at the joints of four cells.

(a) Area ratios

(b) Summary of quality measures

Fig. 5.

Comparison of the vessel area measurements and tumor cell adhesion location classification results by the algorithm (blue) and by the three experts (green, yellow, and cyan) for 20 segments obtained from the experiment. (a) Vessel area ratios by the algorithm and three experts; (b) Summary of quality measures: precisions, recalls and F_1 scores by the algorithm and three experts.

(a)

(b)

Fig. 6. Leukocyte adhesion location classification from a photomicrograph in [14]. (a) gray-scale image of the leukocyte micro-photograph; (b) location classification results.

Table I

Classification results by the three experts against the ground truth for Fig. 4.

Type	On one Cell	At joints of 2 Cells	At joints of 3 Cells	At joints of 4 Cells
True Positive	1, 1, 1	148, 139, 145	50, 44, 47	26, 25, 26
False Positive	0, 0, 0	7, 9, 12	5, 9, 8	5, 7, 3
False Negative	0, 0, 0	14, 23, 17	3, 9, 6	3, 4, 3

Table II

Three quantification results by the same expert against the average for Fig. 4.

Type	On one Cell	At joints of 2 Cells	At joints of 3 Cells	At joints of 4 Cells
True Positive	1, 1, 1	150, 145, 149	51, 51, 48	25, 28, 25
False Positive	0, 0, 0	6, 10, 5	3, 6, 6	3, 6, 5
False Negative	0, 0, 0	12, 17, 13	3, 3, 6	4, 1, 4

Table III

Precision, recall and F_1 scores for the three experts as shown in Tables I and II.

Type	Expert1	Expert2	Expert3
Precision	0.93 (0.95, 0.91, 0.93)	0.89	0.91
Recall	0.92 (0.92, 0.91, 0.91)	0.85	0.89
F_1 score	0.92 (0.94, 0.91, 0.92)	0.87	0.90

Table IV

Classification and quantification results by the proposed algorithm against the ground truth for Fig. 4.

Type	On one Cell	At joints of 2 Cells	At joints of 3 Cells	At joints of 4 Cells
True Positive	0	144	43	21
False Positive	1	7	5	4
False Negative	0	18	11	8

Table V

The proposed algorithm quantification results by the algorithm and three experts for Fig. 6.

Type	On one Cell	At joints of 2 Cells	At joints of 3 Cells	At joints of 4 Cells
True Positive	0,1,1,1	28,30,32,29	13,14,11,15	21,24,22,22
False Positive	0,0,0,0	5,7,3,3	1,0,4,1	2,1,3,4
False Negative	1,0,0,0	6,4,2,5	4,3,6,2	4,1,3,3

Table VI

Precision, recall and F_1 score of the algorithm and the three experts whose average is the ground truth for Fig. 6.

Type	Proposed algorithm	Expert 1	Expert 2	Expert3
Precision	0.89	0.90	0.87	0.89
Recall	0.81	0.90	0.86	0.87
F_1 score	0.84	0.90	0.86	0.88